

A Design for a Primary Personal Information Market

Ross Farrelly: BSc (Hons), M. App. Stat., M. App. Ethics, Grad. Dip. Ed.

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy (Information Systems)

Faculty of Design, Architecture and Building

University of Technology Sydney

30 June 2019

Abstract

As mobile and wearable technologies grow in popularity, ever-increasing volumes of fine-grained personal information are generated as people go about their daily lives. There is a growing consensus that these data are of great value to corporations for a variety of uses and there is a flourishing worldwide secondary market for personal information. From the individual's perspective, this personal information may be exchanged for 'free' services, but there are currently no widely adopted means by which individuals can benefit financially from their personal information. This research addresses the question as to why this is the case.

Although there is a wealth of multi-disciplinary research into related areas such as personal informatics, infomediaries and value co-creation, there is very little research in how a Primary Personal Information Market (PPIM) could operate for the benefit of both individuals and companies. In particular, the application of the integrated Service Innovation Method (*iSIM*), industry platform theory and digital platform theory to the design of a PPIM has not yet been undertaken in the literature.

This presents an opportunity for the development of a novel artefact that would make a contribution to both the theoretical aspects of platform design and also be of practical benefit to those individuals, businesses and entrepreneurs who are looking to monetise personal information.

Drawing on design science methodology, market engineering theory, service innovation and digital platform theories, this dissertation presents a novel, detailed design of an innovative digital platform, which would collect and monetise personal information on behalf of the original creators of that information without compromising privacy or security. It presents a justification for the creation of a Primary Personal Information Market (PPIM) as a means of addressing the problem of personal information monetisation and outlines a theoretical model of a PPIM and the IT infrastructure necessary to support it. This dissertation also presents a critical Service Architecture for the PPIM digital platform to ensure scalability and potential commercialisation.

In line with Design Science Research Methodology, the structure of this dissertation follows the design evolution of a PPIM digital platform based on academic experts' iterative feedback in stages over three years. Commencing with an initial design of a two-sided market and the transaction object, which would be traded on such a market, the design of a PPIM then evolves to a four-sided industry platform and culminates in the detailed design for a four-sided digital platform. The PPIM Service Architecture Framework was also developed. A prototype PPIM is described and its viability from the perspective of the digital platform's four categories of market participants is assessed.

The primary theoretical contribution of this dissertation is to develop a theory for design and action (following the classification scheme of Gregor). It applies the integrated Service Innovation Method

(iSIM) to the design of a novel digital platform. Following iSIM, this dissertation outlines a schematic description of a novel four-sided PPIM design and the associated PPIM Service Architecture and end-to-end operational models to facilitate a scalable implementation of a PPIM. The schematic representation of the four-sided PPIM design and associated Service Architecture and end-to-end operational models are described at an abstract level without recourse to a specific technology.

This dissertation also makes practical (managerial) implications for platform owners and for all four participant types on a four-sided PPIM. Business owners and senior executives in industries which are undergoing “platformisation” will benefit from considering, in the light of this research, how to monetise the personal information they collect. Individuals with a propensity to monetise their personal information will benefit from understanding how this can be facilitated while maintaining privacy and security and model developers and data feed developers will benefit from gaining an insight into how to generate a recurring income stream from personal information.

List of Figures

Figure 1: DSR Knowledge Contribution Framework (Gregor & Hevner 2013)	31
Figure 2: Design and design science process elements from IS other disciplines and synthesis objectives for a design science research process in IS (Peppers et al. 2006)	41
Figure 3: Iterations and atomisation steps in the design process	43
Figure 4: Data contributing individual initiated transactions.....	76
Figure 5: End user initiated transactions	76
Figure 6: The IT infrastructure to support a PPIM	78
Figure 7: A Simplified PITO	87
Figure 8: Creation and exchange of PITO on a PPIM.....	87
Figure 9: Permissions logical data model	98
Figure 10: About screen from survey	99
Figure 11: Scenario screen from survey.....	99
Figure 12: Sign in screen from survey	100
Figure 13: Screen shot 1 of example permissions as presented to respondents.....	100
Figure 14: Screen shot 2 of example permissions as presented to respondents.....	101
Figure 15: Finish and request to share page from survey	101
Figure 16: Four participant types and value propositions on a PPIM.....	110
Figure 17: Provider, user and joint spheres of value creation on a PPIM.....	112
Figure 18: The PPIM service system design.....	114
Figure 19: Service experience blueprint for participation on a PPIM	115
Figure 20: Screen shot of prototype end user facing dashboard	116
Figure 21: Home page of simulated PPIM.....	117
Figure 22: Browse page of simulated PPIM	118
Figure 23: Manage page of simulated PPIM	118
Figure 24: First part of create targeted marketing PDP of simulated PPIM	119
Figure 25: Second part of create targeted marketing PDP of simulated PPIM.....	120
Figure 26: Preview page of targeted marketing PDP on simulated PPIM	120
Figure 27: View and renew page of targeted marketing PDP on simulated PPIM	121
Figure 28: View and execute page of targeted marketing PDP on simulated PPIM.....	121
Figure 29: Success notification page targeted marketing PDP on simulated PPIM	122
Figure 30: First part of create personalised service response PDP of simulated PPIM.....	122
Figure 31: Second part of create personalised service response PDP of simulated PPIM	123
Figure 32: Third part of create personalised service response PDP of simulated PPIM	123
Figure 33: Preview of personalised service response PDP of simulated PPIM.....	124

Figure 34: Screen display when executing personalised service response PDP of simulated PPIM..	124
Figure 35: Model developer develops new model	138
Figure 36: Model developer checks on usage of his or her models and reconciles against payments	138
Figure 37: End user creates PDP from template	139
Figure 38: Platform owner receives notification of disputed payment from a model developer and resolves the dispute	139
Figure 39: Data contributing individual installs data collecting app, selects data feeds to be collected and receives requested for his or her data to be include in PDPs	139
Figure 40: Hierarchy of API implementation.	140
Figure 41: The value chain for the creation of a new PDP	141
Figure 42: The value chain for the creation of a new model	142
Figure 43: The value chain for the creation of a new data feed.....	142
Figure 44: The Bathtub model of innovation as applied to a PPIM, adapted from (Tiwana 2013, p. 204)	144
Figure 45: Service architecture for a PPIM	146
Figure 46: Create model scenario executed on a service architecture for a PPIM	147
Figure 47: A non-predictive personal information app	150
Figure 48: A predictive personal information app	151
Figure 49: Workflow to construct and execute a targeted marketing personal data product	153
Figure 50: Workflow to construct and execute a personalised service response personal data product	155
Figure 51: iSIM interrelationship map (Chew 2015).....	168
Figure 52: Model Developer Portal – login screen	169
Figure 53: Model Developer Portal – create account screen	169
Figure 54: Model Developer Portal – home screen	170
Figure 55: Model Developer Portal – sources screen	170
56: Model Developer Portal – SQL screen	171
57: Model Developer Portal – Jupyter screen	171
58: Model Developer Portal – financials screen	172
59: Model Developer Portal – usage screen.....	172
60: Data Feed Developer Portal – login screen.....	173
61: Data Feed Developer Portal – home screen	173
62: Data Feed Developer Portal – eclipse screen.....	174
63: Data Feed Developer Portal – financials screen	174
64: Data Feed Developer Portal – usage screen.....	175
Figure 65: Data Contributing Individual – login screen	175

Figure 66: Data Contributing Individual – manage data feeds screen	176
Figure 67: Data Contributing Individual – manage PDPs screen	176
Figure 68: Data Contributing Individual – example PDP screen	177
Figure 69: Data Contributing Individual – new PDP invitations screen	177
Figure 70: Data Contributing Individual – example PDP invitation screen	178
Figure 71: Data Contributing Individual – financials screen	178
Figure 72: Platform Owner Market Management Portal – login screen	179
Figure 73: Platform Owner Market Management Portal – data contributing individuals screen	179
Figure 74: Platform Owner Market Management Portal – example data contributing individual screen	180
Figure 75: Platform Owner Market Management Portal – example model developer screen	180
Figure 76: Platform Owner Market Management Portal – data feed developers screen	181
Figure 77: Platform Owner Market Management Portal – end user screen	181
Figure 78: Platform Owner Market Management Portal – policies screen	182
Figure 79: Platform Owner Market Management Portal – financials screen	182
Figure 80: Platform Owner Market Management Portal – Disputes screen	183
Figure 81: Platform Owner Market Management Portal – models screen	183
Figure 82: Platform Owner Market Management Portal – data feeds screen	184
Figure 83: Output from sentiment analysis of workshop participants verbatim responses	192
Figure 84: Effect of applying a sequential entry strategy to a PPIM	199
Figure 85: Effect of applying a simultaneous entry of sides strategy to a PPIM	200
Figure 86: Effect of applying an entry with significant pre-commitment investment strategy to a PPIM	201

List of Tables

Table 1: The research journey	29
Table 2: Design science research contribution types, Gregor & Hevner (2013)	30
Table 3 Summary of theoretical contributions by chapter	32
Table 4: Design Science Process Elements	42
Table 5: Summary of proscribed aspects of successful infomediaries by various authors	49
Table 6: Five stages in personal informatics (Li, Dey & Forlizzi 2010)	50
Table 7: Foundational premises of S-D logic and their alternatives	59
Table 8: Comparison of entities which fulfil vendor and consumer roles in traditional and PPIM contexts.	73
Table 9: Summary of objectives of traditional infomediaries	74
Table 10: Objectives of infomediaries in a traditional and PPIM context	74
Table 11: Summary of design parameters’ relative impact on market quality	85
Table 12: Summary of interviewees’ opinions	88
Table 13: Issues and concerns raised by interviewees	88
Table 14: Classification of opinions and illustrative quotations from interviewees’ responses	90
Table 15: Logistic regression model	103
Table 16: Statistically significant variables and associated changes in odds	104
Table 17: Indicative supportive quotations from interviewees	125
Table 18: Process modules of a PPIM service architecture	130
Table 19: Reuse and variation matrix of process modules of a PPIM service architecture	131
Table 20: Activities, service modules and reuse submodules for end users on a PPIM	133
Table 21: Activities, service modules and reuse submodules for data contributing individuals on a PPIM	134
Table 22: Activities, service modules and reuse submodules for model developers on a PPIM	135
Table 23: Activities, service modules and reuse submodules for data feed developers on a PPIM....	136
Table 24: Activities, service modules and reuse submodules for platform owners on a PPIM	137
Table 25: Comparison of roles fulfilment between standard platform and PPIM	149
Table 26: Survey questions presented in end-to-end evaluation workshop	186
Table 27: Summary of all survey responses	188
Table 28: Summary of survey responses by question category	189
Table 29: Survey responses	190
Table 30 Summary of theoretical contributions by chapter	205

List of Key Terms and Definitions

Personal infomediary: a company facilitates market transactions using personal information for the mutual (not necessarily financial) benefit of both the individuals who created the information and the companies which wish to access it.

Personal Information (PI): Any data generated by an individual, either consciously or unconsciously, as they go about their daily lives, which capture some aspects of the individuals' status, activities, opinions, location, lifestyle or preferences.

Digital persona (DP): A digital persona is an organised, queryable collection of the data which has been generated by a single individual in the course of his or her everyday life. It comprises any data generated by the individual such as Facebook posts, tweets, sent emails, sent SMS messages, GPS data from a mobile phone, automotive sensor data and data from wearable devices.

Personal Information Market (PIM): The concept, trading rules, technology and information systems which comprise a market on which personal information is bought and sold.

Primary Personal Information Market (PPIM): The concept, trading rules, technology and information systems which comprise a market on which the original producer of the personal information exchanges their personal information in return for financial compensation.

Secondary Personal Information Market (SPIM): The concept, trading rules, technology and information systems which comprise a market on which an entity other than the original producer of the personal information exchanges other individuals' personal information in return for financial compensation.

Transaction object: The object which is to be bought and sold on the PPIM.

Permission based PPIM: A PPIM on which the transaction object is permission to use the personal information in certain specified ways. The ownership of the personal information remains with the individuals who produced it.

Ownership based PPIM: A PPIM on which the transaction object is the actual data itself which is transferred in its entirety from individuals to the purchasing entity.

Personal Data Store (PDS): A Personal Data Store is a centralised repository in which the personal data of data contributing individuals are stored.

Industry Platform: “products, services, or technologies that ...provide the foundation upon which outside firms ... can develop their own complementary products, technologies, or services” (Gawer & Cusumano 2014:418).

Digital Platform: “an extensible codebase to which complementary third-party modules can be added” and “Technical elements (of software and hardware) and associated organisational processes and standards” (de Reuver, Sørensen & Basole 2017, p. 3).

Certificate of Original Authorship

I, Ross Farrelly, declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Design, Architecture and Building at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note:
Signature removed prior to publication.

Date: 30th June 2019

Associated Publications

- Farrelly, R. & Chew, E. 2016, 'Designing a Personal Information Transaction Object', System Sciences (HICSS), 2016 49th Hawaii International Conference on, IEEE, Hawaii, pp. 4061-71.
- Farrelly, R. & Chew, E. 2016, 'Towards a Primary Personal Information Market', Twelfth International Conference on Technology, Knowledge, and Society.
- Farrelly, R. & Chew, E. 2016, 'Who's in to Win?: Participation Rate in a Primary Personal Information Market', Proceedings of the 27th Australasian Conference on Information Systems.
- Farrelly, R. & Chew, E. 2017, 'Designing a Primary Personal Information Market as an Industry Platform: a Service Innovation Approach', System Sciences (HICSS), 2017 50th Hawaii International Conference on, IEEE, Hawaii.
- Farrelly, R. and Chew, E.K., 2018. 'Transforming the worldwide market in personal information', Journal of Enterprise Transformation, pp.1-36.

Acknowledgments

First I would like to thank my supervisor, Professor Eng Chew. Eng has been an attentive, responsive and knowledgeable supervisor. He was willing to listen to my somewhat half-baked and ill-conceived proposal when I first went to him in the very early days of this research project. He encouraged me to pursue my passion in this endeavour, to examine my arguments, to gather evidence and to rigorously test my hypotheses. He also encouraged me to aim high and submit numerous papers to A level conferences, an experience which further tempered my research in the heat of academic debate. Without Eng's tireless encouragement and determined questioning, this dissertation would not have been possible.

I would also like to thank the staff at UTS who have assisted me with administrative tasks, booking travel and accommodation, proving reimbursements and of course the library staff who have tirelessly assisted me to locate resources as and when needed.

The writing of this dissertation would not have been possible without the willing participation of numerous interviewees and survey respondents who took the time to understand the research project, consider their views on the matter and provide responses to interviews and survey questions. To these respondents I owe a great deal of thanks.

I would like to make a special mention of thanks to my proofreader, Kristina Proft, who not only proofread this dissertation but also proofread a number of papers published during the course of this PhD. Without her extensive experience and almost universal knowledge of the intricacies of grammar, punctuation and formatting this dissertation would be a far inferior product. Having said that, any errors remaining are of course the responsibility of the author.

Finally, I would like to thank my long-suffering family who have had to accommodate the extensive amount of time I have put into researching and writing this dissertation.

Table of Contents

Abstract	2
List of Figures	4
List of Tables.....	7
List of Key Terms and Definitions.....	9
Certificate of Authorship and Originality	11
Associated Publications	12
1 Introduction	20
1.1 Motivation	20
1.1.1 Ease of collecting personal data.....	20
1.1.2 Increased volumes of personal information	20
1.1.3 Increased value of personal information	21
1.1.4 Ownership of personal information	21
1.1.5 The tide is turning	23
1.2 Overview of a PPIM	23
1.3 The research journey	26
1.4 Theoretical and practical contributions.....	29
1.4.1 Type of theory	30
1.4.2 Depth of development of the theory	30
1.4.3 Theoretical contribution	31
1.4.4 Managerial contribution	32
1.4.5 The novelty of the artefacts.....	33
1.4.6 Utility to a community of users and effectiveness of prototype	34
1.4.7 Relevance	35
1.5 Implications.....	36
1.6 Structure of this dissertation	36
2 Research methodology and justification	37
2.1 Introduction	37
2.2 Research methodologies used in related research.....	37
2.2.1 Typology of research methodologies	37
2.3 Candidate design research methodologies	38
2.3.1 Soft systems methodology	38
2.3.2 Design Science	38

2.4	Adherence to Design Science methodology	41
3	Literature reviews	44
3.1	Intermediaries and infomediaries	44
3.1.1	Proscriptions for a successful infomediary	45
3.1.2	Summary and analysis	48
3.2	Personal informatics	49
3.2.1	Models of personal informatics	49
3.2.2	Data volumes in personal informatics	51
3.2.3	Data sources and amalgamation of personal informatics data	51
3.2.4	Summary	53
3.3	Value co-creation	53
3.3.1	Introduction	53
3.3.2	Terminology	53
3.3.3	Motivation	54
3.3.4	Historical overview of value co-creation	55
3.3.5	G-D logic and S-D logic	57
3.3.6	Co-creation	63
3.3.7	Other aspects of co-creation	64
3.3.8	Conclusion	65
3.4	Personal Information Markets	66
4	Why personal infomediaries failed	67
4.1	Introduction	67
4.2	Theoretical framework	68
4.3	Analysis of personal infomediaries	69
4.3.1	Introduction	69
4.3.2	Guardians	69
4.3.3	Facilitators	70
4.3.4	Brokers and bonus brokers	71
4.3.5	Incentivisers	72
4.4	Concluding remarks on the viability of Infomediaries	72
5	PPIM conceived as a two-sided market	73
5.1	Introduction	73
5.2	Design and development	75
5.3	The transaction object	75

5.4	The IT infrastructure	77
5.5	Demonstration	78
5.6	Evaluation	79
5.7	Conclusion	80
6	The Transaction Object on a PPIM	80
6.1	Method	81
6.2	Design criteria	83
6.3	Design parameters	83
6.4	Proposed design solution	86
6.5	Personal information transaction objects (PITO):.....	86
6.6	Evaluation	88
6.7	Conclusion	89
7	Individuals' appetite to participate on a PPIM conceived as a two-sided market.....	92
7.1	Research questions	92
7.2	Literature review	92
7.2.1	Motivation to share	93
7.2.2	Value of personal information	93
7.2.3	Willingness to share	94
7.2.4	Summary and critique	95
7.3	Method	95
7.4	Company types.....	96
7.5	Uses	96
7.6	Price	97
7.7	Requests	97
7.8	The IT infrastructure	97
7.9	The survey	98
7.10	Results	102
7.11	Conclusions	104
7.12	Discussion	104
8	PPIM conceived as a four-sided industry platform	105
8.1	Introduction	105
8.2	Background	106
8.3	Method	108
8.4	Design solution	109

8.4.1	Service business strategy design	109
8.4.2	Customer type and customer value proposition design	110
8.4.3	Service concept design	111
8.4.4	Service architecture design	112
8.4.5	Service system design	113
8.4.6	Customer experience design	114
8.4.7	Monetisation design	117
8.5	Prototype	117
8.6	Evaluation	125
8.7	Conclusion	126
9	PPIM conceived as a digital platform	127
9.1	A service architecture for a PPIM	129
9.2	Design of base service modules	130
9.3	Core processes involved in a PPIM	132
9.4	Scenarios enabled by the service architecture.....	138
9.5	Interfaces	139
9.6	Hierarchy of API	140
9.7	Value chain enabled by the service architecture	141
9.8	Service architecture enabling platform evolution	142
10	PPIM end to end.....	148
10.1	Comparison of PPIM and standard platform	148
10.1.1	Roles.....	148
10.1.2	Location of app execution.....	149
10.1.3	Definition of a PPIM app	150
10.2	Workflow to create a targeted marketing personal data product.....	151
10.3	Workflow to create a personalised service response personal data product	153
10.4	Governance	156
10.4.1	Decision rights	156
10.4.2	Voting rights	156
10.4.3	Control mechanisms.....	156
10.4.4	Quality control	157
10.4.5	Governance examples	158
10.4.6	Pricing	159
10.5	Congruence between <i>i</i> SIM and end-to-end validation of PPIM	160
10.5.1	Service business strategy	160

10.5.2	Customer type and customer value proposition	162
10.5.3	Service concept design	164
10.6	End-to-end fitness	168
10.7	Evaluation of the end-to-end conception of the PPIM as a four-sided market	168
10.7.1	Workshop design and methodology	168
10.7.2	Attendees.....	184
10.7.3	Results	188
10.7.4	Conclusions to be drawn from the workshop.....	193
11	Successful digital platform adoption strategies as applied to a PPIM	193
11.1	Introduction	193
11.2	Determinants of platform adoption	194
11.2.1	Introduction	194
11.2.2	Importance of first mover advantage	195
11.2.3	Expansion strategies.....	195
11.2.4	Network effects	195
11.2.5	Ease of adoption	196
11.2.6	Variety of product offers	196
11.2.7	User innovations	196
11.2.8	Solving business problems.....	197
11.2.9	Availability of complementary products	197
11.2.10	Number of complementors.....	197
11.2.11	Pricing and network effects.....	197
11.2.12	Addressing the chicken and egg problem	198
11.3	Applying platform evolution best practice to PPIM	199
11.3.1	Addressing the chicken and egg problem	199
11.3.2	Exploiting network effects on a PPIM	202
11.3.3	Developing a network of complementors on a PPIM	203
11.3.4	Timing of introduction of a PPIM.....	204
11.3.5	Openness	204
11.3.6	Pricing and fee distribution on a PPIM	204
12	Conclusion	205
12.1	Contribution	205
12.1.1	Theoretical contribution	205
12.1.2	Managerial contribution	207
12.2	Implications.....	207

12.3	Fulfilment of aims	210
12.4	Limitations	211
13	Further Research	212
14	Appendices	214
14.1	Appendix 1	214
	Selected transcripts of interviews regarding the nature of the transaction object.....	214
14.1.1	Transcript 1: Interview with Interviewee 1	214
14.1.2	Transcript 2: Interview with Interviewee 2	219
14.1.3	Transcript 3: Interview with Interviewee 3	226
14.2	Appendix 2	232
14.2.1	Transcript 4: Interview with Interviewee 4	233
14.2.2	Transcript 5: Interview with Interviewee 5	242
14.2.3	Transcript 6: Interview with Interviewee 6	252
14.3	Appendix 3	264
14.3.1	R code used to generate and analyse data with regards to individuals' appetites to participate on a PPIM.....	264
14.4	R code used to simulate population data in section 8.5	280
15	References	312

1 Introduction

1.1 Motivation

There is a growing sense that while the digital age has undoubtedly brought many benefits, the financial benefits of Web 2.0 have been distributed unevenly. It has been argued that companies, rather than individuals “...have appropriated the largess of the digital age asymmetrically” (Tapscott 2016).

Here in Australia, Harris & Cilento (2016) found in their recent Productivity Commission report that individuals ought to have the right to “retain the power to view information held on them, request edits or corrections, and be advised of disclosure to third parties; have improved rights to opt out of collection in some circumstances; and have a new right to a machine-readable copy of data, provided either to them or to a nominated third party, such as a new service provider.”

1.1.1 Ease of collecting personal data

In a way, however, the ability for individuals to generate and collect data about themselves is not new. Throughout history people have been curious about their own behaviour and have kept diaries recording their daily activities. Benjamin Franklin kept detailed diaries recording whether or not he had adhered to certain virtues such as temperance, sincerity and justice (Franklin 2013, p. 73). Buckminster Fuller kept an extensive collection of facts and artefacts documenting his life, amounting to 270 linear feet by the time he passed away (Weintraub 2013).

But with the advent and availability of portable and wearable electronics and low-cost data storage it is now easier than ever to record an ever-increasing number of facts about one’s life at an increasingly granular level of detail. This growing phenomenon of collecting, analysing, sharing and learning from data oneself been given a number of names including the “quantified self” movement (Wolf 2010), “life-logging” (Quantified Self 2013) and “personal informatics” (Personal Informatics 2014).

1.1.2 Increased volumes of personal information

Further to the fact that the volume of personal information being generated by individuals is rising, ‘almost every computational system a person interacts with keeps a detailed log of that person’s behaviour’ (Wiese 2013, p. 41) Therefore, we can expand the concept of personal informatics beyond the confines of devices and data which are specifically designed to record facts about a person’s life and include the data generated by any interaction a person has with a data generating device. The use of social media sites (2012), instant messaging (Taitague 2013), email (Levenstein 2013) and mobile devices (Radicati 2013) are all on the rise and all generate growing volumes of personal information.

1.1.3 Increased value of personal information

Indeed, there is a growing consensus that this personal information is valuable. Detailed personal information about the demographics, preferences, buying habits and geographic location of consumers is valuable because it allows companies to personalise their marketing messages and thereby increase sales. The economic benefit that digital identities could deliver in Europe alone by 2020 has been estimated at €330 billion annually for organisations and €670 billion annually for individuals (Rose, Rehse & Röber 2012).

In 2009, the then European Consumer Commissioner Melena Kuneva pronounced that “Personal data is the new oil of the Internet and the new currency of the digital world” (Kuneva 2009). The World Economic Forum deemed personal information to be a new asset class (World Economic Forum and Bain & Company 2011). Furthermore, there is a large and growing secondary market in personal information (Stahl, Schomm & Vossen 2014), in which companies such as Rapleaf (now Towerdata) collect, aggregate and sell personal information (Riederer et al. 2011). Schwartz (2003, p. 2056) describes the value of personal information as follows:

Personal information is an important currency in the new millennium. The monetary value of personal data is large and still growing, and corporate America is moving quickly to profit from this trend.

Note that he stresses that *corporate America* is benefitting from this data, not the individuals who generated it.

Joined streams of personal information identified with the individuals who generated them allow finely grained segmentation of a customer base. Queries with multiple conjunctions can be run on these data which allow organisations to identify and develop niche markets (Dalgic & Leeuw 1994, p. 40) and to personalise products and services (Searby 2003). Companies use detailed aggregated data to create targeted offers, innovate coupon offers and improve loyalty (Acquisti 2010).

1.1.4 Ownership of personal information

This dissertation has argued that an increasing volume of valuable personal information is now being generated. The next issue to address is the ownership of these data. There are numerous candidates who could claim to own personal information. For example: the original creator of the data, the compiler who organises it into a useful form, the funder who financed the creation of the data, or the data subjects who are described in the data (Johnson 2012).

In the context of personal information, it is argued that personal information ownership is a human right of the original creator of the data. It is proposed that the aggregated personal information store is an extension of the person himself – a ‘digital aggregated self’ (Williams & Neal 2012), and that the

individual who is at the centre of this data generating activity has the right to control access to it – a concept known as the “new deal” of data ownership (Berinato 2014).

However, there are social and economic benefits to both companies and individuals in sharing personal information (Acquisti 2010, p. 7), so it is important that the individuals have the ability to grant access to their personal information when they see fit and that they have the means to do so. To exploit the value in personal information the individual must control access to it and have visibility of that access (Prins 2004, p. 7). Clearly, individuals cannot control the diverse data streams they generate because they are captured and stored on multiple platforms, but if properly enabled, they could control and monitor access to the aggregated personal information repository. Williams & Neal (2012, p. 170) argued that

...while server owners may possess the disaggregated user data stored on their servers, individuals should hold the rights to their set of aggregated data that is stored throughout the entire network of online servers.

The ownership and proper use of the aggregated personal information store can be realised through an appropriate ownership model. Schwartz (2003) drew on two key concepts, inalienability and use-transfer restriction, to devise such a model. Inalienability is the principle that an individual has a right to use his data as he sees fit, and use-transfer restriction is the principle that permission to access personal information may be granted to an entity but only if the individual has the ability to prevent further access to that data by additional entities. Schwartz argued that a hybrid model of personal information ownership which combines inalienability with use-transfer restriction and a default opt-in has the potential to prevent privacy breaches while allowing the value of personal information to be realised.

Clearly there is no widespread primary market in personal information on which an individual can sell access to his or her personal information for financial gain.

However there is a secondary market in personal information, in which companies collect personal information, and on-sell it in a variety of ways (Spiekermann & Novotny 2014, p. 2). The secondary market in personal information is opaque and operates without the explicit consent of the individuals who generated the data (Conger, Pratt & Loch 2013, p. 406). Unauthorised secondary use of personal data is one of highest ranked personal information privacy concerns (Milberg, Smith & Burke 2000, p. 39).

Classical economics states that complete information and transparency creates economic efficiency (Miller & Holmes 2010). To realise that economic efficiency, it is clear that there is an opportunity to establish a primary market in personal information. This thesis argues that the raw materials of a valuable information asset are being generated, but individuals currently lack the means to capture,

aggregate and trade permission to access their digital aggregated assets. A primary market in personal information would address this need. However, according to Tim Berners-Lee, one of the founders of the internet, it would represent a “paradigm shift” in the way we think about and use personal information (Zomorodi 2017).

1.1.5 The tide is turning

There is growing discontent with this one-sided relationship between corporations which gather and exploit the value of individuals’ personal information and the relatively powerless individuals who originally created the data. The emergence of a number of organisations, research projects and companies which aim to adjust the balance of power in favour of the individual is evidence of this change in attitude. To name just a few, The Hub of all Things (HAT Research and Development 2017) is an initiative driven by a consortium of universities in the UK which is designed to allow individuals to capture, store, use and monetise their personal data. Cross Cloud (Crosscloud 2017) is an MIT initiative which aims to help individuals to integrate, share and analyse multiple personal data streams in a secure and transparent way. Mydex (Mydex 2017b) is a community interest company established in to give individuals:

...greater control of their lives with the freedom to choose how and what personal data to share in order to have a balanced relationship with each other, society as a whole, governments and organisations. (Mydex 2017a)

The emergence of these and many similar organisations shows that the ownership, control and monetisation of personal data by individuals is a pressing issue in today’s modern interconnected world and demonstrates the relevance of the research undertaken in this dissertation.

1.2 Overview of a PPIM

The conception of an appropriate design for a Primary Personal Information Market (PPIM) has evolved during the execution of the research – as would be expected with the iterative process of design science. See for example Venable, Pries-Heje & Baskerville (2012).

As developments in mobile and wearable computing lower the barrier to the collection of individuals’ digital exhaust, it is now easier than ever for people to collect and store a valuable trove of personal information about themselves. These data are valuable to companies, but currently there is no easy way for individuals to reap a financial benefit by selling the personal data they generate as they go about their daily lives. This dissertation aims to take the first step in addressing the question as to why this is the case, and what could be done to remedy the situation.

So the question is: why is there no primary market in personal information? If such a market were to be built, what are the design features that should be included in it? This dissertation asserts that there is an

opportunity to design and build a viable primary personal information market. It argues from first principles for the need for such a market and addresses the question as to how such a market could be designed. The following sections of this dissertation take the form of an argument. These sections comprise a series of observations and assertions (each with supporting evidence) and one conclusion, namely that there is a need for a Primary Personal Information Market (PPIM). We then outline the methodology used to design such a PPIM and describe the theoretical framework upon which a PPIM could be built. This is then followed by an overview of the IT infrastructure required to support a PPIM, three theoretical case studies to demonstrate how a PPIM would work in practice and a discussion of the privacy implications of a PPIM.

There is a growing recognition that the personal information generated by individuals as they go about their daily lives embodies significant economic value (Rees 2014). These data allow companies to identify and develop niche markets (Dalgic & Leeuw 1994, p. 40) and to personalise products and services (Searby 2003). The potential economic impact of digital identities in the coming decades in the European economy alone has been estimated at €330 billion annually for organisations and €670 billion annually for individuals (Rose, Rehse & Röber 2012).

The value of this personal information is further demonstrated by the thriving secondary market in personal information, in which companies collect personal information and on-sell it in a variety of ways (Spiekermann & Novotny 2014, p. 2). However, this secondary market is opaque and operates without the explicit consent of the individuals who generated the data (Conger, Pratt & Loch 2013, p. 406).

A PPIM is a market in which the original producer of personal information, also referred to as the data subject, sells access to that information in some way to a consumer who wishes to benefit from access to that information. When ownership of the data is granted the resulting market is an ownership-based PPIM, when permission to use the personal data is granted it is a permission-based PPIM. These primary markets stand in contrast to a secondary market in personal information in which the original producer of the personal information is not a direct participant.

Farrelly & Chew (2016b) make the case for a PPIM. They argue that, with recent developments in mobile and wearable technology, the volume and granularity of personal information is rapidly increasing. Furthermore, they observe that this personal information is valuable to companies. They cite as evidence for this the thriving secondary market in personal information.

The widespread adoption of Social Media changed the way personal information is generated, stored and used (Rawassizadeh et al. 2013; Wiese 2013). There is now more personal information generated than ever before (Radicati 2013). That information is increasingly detailed and captures an increasing number of aspects of our lives, and as a consequence it is increasingly valuable to organisations (Rees

2014). Furthermore, it is well known that business-to-customer (B2C) companies have multiple uses for the social media and other personal information of both customers and prospects (Gualtieri 2013; Holdren & Lander 2014; Rose, Rehse & Röber 2012; Teevan, Dumais & Horvitz 2005) and the strategies companies employ to access this information has been well researched (Marsden 2011). Yet research into the effect of offering financial compensation for access to individuals' social media and other personal information is still lacking. The question of obtaining legitimate access to the social media data of both customers and prospects is an important one and needs to be answered by companies as they adapt to the wide ranging impact social media is having on the way many businesses operate (Aral, Dellarocas & Godes 2013).

Currently there is no widely adopted means by which companies can offer financial compensation for access to individuals' personal information. de Montjoye et al. (2014, p. 2) argue for the need for individuals to retain ownership of their personal information and for the establishment of a "fair and efficient market" in personal information. It is argued elsewhere (Farrelly & Chew 2016a, 2016b) that one way to address this is the Primary Personal Information Market (PPIM).

A PPIM is a market in which the original producer of the personal information sells access to that information in some way to a consumer who wishes to benefit from it. This stands in contrast to the secondary personal information market (Conger, Pratt & Loch 2013) in which a secondary party gains financial benefit by selling other individuals' personal information.

As mobile and wearable technology continues to evolve it is becoming easier than ever to capture an ever-increasing volume of personal information (Radicati 2013), information which is not only voluminous, but also increasingly detailed and therefore of great value to organisations (Rees 2014). However, currently there are no readily available means by which an individual can receive financial compensation for granting access to his or her personal information. One solution to this is the PPIM.

Our research and evaluation of an individual facing prototype shows that a PPIM is, broadly speaking, viable (Farrelly & Chew 2016c) and should be permissions based (Farrelly & Chew 2016a) and that there is likely to be significant appetite from individuals to participate on such a market (Farrelly & Chew 2016c).

A second iteration in the design process is required to redesign for scalability and to explore the potential for commercialising a PPIM through the provisioning of a variety of possible personal data products (PDPs). Consequently, this dissertation employs the integrated Service Innovation Method (*iSIM*) (Chew 2015) as a framework by which to design a PPIM as an industry platform for innovative personal data service delivery (Krueger et al. 2015).

iSIM is "an integrated cross-disciplinary, holistic method to design and commercialize service innovation" (Chew 2015:1). It is a method to simultaneously design an innovative service offering an

associated business model with the aim of engendering sustained value co-creation between customer and firm.

1.3 The research journey

This section summarised the research journey undertaken in the course of this research project (see).

Following Alvesson & Sandberg (2011), who propose gap spotting and problematising as two distinct but not mutually exclusive methods for developing research questions, this dissertation outlines the objectives of this research as follows.

The main objective of this research is to make progress towards designing a workable PPIM. The design of a workable PPIM in its entirety is a major cross-disciplinary undertaking comprising multiple lines of research, which would be beyond the scope of a single PhD dissertation. Nevertheless, in order to make progress toward this goal we outline below a number of research questions we seek to address in this work. The research objectives are to answer the following questions.

1. What are the useful theoretical frameworks we can draw upon when designing a PPIM? How do the theories of market engineering, value co-creation and industry platforms inform our understanding of PPIM design?
2. Have attempts been made in the past to establish a PPIM? If so, why did they fail? Are there any current initiatives, either commercial or non-commercial, which address the problem outlined in section 1.1 above?
3. How might a PPIM be designed as a two-sided market, which joins individuals who are willing to sell their personal information with buyers who are willing to purchase that information?
4. What is the nature of the personal information transactions object which is to be bought and sold on the PPIM?
5. How might a PPIM be designed as an m-sided industry platform, which connects contributions from individuals who are willing to sell their personal information and developers who wish to develop mechanisms to extract value from that personal information with buyers who are willing to purchase the resulting personal data products?
6. Would a PPIM have sufficient 'thickness' to be viable? That is to say, would there be a sufficiently large number of transactions conducted on a PPIM in order for the PPIM to be viable? This question must be answered in two parts:
 - a. To what extent are individuals willing to participate on a PPIM? What factors determine their willingness to participate?

- b. To what extent are companies willing to participate on a PPIM? Would the personal information offered on a PPIM be a sufficient value for companies to make purchases on the market?
7. What steps would be taken to establish a PPIM in practice? How would the chicken-and-egg problem encountered when establishing a nascent market be addressed? How would a sufficient number of participants on all sides of the market be induced to participate in order to make the market viable?

Our research journey began with an examination of why previous attempt to monetise personal information has not been successful. Our research focused on the so-called ‘infomediaries’ which were founded in the late 1990s and early 2000s.

Having established the reasons that these infomediaries failed to successfully establish themselves, we designed, evaluated and published an initial conceptual model for a two-sided PPIM (Farrelly & Chew 2016b).

The next step in the research journey was to design, evaluate and publish a conceptual model for a transaction object which could be traded on a PPIM and also contribute a conceptual model of a personal information transaction object (Figure 7).

Having established a model of for a two-sided PPIM and a transaction object which could be traded on it, we then designed a simulation which can be used to test the appetite of individuals to participate on a PPIM (Farrelly & Chew 2016c). Two hundred and nine (209) individuals engaged with this simulation and the results of this study showed that a large proportion of digital citizens would be willing to grant permission to use their personal data via a PPIM while retaining ownership of those data.

Following the presentation of these results, it was clear that there was a need to improve and refine the design of a PPIM in the light of the evaluation from our research so far. It was decided to reconceptualise a PPIM as an industry platform in order to facilitate co-creation of value by PPIM stakeholders with a view towards PPIM commercialisation. This dissertation utilised *iSIM* as an innovation method for industry platform design. The result of this research was to expand the PPIM from a two-sided market to a four-sided market with the inclusion of model developers and data feed developers into the platform (Farrelly & Chew 2017).

After the evaluation of the PPIM designed as a four-sided industry platform it was clear that a more detailed design was required to progress towards the realisation of a workable PPIM. When searching for a suitable paradigm within which to provide this additional level of detail it became clear that there was a more suitable alternative to industry platform theory.

Industry platforms have been defined as “products, services, or technologies that ... provide the foundation upon which outside firms ... can develop their own complementary products, technologies, or services” (Gawer & Cusumano 2014, p. 418). Most (but not all) of today’s industry platforms are digital and many have physical elements included in the product offering (Evans & Gawer 2016).”

On the other hand, digital platforms exist purely in the digital realm. They have been defined from a technical viewpoint as “an extensible codebase to which complementary third-party modules can be added” and from a sociotechnical view as “Technical elements (of software and hardware) and associated organisational processes and standards” (de Reuver, Sørensen & Basole 2017, p. 3). Since the purpose of a PPIM is the collection, analysis and monetisation of personal information, which takes place in a purely digital realm, we claim that a PPIM can be better classified as a digital platform.

As a consequence, the next step on our research journey was to reconceptualise a four-sided PPIM as a digital platform and develop a detailed service architecture, API specification, and value chain within the context of a digital platform. We again utilised *iSIM* to address this research problem. The results of this research are to be published shortly (Farrelly & Chew forthcoming).

The next step in the research journey was to synthesise the findings from our previous research and to add to our conceptual model of a workable PPIM by focusing on the end-to-end workings of the market and the governance structures necessary to ensure both market evolution and market innovation. This research resulted in a second prototype and was extensively evaluated by a panel comprising both academic and industry experts.

We conclude the research journey by applying platform evolution and adoption theory to a PPIM to develop strategies and guidelines, which could be implemented to ensure PPIM adoption in practice.

Table 1: The research journey		
	Theory of design and action	Generalisations in the form of a design artefact
Chapter 5	A partial theory of design and action which says how to construct a two-sided PPIM	
Chapter 6	A partial theory of design and action which says how to construct a transaction object which could be traded on a PPIM	Conceptual model of a personal information transaction object (PITO)
Chapter 7		1) A simulation to test the appetite of individuals to participate on a PPIM 2) A data model to store information on a two-sided PPIM
Chapter 8	A partial theory of design and action which says how to construct a four-sided PPIM as an industry platform (developed using <i>iSIM</i> and industry platform theory)	1) A prototype PPIM 2) Service experience blueprint for participation on a PPIM
Chapter 9	A partial theory of design and action which says how to construct a four-sided PPIM as a digital platform (developed using <i>iSIM</i> and digital platform theory)	A service architecture for a four-sided PPIM
Chapter 10	An extension to the partial theory of design and action developed in chapter 9	An extension to the PPIM prototype developed in chapter 7

1.4 Theoretical and practical contributions

To argue that we have made a theoretical contribution to knowledge, we must first clarify what constitutes a legitimate contribution to knowledge in the context of design science research. We examine this question through two different lenses, one which examines the type of theory under consideration and the other which examines the depth and level of development of the theory.

1.4.1 Type of theory

For this dissertation, we take theories to be defined as ‘abstract entities that aim to describe, explain, and enhance understanding of the world and, in some cases, to provide predictions of what will happen in the future and to give a basis for intervention and action.’ (Gregor 2006, p. 616). Gregor (2006) outlines five different types of theory which are germane to information systems research: theory for analysing, theory for explaining, theory for predicting, theory for explaining and predicting and theory for design and action. She argues that prototyping and design science are alternative terms for this type of theorising. Since our research follows a design science methodology and includes the development of a prototype and other design artefacts, we claim that our research makes a theoretical contribution of the fifth of Gregor’s types, a theory for design and action.

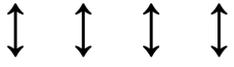
As a second perspective, theoretical contributions in design science research, we also reference Dwivedi, Puroo & Straub (2014) who define contribution types as:

- Construct: A conceptualisation to describe problems and their solutions within a domain
- Model: A set of propositions or statements expressing relationships among constructs
- Method: A set of steps used to perform a task. Methods are based on a set of underlying constructs (language) and a representation (model)
- Instantiation: Realisation of an artefact in its environment

1.4.2 Depth of development of the theory

Furthermore, Gregor & Hevner (2013, p. 339) argue that contributions to knowledge can take a number of forms, not with regard to the type of theory but with regard to the depth to which the theory has been developed. They argue that a contribution can be made at three different levels of development as outlined in Table 2.

Table 2: Design science research contribution types, Gregor & Hevner (2013)

	Contribution Types	Example Artifacts
More abstract, complete, and mature knowledge	Level 3. Well-developed design theory about embedded phenomena	Design theories (mid-range and grand theories)
	Level 2. Nascent design theory—knowledge as operational principles/architecture	Constructs, methods, models, design principles, technological rules.
More specific, limited, and less mature knowledge	Level 1. Situated implementation of artifact	Instantiations (software products or implemented processes)

They argue that a contribution can be made at three different levels of development: a well-developed design theory, but it can also be a ‘partial theory, incomplete theory or even some particularly interesting

and perhaps surprising empirical generalisations in the form of a new design artefact' (Gregor & Hevner 2013, p. 339).

Since our research is prototypical in nature and explores a nascent and extensive field of endeavour, we claim that our research makes a contribution in the second and third of Gregory and Hevner's levels of development. The nature of our contribution will be further described in the following paragraphs.

Gregor & Hevner (2013) also present the DSR Knowledge Contribution Framework in terms of solution maturity and application domain maturity (Figure 1). Since our artefact is high in application domain maturity but develops and enhances other platforms which collect and store but do not monetise personal data, we claim that the contribution of this research falls into the improvement quadrant in Gregor and Hevner's 2x2 matrix.

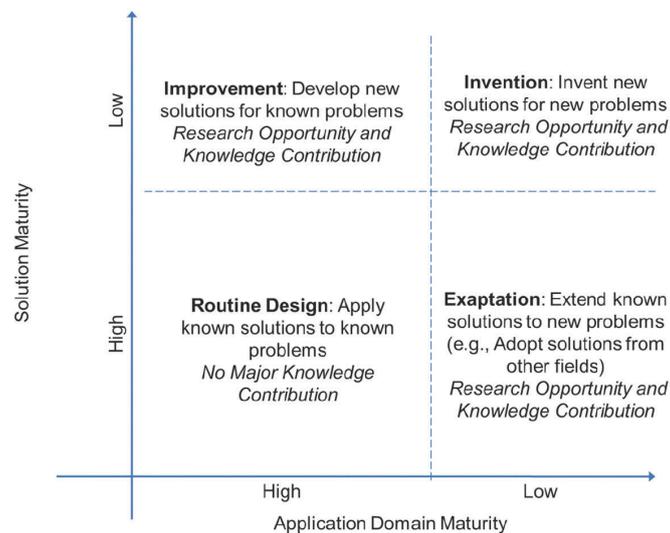


Figure 1: DSR Knowledge Contribution Framework (Gregor & Hevner 2013)

1.4.3 Theoretical contribution

The major theoretical contribution made by research is a detailed and multifaceted theory of design and action to construct a four-sided PPIM as a digital platform. This theory has been developed by applying *iSIM* to digital platform design theory in the context of a PPIM. The first iteration of this theory can be in chapter 9 and the second in chapter 10.

This theory of design and action is supported by a second major theoretical contribution, a generalisable design artefact in the form of a PPIM prototype. This prototype has been through two design iterations, the first in chapter 8 and the second in chapter 10.

A number of other supporting contributions were made in the course of the research journey. A partial theory of design and action which says how to construct two-sided PPIM which can be found in chapter 5. Chapter 6 contributes a partial theory of design and action which says how to construct a transaction object which could be traded on a PPIM and also contributes a conceptual model of a personal information transaction object (Figure 7). Chapter 7 contributes two generalisable design artefacts: a simulation which can be used to test the appetite of individuals to participate on a PPIM and a data model to store information on a two-sided PPIM (Figure 9). Three theoretical contributions can be found in chapter 8. The first is a partial theory of design and action, which says how to construct a four-sided PPIM as an industry platform. This theory was developed by drawing upon industry platform theory and *iSIM*. This chapter comprises a novel design approach of utilising *iSIM* to design an industry platform for a specific use case (in this case – the design of a PPIM).

1.4.4 Managerial contribution

Throughout this dissertation there are practical (managerial) contributions for potential PPIM platform owners and all four participant types. Taken as a whole, this dissertation demonstrates a novel approach to monetising personal information and for model developers and data feed developers to generate a passive income stream. Business owners and senior executive in industries which are undergoing ‘platformization’ (Constantinides, Henfridsson & Parker 2018, p. 1) will benefit from considering how to monetise the information they collect in the light of this research.

Table 3 summarises the theoretical contributions made in each chapter of this dissertation.

Table 3 Summary of theoretical contributions by chapter		
	Theory of design and action	Generalisations in the form of a design artefact
Chapter 5	A partial theory of design and action which says how to construct two-sided PPIM	
Chapter 6	A partial theory of design and action which says how to construct a transaction object which could be traded on a PPIM	Conceptual model of a personal information transaction object (PITO)
Chapter 7		1) A simulation to test the appetite of individuals to participate on a PPIM

		2) A data model to store information on a two-sided PPIM.
Chapter 8	A partial theory of design and action which says how to construct a four-sided PPIM as an industry platform (developed using <i>iSIM</i> and industry platform theory)	1) A prototype PPIM 2) Service experience blueprint for participation on a PPIM
Chapter 9	A partial theory of design and action which says how to construct a four-sided PPIM as a digital platform (developed using <i>iSIM</i> and digital platform theory)	
Chapter 10	An extension to the partial theory of design and action developed in chapter 9	1) A service architecture for a four-sided PPIM 2) An extension to the PPIM prototype developed in chapter 7

1.4.5 The novelty of the artefacts

We claim that the artefacts developed in this research are novel. Clearly there is no workable PPIM in existence upon which individuals can exchange access to their personal data in return for financial compensation. In addition to this, although there are a number of current research initiatives which share parts in common with the research in this dissertation, none are identical or sufficiently similar to undermine our claim of novelty.

For example, The Hat Project (HAT Research and Development 2017) ‘enables individuals to collect their own personal data through IoT-enabled objects and to control and combine it in any way into data “mash-ups” that can be (sic) shared in a privacy-preserving manner, to help us make better and more informed decisions in our lives and allow companies to offer us more personalised and customised offerings.’

This research initiative recognises that ‘Personal data is both a digital asset to the individual and a potential commodity for trade’ (HAT Research and Development 2017), however, it, and the research conducted by its members – see for example (Kumar et al. 2016; Ng et al. 2015; Ng & Wakenshaw 2016; Perera et al. 2016) does not explain how personal data can be commoditised into PDPs, how permission to access personal data could be traded, or how an m-sided market could be constructed to benefit data contributing individuals, model developers, data feed developers and end users.

A second related initiative, Solid (Conner-Simons 2015), is a project led by Prof. Tim Berners-Lee, which aims to create ‘true data ownership as well as improved privacy’ (The Solid Project 2016). This project is related to the research conducted in this dissertation in that it addresses questions of ownership

and privacy but is radically different in that it does not address the question of monetisation of personal information nor does it attempt to create a working marketplace for personal information products.

A third related but distinct research project is the OpenPDS/SA project (de Montjoye et al. 2014; MIT Human Dynamics Group 2017). This initiative focuses on control, privacy, anonymisation and auditability of personal information but does not touch on the monetisation aspect of this problem and is therefore distinct from the research in the dissertation.

Further evidence of the novelty of this research are numerous emails in my possession (R. Heard, S. Hollender, G. Lightowlers et. al., personal communication, Sept. 2016) from one of the leading patent attorney firms in Australia indicating that the research very probably contains sufficient original intellectual property to be patentable and agreeing to act on behalf of the author to prepare a patent application.

Our claim is that, although there have been numerous similar and related research projects, this research is sufficiently distinct from these to constitute an original contribution to knowledge.

1.4.6 Utility to a community of users and effectiveness of prototype

The artefacts presented in this research have been formally assessed by two communities of users: data contributing individuals and end users. The artefacts have also been informally assessed by model developers who represent a third group of users of a PPIM.

In the course of this research, we found that we can be 90% confident that the proportion of digital citizens who would participate in a permissions-based PPIM lies between 74.6% and 84.3%. We assert that this demonstrates that the artefact presented in this chapter in the form of a simulated PPIM (Figure 13 and Figure 14) was found to be of utility by a significant number of members of the communities of users.

The utility of the four-sided PPIM was assessed by an expert panel of representatives from eight businesses, one from each of Education, Cultural and Recreational Services; Accommodation, Cafes and Restaurants; Government Administration and Defence; and Communication Services and two from each of Finance and Insurance and Personal and Other Services. There was one micro, two small, three medium and two large businesses. Despite the fact that many interviewees suggested improvements and novel uses of the artefact, the assessment from this panel was, broadly speaking, supportive and expressed the view that the artefact presented would be useful in their businesses.

This research and the design for a PPIM was also presented to a group of approximately 200 data scientists at a data science meet-up in Sydney (Farrelly 2017). Although there was no formal assessment following the talk, anecdotally I can report that the audience was receptive to the ideas and indicated they would be willing to participate on such a market should it be brought into existence.

1.4.7 Relevance

It is one thing to make a contribution, but it is also necessary in Design Science to make a relevant contribution (Hevner et al. 2004). We claim this research is extremely relevant and timely for a number of reasons.

This research is relevant because it is in line with changing attitudes with respect to the ownership and control of personal data in Australia and internationally. Evidence for this change in attitude can be found in changes to Australian federal law which give individuals increased control and visibility of the fate of their personal data. For example, in February 2017 amendments to the Privacy Act were passed meaning that ‘if an organisation subject to the Privacy Act incurs an “eligible data breach”, it will have to alert ... the people whose data has been compromised.’ (Redrup 2017). Furthermore, the Australian Productivity Commission (Harris & Cilento 2016) in its 2016 draft report of the value of both private and government data assets recommended:

greater control for individuals over data that is collected on them by defining a new Comprehensive Right for consumers. This right would mean consumers: retain the power to view information held on them, request edits or corrections, and be advised of disclosure to third parties. (Harris & Cilento 2016, p. 2)

Internationally, Europe has strengthened its privacy and data protection legislation with the introduction of the General Data Protection Regulation (GDPR) in May 2018 (General Data Protection Regulation 2017).

These examples indicate that there is a move towards recognising the value of personal information and that individuals should have increased control and visibility of how their data are being used.

Further evidence for this change of attitude can be found in a number of conferences, both academic and commercial, which focus on the issue of ownership of personal information. Examples of these conferences are: the Preparing For The Era Of Consumer Controlled Data Workshop (Infact Decisions 2017), the Wolfson-HAT Annual Symposium Series (Wolfson College Cambridge & HAT Community 2017). There are also examples of individuals taking class actions against companies who misuse their personal data (Cormack 2017), and as wearables become more and more invasive, the potential for similar suits grows.

In conclusion, we argue that this dissertation makes a significant theoretical contribution to the literature in a number of important ways and that it contains useful, relevant and effective knowledge. This chapter has been read by Shirley Gregor, one of the leading Design Science theorists in the world and, in private correspondence, she has indicated that these claims to theoretical contribution are sound and valid (Gregor 2018).

1.5 Implications

The implications of this research are twofold. There are implications for the field of platform design and implications for society as a whole.

Regarding the implications for the field of platform design, the novel use of *iSIM* for the design of both an industry platform and a digital platform shows that the application a service innovation approach to platform design can be a fruitful one.

The implications of this research for society as whole are wide ranging and far reaching. If implemented, a PPIM has the potential to act as a platform for the exchange of access to personal data of any digitally connected e-citizen. This research lays the foundation for further research into a workable PPIM and ultimately, if implemented, the establishment and maintenance of such a market. This could radically transform the lives of many digitally connected e-citizens, especially in developing countries as they increasingly gain access to the internet. The full implications of this research are explored in detail in chapter 12.2.

1.6 Structure of this dissertation

This dissertation is based on the original publications listed above, and the structure of the dissertation reflects this fact. In chapter 1 the research project is introduced by outlining the motivation behind the theoretical and the practical contribution that the research has made. Then an overview of a Primary Personal Information Market (PPIM) is presented, which is the central subject of this dissertation.

In chapter 2 a number of candidate research methodologies are surveyed and the justification for adopting Design Science as the research methodology for this project is presented.

On account of the multi-disciplinary nature of this research project a number of literature reviews were conducted in the course of the research. These are presented in chapter 3 which comprises the literature review which was conducted in the initial stages of this research project together with the literature reviews conducted for each of the papers published in the course of this research.

Chapter 4 explores why previous attempts to establish personal infomediaries failed while chapter 5 comprises the initial conception of a PPIM as a two-sided market and is based on Farrelly & Chew (2016b).

Chapter 6 comprises the design and description of the transaction object which would be traded on a PPIM. It is largely based on Farrelly & Chew (2016a).

Having outlined a design for a two-sided PPIM and defined the transaction object, in chapter 7 these designs are tested by researching individuals' appetite to participate on a PPIM. This chapter is based on Farrelly & Chew (2016c).

In chapter 8, the design for a PPIM as a four-sided industry platform is presented. This chapter is based on Farrelly & Chew (2017) and represents a second iteration in the design science process in which the initial conception of the PPIM is further refined in light of the feedback received from previous publication.

In chapter 9 this research pivots from designing an industry platform to the design of a digital platform. This chapter comprises detailed design for a modular service architecture for a PPIM conceived as a digital platform. Much of the material in this chapter was published in (Farrelly & Chew forthcoming).

In chapter 10 the end-to-end design of the PPIM is presented in more detail. This chapter comprises detailed workflows, a detailed API design, an exploration of the governance structure needed for the successful operation and evolution of the PPIM and the results of an evaluation workshop.

Chapter 11 outlines how adoption strategies for the establishment and growth of a PPIM could be implemented. In chapter 12 a discussion on contributions and implications for both research and practice are presented; and suggestions for further research are presented in chapter 13. Appendices containing the code utilised for statistical analysis and some selected interviews are presented in chapter 14.

2 Research methodology and justification

2.1 Introduction

This chapter examines the possible research methodologies which could be employed to address the research aims, namely: to develop a sound theoretical basis upon which to develop a workable, efficient, robust, successful Personal Information Market (PIM). This development may possibly extend and improve the four space model of Spiekermann & Novotny (2014). Strengths and weaknesses of each of the candidate research methodologies are identified, the most suitable candidate is selected and a justification for the selection is presented.

2.2 Research methodologies used in related research

2.2.1 Typology of research methodologies

There are a number of research methodologies available for research in Information Systems and there is an important distinction that divides these methodologies into two main groups. One group of IS research methodologies is drawn from the behavioural sciences paradigm. These methodologies are ‘concerned with explaining and truth finding’ (Goes 2014, p. iv). They are ‘reactive in the sense that the information systems or IT artefact is given’ (Hevner et al. 2004, p. 98). Methodologies which fall into this category include: library research, observation research, case studies, statistical techniques,

simulations, forecasting, action research and experimental research. These are known as *interpretive methodologies*.

The second group of methodologies is proactive and concerned with ‘creating and evaluating innovative IT’ (Hevner et al. 2004, p. 98). Methodologies such as: collaborative practice research, design science, soft design science, soft systems methodology and prototyping fall into this group. These will be called *design methodologies*.

There is also a third group of research methodologies, known as combination or mixed methodologies, which draw on both interpretative and design methodologies.

It is clear from the research questions and research problem that this dissertation is interested in designing and creating a new artefact, vis. the theoretical basis for and practical implementation of a sound, workable personal information market. Therefore, only *design methodologies* need be considered for the design phases of this research.

2.3 Candidate design research methodologies

2.3.1 Soft systems methodology

Soft systems methodology (SSM) is ‘an action-oriented process of inquiry into problematic situations in which users learn their way from finding out about the situation, to taking action to improve it.’ (Checkland 2013). SSM has been used in the design of artefacts as diverse as a net-centric cyber defence system (Craig et al. 2014) and a search application for the BBC (Fenning, Dogan & Phalp 2014). One of the strengths of SSM is that it encourages the participants in the problem-solving exercise to examine the problem from multiple world views and to design a solution which is (at least partially) acceptable to all of them.

However, SSM is not well suited to the current research project as the primary focus is not to ‘take action to improve’ an existing artefact, rather it proposes to design a new artefact from scratch. Furthermore, it may not be feasible to gather multiple opinions from stakeholders regarding the issues around personal data that this research seeks to address, nor to gather the multiple conflicting world views which pertain to the use of personal data, thereby indicating that SSM may not be the most suitable research methodology.

2.3.2 Design Science

An early work in the history of what we would now call Design Science described a framework to improve design in IT systems in terms of constructs, models and methods (March & Smith 1995). In this conception, a construct is considered to be the language and concepts which allow the expression of the problem, envisioned solutions and the solution space within which the solution may lie. The

model then draws upon these constructs to express the problem and envisioned solutions. The method is considered to be analogous to a search within the solution space to find incrementally better solutions. The method also comprises the building of an instantiation of the preferred solution.

Simon (1996), in a book now considered to be one of the seminal texts in Design Science, argued that there was a need for a ‘science of the design that is tough, analytic, partly formalizable, partly empirical, teachable doctrine’ (Simon 1996, p. 112). In the fifth chapter entitled ‘The Science of Design: Creating the Artificial’ Simon argues that a design science is desirable and that at the time of writing there was an emerging trend towards a more rigorous approach to design in a number of leading universities. He established the fact that design was concerned with meeting set goals and he argues that utility theories are the appropriate approach to design systems to meet those goals but also canvasses other approaches such as design as the allocation of resources and design as a series of generate–test cycles.

Cross (2001) sought to further clarify the role of Design Science by drawing the distinction between scientific design, Design Science and the science of design. Scientific design simply means using scientific knowledge in your design process – for example understating the properties of the materials you will be using in your design artefact. It is a basic requirement in any project which seeks to create a physical instantiation of an artefact. Design Science is an attempt to formulate a method by which design must be done if it is to be called truly scientific. A science of design, on the other hand, is a study of design, which studies how design has been done in order to understand the principles behind it. It is analogous to the ‘science of science’ which studies how science has been done over the years.

Hevner et al. (2004) stated that Design Science is a research paradigm which pertains to the creation of ‘new and innovative artefacts’ (Hevner et al. 2004, p. 75). They built on March and Smith’s three aspects of Design Science (March & Smith 1995) by outlining a seven aspect conceptual framework for Design Science: design as an artefact, problem relevance, design evaluation, research contributions, research rigour, design as a search process and communication of the research. By paradigm it is meant ‘... a theoretical framework or a set of beliefs that determine the problems that are regarded as crucial, the ways such problems are conceptualized, the appropriate method of inquiry and the relevant standards of judgments of knowledge claims’ (Aljafari & Khazanchi 2013, p. 3745)

In an attempt to emphasise the iterative nature of Design Science, Hevner (2007) described the paradigm in terms of three cycles: a relevance cycle, a rigor cycle and a design cycle. Since Design Science is motivated by a desire to solve real world problems, Hevner argues that there is a need for a Relevance Cycle, which assesses the needs to be met by the research project and the importance of these needs. It ensures that the problem is worth solving and that it is of sufficient magnitude to warrant the time and effort to be spent on developing the solution. However, to give the solution the best chance of succeeding, it needs to be firmly based in previous research and established theories. This, according to Hevner, is the role of the Rigor Cycle, which grounds the research in a sound theoretical basis and

also provides an avenue for new theories developed in the course of the research project to be contributed to the knowledge base. Having established the need and designed a theoretically sound solution outline, the Design Cycle comprises the instantiation of the design artefact and an evaluation of the degree to which it has met the need it was designed to address.

In summary, Design Science can be understood to be a rigorous, scientific approach to identifying, scoping, researching and clarifying real world problems and then designing, building and testing new artefacts which are intended to solve those problems.

Peppers et al. (2006) and Peppers et al. (2007-8) argue in favour of an explicit Design Science Research Methodology. Their work in summarising previous work on DSR and distilling a DSRM is very applicable to this project. Following the methodology of (Peppers et al. 2006) and Peppers et al. (2007-8) the following steps would be suitable to achieve the research goals.

The *problem identification and motivation* stage involves defining the research problem, possibly by atomising it into a number of smaller sub-problems. It also involves justifying the value of a solution to the problem.

In light of the complexity of the task, it will be important to atomise the problem into sub-problems and attack each one separately. For each aspect of the project the following steps to design will be executed and a solution will be evaluated.

The *definition of the objectives for a solution* stage involves formulating the objectives of a solution based on the problem definition and taking into account the feasibility of alternative solutions. Where possible the research should attempt to quantify the benefits of the proposed solution over the current state.

The *design and development* stage involves the design and creation of an artefact, which is broadly defined to include constructs, models, methods and instantiations. This necessitates determining the required features of the artefact to meet the objectives identified in the previous stage.

The *demonstration* stage involves demonstrating that the developed artefact meets the objective of the solution by using the artefact to solve one or more examples of the problem outlined in the *problem identification and motivation* stage. This demonstration can take the form of a proof, simulation, case study or proof of concept project.

The *evaluation* stage involves comparing results from the demonstration stage with the objectives of the solution outlined in the *definition of the objectives for a solution* stage. Depending on the nature of the problem and solution this comparison could be done by such means as simulations, user feedback interviews or customer satisfaction surveys.

The communication stage of the research project will take place through the publication of academic papers and finally through the publication of this dissertation.

2.4 Adherence to Design Science methodology

Although there is a variety of views on how Design Science is best defined and the status of its philosophical underpinnings (Aljafari & Khazanchi 2013), the literature shows that Design Science is known widely as a methodology well suited to the design and evaluation of novel, ‘new-to-the world’ IS artefacts (Tuunanen, Baskerville & Beck 2017, p. 4664). It has been described as a ‘creative process that involves generation of new thoughts and imaginative jumps to future possibilities’ (Purao 2002, p. 17).

For example, Hevner et al. (2004) distinguish design science from behavioural science stating that whereas behavioural science ‘seeks to develop and verify theories that explain or predict human or organizational behaviour’, design science ‘seeks to extend the boundaries of human and organizational capabilities by creating new and innovative artifacts’. They propose IT artefacts as ‘constructs (vocabulary and symbols), models (abstractions and representations), methods (algorithms and practices), and instantiations (implemented and prototype systems)’ Hevner et al. (2004, p. 77).

Not only do numerous authors agree that design science is a methodology well suited to the design and evaluation of novel artefacts, there is also some consensus and commonality among authors on the design science process elements necessary to execute a design science research project. Peffers et al. (2006) presented a summary of these process elements (see Figure 2).

Objectives for a design science research process model	Archer (1984)	(Takeda et al. 1990)	Ekels and Rozenburg (1991)	Numamaker et al (1991)	Walls et al (1992)	(Rossi et al. 2003)	(Hevner et al. 2004)
1. Problem identification and motivation	Programming Data collection	Problem enumeration	Analysis	Construct a conceptual framework	Meta-requirements Kernel theories	Identify a need	Important and relevant problems
2. Objectives of a solution			Requirements				Implicit in “relevance”
3. Design and development	Analysis Synthesis Development	Suggestion Development	Synthesis, Tentative design proposals	Develop a system architecture Analyze and design the system. Build the system	Design method Meta design	Build	Iterative search process Artifact
4. Demonstration			Simulation, Conditional prediction	Experiment, observe, and evaluate the system			
5. Evaluation		Confirmatory evaluation	Evaluation, Decision, Definite design		Testable design process/product hypotheses	Evaluate	Evaluate
6. Communication	Communication						Communication

Figure 2: Design and design science process elements from IS other disciplines and synthesis objectives for a design science research process in IS (Peffers et al. 2006)

Since Peffers presented this summary in 2006 there have been a number of refinements and alternative steps suggested for the design science process steps, which are summarised in Table 4.

	(Peffers et al. 2006)	Peffers et al. (2007-8)	(Jones 2008)	(Gregor & Hevner 2013)
Problem identification and motivation	Problem identification and motivation	Problem identification and motivation	Introduction, literature review	Introduction Literature Review
Objectives of a solution	Objectives of a solution	Define the objectives for a solution		
Design	Design and development	Design and development	Specification of designed artefact	Artefact description
Demonstration	Demonstration	Demonstration.	Instantiation	
Evaluation	Evaluation	Evaluation	Evaluation	Evaluation
Communication	Communication	Communication	Discussion, conclusion	Discussion, Conclusion

Drawing on the literature described above, the steps included in this design process are described here:

The *problem identification and motivation* stage involves defining the research problem, possibly by atomising it into a number of smaller sub-problems. It also involves justifying the value of a solution to the problem.

In light of the complexity of the task, it will be important to atomise the problem into sub-problems and attack each one separately. For each aspect of the project the following steps to design will be executed and a solution evaluated.

The *definition of the objectives for a solution* stage involves formulating the objectives of a solution based on the problem definition and taking into account the feasibility of alternative solutions. Where possible the research should attempt to quantify the benefits of the proposed solution over the current state.

The *design and development* stage involves the design and creation of an artefact, which is broadly defined to include constructs, models, methods and instantiations. This necessitates determining the required features of the artefact to meet the objectives identified in the previous stage.

The *demonstration* stage involves demonstrating that the developed artefact meets the objective of the solution by using the artefact to solve one or more examples of the problem outlined in the *problem identification and motivation* stage. This demonstration can take the form of a proof, simulation, case study or proof of concept project.

The *evaluation* stage involves comparing results from the demonstration stage with the objectives of the solution outlined in the *definition of the objectives for a solution* stage. Depending on the nature of the problem and solution this comparison could be done by such means as simulations, user feedback interviews or customer satisfaction surveys.

The *communication* stage of the research project has taken place through the publication of academic papers (see list of associated papers above) and finally through the publication of this dissertation.

Each of these steps has been executed on our design cycles, with the communication aspect of design science methodology being fulfilled by the publication of associated papers (see page Associated Publications section above) and the publication of this dissertation.

Furthermore, iteration (Peffer et al. 2006; Puro 2002), also known as ‘design as a search process’, in which repeated design / instantiation / demonstration / evaluation / cycles are undertaken to incrementally improve the artefact under design, is a key aspect of design science. This research comprises three iterations of the design process as represented in Figure 3.

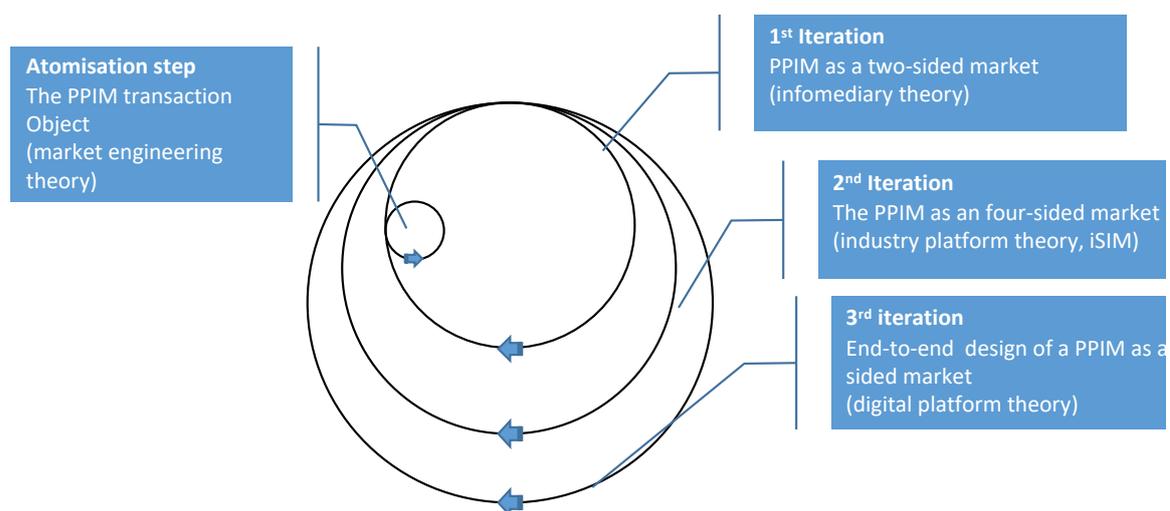


Figure 3: Iterations and atomisation steps in the design process

A second aspect of design science applied to the large and challenging problem of designing a PPIM is the concept of *atomisation*, the identification of sub-problems within a larger problem space to better understand the complexity of the challenge (Peffer et al. 2007-8). This process identified the need for a transaction object to be designed and evaluated, a process which is described in chapter 6.

3 Literature reviews

3.1 Intermediaries and infomediaries

The references in these literature reviews were sourced using a process similar to Kitchenham's systematic literature review methodology (Kitchenham et al. 2009) by searching Business Source Complete, ACM Digital Library, IEEE Explore, ProQuest and Scopus for the following words and phrases: intermediary, infomediary, data ownership and two sided markets.

The concept of the intermediary in commerce is not a new one. Since the 12th century intermediaries such as shippers, brokers and wholesalers played a part in facilitating trade and reducing transaction costs (Garcia 2000). In a climate of asymmetric information, intermediaries have traditionally played an important role in commerce; assisting vendors to find well-qualified prospects and helping potential buyers find vendors, which sell appropriate products. As goods and services become increasingly complex, intermediaries have played a greater role in reducing search costs and facilitating the classification and distribution of goods (Leickly 2004, p. 9).

In the early 1980s, the world wide web began its relentless expansion from government and academia into mainstream commerce and the everyday lives of millions of people (Leiner et al. 2009). From this expansion it was obvious that information technology would have a profound effect on how companies did business. The ability of companies to quickly and cheaply acquire, communicate and process large amounts of information had profound effects on the management and structure of firms, on the commercial marketplace as a whole and on the value chain in particular (Sarkar, Butler & Steinfield 1995, p. 3).

As consumers increasingly had direct access to the information they sought, 'disintermediation' started to take place in many sectors of the economy, including financial institutions, libraries and government (Gellman 1996).

In light of these changes, it was argued by some that the role of the intermediary would be greatly curtailed or even in some cases eliminated. Benjamin & Wigand (1995) argued that there was a possibility that retailers and wholesalers may be eliminated completely from the value chain. The

argument was that information technology would enable consumers and manufactures to deal directly with one another. By elimination, intermediaries and transaction costs would be reduced and consumers would have a wider choice of products by dealing directly with manufacturers (Sarkar, Butler & Steinfield 1995, p. 3).

However, as the WWW evolved and become more ubiquitous, the role of the intermediary continued to play a part in commerce. As the amount of information increased, the percentage of *relevant* information to a particular consumer looking to satisfy a particular need necessarily decreased, which had the effect of increasing search costs. Furthermore, e-commerce transacted at a distance meant that buyers were confronted with the choice of whether or not to trust the vendor to provide goods as described. With a growing amount of information available there was also the potential that buyers would simply not be aware of certain goods and would make wrong decisions about which goods were most suitable for a particular purpose. These inhibitors or friction points in the purchasing process – search cost, trust, missed opportunities and wrong decisions – provided a role for intermediaries in the new economy (Rose 1999, pp. 1 - 3).

In the late 1990s, Hagel & Singer (1999) published *Net Worth*, an influential book on the subject of intermediaries in the information economy and coined the term ‘infomediary’. As we shall see, this book attempted to redefine the role of the intermediary in the context of the networked world of e-commerce and outlined how intermediaries could benefit both buyers and sellers. There were number of well-funded attempts to put the concepts outlined in *Net Worth* into practice. Many investors and entrepreneurs were convinced of the validity of these enterprises and they attracted significant interest (Evans 2009). Companies such as Lumeria, AllAdvantage and Firefly were started in the 1990s but they failed to attract both sellers and buyers and in the end none of them were successful (Evans 2009).

3.1.1 Proscriptions for a successful infomediary

This section summarises proscriptions for successful infomediaries which have been put forward by a number of authors. In section 3.1.2 these proscriptions are compared and contrasted and findings are summarised. Discussions are restricted to infomediaries that seek in some way to exploit the value of the individual’s personal information as opposed to infomediaries that simply act as directories or centralised search hubs, such as those surveyed by Palvia & D'Aubeterre (2007).

Hagel & Singer (1999) outline a wide range of possible activities that an infomediary may undertake. They describe an infomediary as a company that captures, manages and maximises the value of personal information on behalf of individuals (pp. 33ff.), and also collects and aggregates vendor data in order to facilitate interactions between buyers and sellers (pp. 42ff).

Next, this conception of an infomediary as a company is summarised. These are companies that are involved in:

- 1) the **collecting** of personal information and vendor offers
- 2) the **selling** of access to personal information, by which the individual gains financial reward for granting access to his or her personal data
- 3) the **aggregating** of both personal information and vendor offers
- 4) **protecting** the individual from unwanted privacy intrusions and unwanted advertising via a privacy toolkit such as anonymised email addresses, anonymous web browsing software and infomediary issued credit cards
- 5) the **matching** of an individual's needs with vendor offerings. This is done by filtering and selecting the most relevant offers from vendors
- 6) **negotiating** with vendors on behalf of individuals, via reverse auctions and using combined market power to obtain better offers

Rose (1999) defines an intermediary as: 'an independent, profit-maximising economic agent mediating between two market sides in presence of market imperfections' (Rose 1999, p. 51) and the act of intermediation as 'the bridging (of) the incompatibilities between two (market) sides involved in a transaction by transforming of output attributes of the supply market side to appropriate input attributes of the demand market side.' (Rose 1999, p.51)

Rose (1999) lists a number of critical success factors for information intermediaries. Some are structural (such as the gaining of a critical mass of users) and describe the *form* the successful infomediary must have while others describe the *function* of the successful infomediary. These include:

- 1) provision of personalised service
- 2) provision of value-adding services in addition to the services of information mediation
- 3) focus on a thematic domain, i.e. providing a service in a specialised sector of the market (Rose 1999, p. 162)

Leickly (2004) argues that the infomediary as conceived by Hagel & Singer (1999) is a flawed concept because, by emphasising the protection of individuals' data it was excessively customer facing and consequently there was insufficient benefit to companies to make the concept viable (Leickly 2004, p. 10).

Her major point of difference with Hagel and Singer is that to be successful, an infomediary must not be in the business of commodifying consumer information and privacy (Leickly 2004, p. 57). Rather, she argues that 'the intermediary can ... be interpreted as a mechanism for controlling uncertainty and sharing risk to reduce agency costs' (Leickly 2004, p. 20). She lists seven rules for success for digital-age intermediaries:

- 1) provide value for businesses and consumers
- 2) promote trust in online transactions
- 3) become experts in specialised industries
- 4) respect information privacy
- 5) achieve critical mass
- 6) gain first-mover advantage
- 7) take advantage of information commodities

Some of these success factors describe the function of the infomediary (1, 3, 7) while the others describe *how* these functions should be done and are not specific to infomediaries. All businesses that have a business model which depends on a large customer base would aim to gain the trust of their customers, respect their privacy and achieve a critical mass of users.

Leickly also argues that infomediaries provide the most value and have the highest chance of success when they work in markets that have high information asymmetry (Leickly 2004, p. 16), due to the complexity of the products traded, and high transaction costs in the form of search costs (Leickly 2004, p. 46).

We can summarise Leickly's prescription for a successful infomediary as a company that:

- 1) provides value for both businesses and consumers
- 2) works in a specialised industry which has high search cost and has high information asymmetry due to complex products
- 3) deals in information commodities
- 4) does not commodify consumer information and privacy.

Grover & Teng (2001) argue that infomediaries are set to play an increasingly important role in e-commerce and that there are seven main functions successful infomediaries should perform (Grover & Teng 2001, p. 80). These are:

- 1) **search and complexity services** – whereby the infomediary facilitates the search process for the buyer
- 2) **matching services** – whereby buyers and products are matched, benefiting both buyer and vendor
- 3) **content services** – for example independent reviews of products
- 4) **community services** – whereby customers can share and discuss opinions

- 5) **informational services** – whereby customers are informed of new products based on their previous buying patterns
- 6) **privacy protection services** – whereby the customer’s privacy is protected to prevent unwanted advertising
- 7) In some cases, the infomediary could provide **infrastructure services** such as facilitation of secure transactions.

Bakos & Bailey (1997) conducted a survey of thirteen e-commerce firms in order to examine the changing roles of intermediaries. They identify the main roles played by a successful intermediary as:

- 1) aggregating buyer demand and/or supplier offerings
- 2) matching suppliers and customers
- 3) building trust (between buyer and seller)
- 4) providing inter-organisational market information.

3.1.2 Summary and analysis

Section 3.1.1 shows that there is considerable consensus amongst the authors studied regarding certain attributes of a successful infomediary. All the authors agree that reducing search cost and reducing uncertainty / gaining trust are key roles of successful infomediaries and four of the five authors agree that matching buyer requirements with vendor offerings is an important role.

However, one clear dichotomy emerges – that of the commodification of personal information. Only Hagel advocates collecting, storing and selling access to personal information while the other authors do not.

From section 3.1.1 we identify three main types of proscriptions for successful infomediaries, summarised in Table 5.

- 1) Infomediaries whose primary purpose is to aggregate buyer purchasing power. These benefit companies by helping to sell a large volume of goods or services and help the buyer by negotiating better deals. Companies such as Groupon and Mercata fall into this category.
- 2) Infomediaries whose primary purpose is to amalgamate vendor offerings into one easily searchable platform. This benefits both buyer and seller by reducing search costs and it benefits the buyer by enabling price comparisons. Companies such as eBay, Amazon and Trademe fall into this category. Palvia & D'Aubeterre (2007) list many examples of this type of infomediary.

- 3) Infomediaries whose primary purpose is to commoditise the personal information of individuals. We will call these *personal infomediaries*. To date there have been no widely adopted successful implementations of the personal infomediary.

Table 5: Summary of proscribed aspects of successful infomediaries by various authors

	(Hagel & Singer 1999)	(Rose 1999)	(Leickly 2004)	(Grover & Teng 2001)	(Bakos & Bailey 1997)
Collecting and storing personal information	✓				
Selling access to personal information	✓				
Protecting privacy of individuals	✓			✓	
Negotiating better deals on behalf of individuals	✓				
Matching requirements of individuals with products offered by vendors	✓	✓		✓ ¹	✓
Aggregating buyer purchasing power	✓				✓
Reducing search cost by aggregating vendor offerings	✓	✓	✓	✓	✓ ²
Reducing uncertainty		✓	✓	✓ ³	✓ ⁴

3.2 Personal informatics

3.2.1 Models of personal informatics

The references in these literature reviews were sourced using a process similar to Kitchenham's systematic literature review methodology (Kitchenham et al. 2009) by searching Business Source Complete, ACM Digital Library, IEEE Explore, ProQuest and Scopus for the phrases personal informatics, informatics, quantified self and personal information.

There have been a number of attempts to formulate models for personal informatics. Li, Dey & Forlizzi (2010) propose an iterative five-stage model of personal informatics with the aim of improving the diagnosis, assessment and prediction of problems in personal informatics systems – including barriers

¹ through informational services

² through inter-organisational market information

³ through reviews and community comments

⁴ by providing trust

faced by users at each stage of the personal informatics process. The five stages they identify are shown in Table 6.

Table 6: Five stages in personal informatics (Li, Dey & Forlizzi 2010)

Stage	Description
preparation	Identifying the motivation, method and scope of personal data collection
collection	Collecting the personal data
integration	Preparing, transforming and analysing the information
reflection	Understanding and thinking about the implications of the data which has been collected
action	Taking actions based on the insights provided by the data

Related to the ‘reflection’ stage, Sellen & Whittaker (2010) use theories of human psychology to explore the reasons why people might be motivated to capture personal data about their lives with a view to augmenting human memory. They summarise their findings in terms of five Rs: recollecting, remembering intentions, reminiscing, reflection and retrieving.

Whittaker et al. (2012) attempted to derive design principles for life logging platforms by studying existing digital tools. Their motivation was to formulate principles which would facilitate the design and instantiation of life logging platforms which have a greater chance of being taken up by a wider audience. They developed four principles in the design of life logging platforms: Selectivity, Embodiment, Synergy, and Reminiscence. *Selectivity* is an aspect of a life logging platform which, either automatically or manually, selects or flags certain events as more important than others, thus facilitating more meaningful recall of events. *Embodiment* is an aspect that better embeds digital mementos into the user’s everyday life. *Synergy* is an aspect that allows a life logging platform to complement user memory as and when needed. *Reminiscence* is an aspect that goes beyond simple recall of facts and helps users build narratives around their mementos in their digital archives.

Wiese (2013) proposes to build a platform which supports multiple steps in the personal informatics phenomenon: ‘data aggregation and linking, feature extraction, ground truth collection, model building and refinement, and access to inferences through a unified API’ (p. 42). He proposes to store personal informatics data from multiple sources using a data model comprising: a data source table, a behavioural

data table, an identity table, and an attributes table. His proposed platforms will also accept ground truth inputs for many different types of users.

The approach taken by Whittaker et al. (2012) appears to be a fruitful and useful approach to developing a design principle and could be adapted and applied to designing a personal data commodification platform. The approach proposed by (Wiese 2013) also shows promise but there is a need for a lot more detail to be added to the proposed data model.

3.2.2 Data volumes in personal informatics

Main stream, high volume personal informatics platforms are of particular interest in the exploration of a personal data market because they generate a large volume of data. It has been estimated that, worldwide there were:

- 1.4 billion monthly active users on Facebook as of December 31, 2017 (Facebook 2017)
- 3.578 billion people using the internet in 2018 (International Telecommunication Union 2017)
- 3.4 billion instant messaging accounts in 2013 growing to over 4.4 billion by 2017 (Taitague 2013)
- 3.9 billion active email accounts in 2013 increasing to over 4.9 billion by the end of 2017 (Levenstein 2013)
- 7.7 billion mobile devices (including tablets) in use in 2014 rising to over 12.1 billion by 2018 (Radicati 2013)

These figures show that large numbers of individuals are using data generating products and services and consequently are generating an enormous amount as they go about their daily lives.

3.2.3 Data sources and amalgamation of personal informatics data

Although not studying the amalgamation question per se, Li, Dey & Forlizzi (2010) studied a related topic, namely, the variety of sources of personal informatics data. They conducted a survey (n=68) to discover the most common types of personal informatics data collected by users. They grouped the sources into two classes, those which were automatically generated and those which were collected manually. The most common automatically collected data sources were (in decreasing frequency): bank statements, email history, credit card bills, phone call history, SMS history, IM history, financial software, electricity bills, browsing history and search history. The most common manually collected data sources were: calendar events, status updates, work activities, blog posts, weight, exercise, browser bookmarks, time at work, social bookmarks and mood.

Furthermore, there has been a considerable amount of research carried out on various ways in which personal data can be amalgamated into a coherent framework with the aim of discerning interesting correlations between disparate data sources.

The importance of the integration stage of informatics has been studied by Li, Dey & Forlizzi (2010). They found that barriers at the integration stage can significantly deter users from utilising personal informatics platforms.

Kyriacou & Davis (2008) address the problem of ‘user modelling’, which they describe as ‘the ability to model a dynamic and changing user throughout lifetime interactions with a variety of resource providers’ (p. 647). They take the approach of developing a scrutible model expressed in Semantic Web technologies such as XML, RDF, OWL, etc. which allows each user to see which aspects of their personal information has been modelled. It also allows users to control the privacy setting for each aspect of their user model.

Elliott et al. (2009) propose a very high level unified architecture for life-long user profiles. The proposed architecture outlines aspects that are necessary for the building of a user profile. They describe the necessity of choosing appropriate ontologies to categorise the entities and events in a user’s life – a step which they propose can be instantiated through a Life-Long User Model (LLUM) Application Programming Interface (API) . They also propose entity extraction, representation of features, automated logic rules, recommendation and representation of the data to the user as other key stages in the unified architecture.

In work related to the LLUM API and the recommendation stages of the unified architecture for life-long user profiles of Elliott et al. (2009), Kleek et al. (2010) designed and built a personal information assistance engine which incorporated data feeds from social networking and life-tracking and fed them into a unified model which represented people, places and things at different times and states. They then used this platform to simulate a simplified personal assistant-like platform which could remember and act on simple Boolean logic based rules such as automatically updating your Facebook status when attending a concert.

Abel et al. (2010) examine techniques for and benefits of aggregating user profile information from multiple sources on the web. They study integration of both explicitly provided profile information and implicitly generated data such as tags assigned to bookmarks or images.

Albatal et al. (2013) present an ‘extensible and customisable cloud-based semantic analytic service for Lifelogging’ (p. 144). Using the SenseSeer mobile application they collect information from the user’s mobile phone and store it in a cloud server. The app then presents the data in three use cases: ‘My Health’, ‘My Location’ and ‘My Social Activity’. Their platform seeks to auto-categorise the user’s time based on input from a number of sensors and apps.

Rawassizadeh et al. (2013) explored the use of the smart phone as a platform to collect life-logging data. Their aim was to move beyond custom built, dedicated life logging tools which collect only a few data streams and require the user to carry an addition device. They present a framework architecture, data model and implementation of 'UbiqLog', a life-log framework, which uses a mobile phone as a device for life logging.

3.2.4 Summary

We can see from this survey of personal informatics platforms that there have been many disparate attempts to design and build data aggregation platforms for personal informatics. An ideal platform that combines the best of many of these attempts would include: an extensible system which can easily accommodate multiple data sources, an open API which allows developers to build apps which connect to the system, the functionality to accept ground truth input from the user (to facilitate the building of a reliable training set), the ability to auto categorise user data into activities and a cloud based service which stores the user's data and can be accessed from many locations.

3.3 Value co-creation

3.3.1 Introduction

The references in these literature reviews were sourced using a process similar to Kitchenham's systematic literature review methodology (Kitchenham et al. 2009) by searching Business Source Complete, ACM Digital Library, IEEE Explore, ProQuest and Scopus for the phrases value creation, value co-creation, goods-dominant logic and service-dominant logic.

In this literature review we examine the research which has been conducted into value, value creation and value co-creation. A brief historical overview of the subject is presented including a summary of the contribution to this subject made by the major schools of thought in marketing research. A number of models for value and value creation are presented and various aspects of value and value co-creation are identified and discussed. The main tenets of goods-dominant logic (G-D logic) and service-dominant logic (S-D logic) are compared and contrasted. The review concludes with some remarks on the place and implications of value co-creation theory in the commodification of personal data.

3.3.2 Terminology

In this literature review we will use the term 'goods' to mean physical manufactured items produced by a firm. We will use the term 'services' to mean services as a *process* provided by a firm for the benefit of individuals (beneficiaries in general) and itself (Vargo & Lusch 2004), and the term 'products' to mean the output of an activity (be it manufacturing or service activity) either in the form of goods or services or both (Vargo & Lusch 2004).

3.3.3 Motivation

As developments in mobile and wearable computing lower the barrier to the collection of individuals' digital exhaust, it is now easier than ever for people to collect and store a valuable trove of personal information about themselves. This data is valuable to companies, but currently there is no easy way for individuals to reap a financial benefit by selling the personal data they generate as they go about their daily lives. My research aims to address the question as to why this is the case, and what could be done to remedy the situation.

When both producer and consumer act together to create value we say they engage in *co-creation*. Or, to put it more succinctly, 'Co-creation is the process by which mutual value is expanded together' (Ramaswamy 2011, p. 195). Vargo & Lusch (2004, p. 11) pioneered the services dominant logic of marketing which recognised that the 'separation of production and consumption is not a normative goal', and recognised the necessity 'of viewing the consumer as a coproducer.'

Value co-creation will play a major part in the manufacture, commodification and marketing of personal data products. This is because, on one hand, individuals will be the source of the personal data but they will also have control over how their personal data is shared with firms. On the other hand, firms will have a say in what data products will be valuable to them and how they wish to use the personal data products which they purchase access to.

It is for this reason that the following literature review of value co-creation has been undertaken.

3.3.3.1 Parameters of this literature review

This literature review will examine the concepts of value, value creation and value co-creation as they relate to the buying and selling of products. As these topics are intimately related to marketing, the literature on the two main conceptions of marketing, goods-dominant logic (G-D logic) and service-dominant logic (S-D logic) will also be reviewed.

3.3.3.2 Structure of the review

Section 3.3.4 provides a historical overview of the development of value co-creation. In section 3.3.5.3 concepts related to value co-creation are examined. I examine how the concept of value has been conceived in the literature and I examine the two main conceptions of marketing, G-D logic and S-D logic in section 3.3.5. Section 3.3.6 examines the various conceptions and aspects of co-creation. Section 3.3.7 discusses additional aspects of value co-creation in the context of the commodification of personal data. Section 3.3.8 concludes the value co-creation literature review.

3.3.4 Historical overview of value co-creation

3.3.4.1 Aristotle on value

Aristotle was one of the first thinkers in the western tradition to identify two aspects of value: *value-in-use* and *value-in-exchange* (Johnson 1939, p. 448). Value-in-use is the benefit the owner experiences by using the goods he owns, and value-in-exchange is the benefit the owner may accrue by exchanging the good he owns for something else. Clearly, they are related concepts as there is no value-in-exchange if there is no demand and there is no demand if there is no perceived value-in-use on the part of the buyer. Although the term ‘value’ in the study of economics has usually referred to value-in-exchange, the distinction between the two aspects of the concept has proved valuable in many contexts (Richins 1994).

The Classical and Neoclassical economists (Malthus 1803; Smith 1776) considered value to be embodied in manufactured goods. The role of the firm was to manufacture products which embodied value. Individuals became wealthy by the acquisition of goods.

Following the Industrial Revolution in the 18th and 19th centuries and the accompanying profusion of readily available manufactured goods, marketing began to emerge as a field of academic study. In the early 1900s the fledgling discipline of marketing was primarily concerned with describing commodities to the market (Copeland 1923) and, with time and place, utility of commodities. It also concerned itself with the facilitation the change of ownership of commodities from firm to customer and with the facilitation of bringing goods to market (Weld 1917).

As world economies become more industrialised, resources came to mean something less tangible and more dynamic and, most importantly, they were seen to include human resources, skills and knowledge. As this shift in understanding of the nature of wealth took place, the understanding of marketing also changed. In the latter half of the 20th century the focus turned more towards the customer. The concepts of customer focus and the idea that value is determined by the customer (Levitt 1960) were introduced. The notion that firms embed value in goods was nuanced by the insight that the value only existed if it was recognised by the customer. There was a growing recognition that customers bought goods to satisfy needs and the concept of value-in-use started to gain wider recognition. Furthermore a number of technological innovations transformed value creation from a linear phenomenon to a phenomenon in which value could be created in a more parallel manner (Ramirez 1999).

Hand in hand with this increasing focus on the customer came a reconceptualisation of the firm as a service providing entity and consequently the marketing of services started to slowly emerge as an academic sub-discipline (Vargo & Morgan 2005). With services marketing came the idea of co-creation on account of the immediacy of service delivery and the subjective nature of service value.

Three distinct schools of thought in service marketing – each with a different perspective on value co-creation – developed in the 20th century: the French School, the Nordic School and the North American School.

3.3.4.2 The North American School

From 1979 onwards in America there was a growing awareness of the role customers can play in co-production and co-creation of both goods and services. Firms started to explore how self-service and customer participation could play a role in their operations (Bendapudi & Leone 2003). The American Marketing Association and the Marketing Science Institute played key roles in developing the science of service marketing as distinct from goods marketing. In 1977 MSI published its first report comprising five papers which distinguished service marketing from goods marketing (Berry & Parasuraman 1993). In the late 1990s and early 2000s marketing began to focus on the marketing of services (Grönroos 1982, 1984), relationship marketing (Duncan & Moriarty 1998) and a plurality of other aspects of an evolving marketing landscape. This diversity of approaches to marketing came together in the S-D Logic of (Vargo & Lusch 2004) which is discussed in detail in section 3.3.5.2.

3.3.4.3 The French School

The main contribution of the French School was the servuction model of co-creation. This model identified and classified the resources necessary for co-creation of value in firm/customer interactions. The model identifies four main components: contact employees, physical resources, the focal customer and fellow customers who may also be present during the service process (Grönroos 2012, p. 1524). The model emphasises the role of the focal customer and the fellow customers in the value creation process. The physical resources component has been expanded to include emerging technologies, such as web and mobile platforms.

3.3.4.4 The Nordic School

In parallel to the work of the French School, researchers in Finland and Sweden developed their own theories of service marketing. They explored concepts such as employees engaged in service delivery acting as part-time marketers to influence brand awareness (Gummesson 1991) and focused particularly on firm/customer interaction as a value creation opportunity. Grönroos (1982) introduced the notion of interactive marketing and explained the need for it by arguing that, in certain fundamental ways, goods and services were different in nature. For example, with services, the creation and consumption often takes place simultaneously. Furthermore, since with services, the customer is not a passive recipient but reacts and expresses their experience of the service as it is being received, he or she can influence how other customers experience the service and in this way co-create the service. Hence the need for a more interactive service marketing approach for services. Grönroos (1984) drew a distinction between technical and functional quality, arguing that the customer is concerned not just about *what* she receives,

but *how* she receives it. The experience of receiving the service contributes to the perceived service quality and, since the perception is subjective, the customer in a sense co-creates the service. Lehtinen & Lehtinen (1991) used interaction quality as one of two dimensions to measure service quality. Collectively, this body of work came to be known as the Nordic school of thought (Grönroos 2012, p. 1524).

3.3.5 G-D logic and S-D logic

There are two main perspectives on the process of value creation: G-D logic and S-D logic.

3.3.5.1 G-D logic

Viewed from the perspective of goods-dominant logic (G-D logic), value is created by the firm in the form of discrete products. These products are then sold to the customer for a price which reflects the value created by the firm and the value creation process with respect to that product comes to an end. The roles of customers and firms are viewed as distinct and the firm is considered to be the sole creator of value (Vargo, Maglio & Akaka 2008, p. 146).

In G-D logic, the primary unit of exchange is the physical object which has been manufactured by the firm. These physical objects are operand resources on which the firm acts to imbue with value. The firm determines the manner in which it will impart value to the manufactured goods and that value is recognised as value-in-exchange at the time of sale. The customer receives value by obtaining and using ('consuming') the goods, and the firm also treats the customer as an operand resource to which it markets, promotes and distributes its goods.

G-D logic focuses on the economic transaction in which operand resources (those resources on which an act or operation is performed) are bought and sold, on concrete physical objects and the value which is imbued in them by the firm (Vargo & Lusch 2004, p. 1). In order to maximise profit and effectively control the production of goods, firms manufacture standardised goods away from the market, store them as inventory and then bring them to market when required (Vargo & Lusch 2004, p. 7).

3.3.5.2 S-D logic

This section comprises an overview of S-D logic with special attention to the question of co-creation which is the topic of this literature review. However S-D logic is of central importance to the topic of co-creation because 'Six of the 10 premises (1, 3, 6, 7, 9 and 10) relate to value creation and co-creation and their marketing implications' (Grönroos 2011a, p. 280).

While marketing traditionally concerned itself with the exchange of commodities as outlined in section 3.3.5.1, in the 1950s marketing researchers began to focus more on the need of the customer (Vargo & Lusch 2004, p. 1). This was followed in the 1980s by a diversification of views as to what the main aim

of marketing should be. Conceptions such as relationship marketing, quality management, supply and value chain management and services marketing began to emerge. By the end of the 20th century, this multiplicity of conceptions leads to a recognition that a paradigm shift in marketing theory was needed to consolidate and comprehend these emerging trends (Vargo & Lusch 2004, p. 1).

That new paradigm shift, now known as S-D logic, was initially formulated by Vargo & Lusch (2004) with an aim to integrate both goods and services marketing in the same logic. The logic put forward by Vargo & Lusch (2004) came to be known as S-D logic to indicate the view that ‘all businesses are seen as service businesses’ (Aitken et al. 2006, p. 276). Central to the emerging logic was a change in emphasis from *operand resources* (fixed, finite physical raw materials upon which operations were done) to *operant resources* – resources such as human skill, knowledge, processes and procedures that are capable of acting on other resources (operand or operant) to create value. This new role of the firm was described in terms of attracting and developing knowledge and skills which could give the firm a competitive advantage, identifying potential customers who would benefit from these knowledge and skills, developing value propositions (Polo Peña, Frías Jamilena & Rodríguez Molina 2014, p. 1044) to meet customers’ needs and gathering customer feedback to adjust the value proposition accordingly. This interaction between firm and customer lead to the conclusion that value is defined by and *co-created* with the consumer rather than being embedded in output of the firm. This value co-creation arises from integration of resources of both client and provider for mutual benefit. They even went as far as claiming that the customer is *always* a co-producer (later changed to co-creator).

The S-D logic outlined by Vargo & Lusch (2004), and since updated by Vargo & Lusch (2016), met with broad agreement from many scholars (Day et al. 2004) although a number of points of difference emerged. For example, Ballantyne & Varey (2006) argued that *knowing*, *relating* and *communicating* are the three strands of value creating activities which take place between firm and customer. They also argue that, only when value propositions are reciprocal between firm and customer, will the exchange be sustainable.

Vargo & Lusch (2008) further refined and extended their conception of value creation in S-D logic by discussing the interactive and networked aspects of value creation. They clarified that value co-creation does not just take place between firm and customer at the level of the individual but also within and between organisations. They claim S-D logic can be applied to entities, such as families, firms and countries, which engage in exchange with one another (Vargo & Lusch 2008, p. 5).

Vargo & Lusch (2004) formulated eight foundational premises (FPs) of S-D logic which they later modified and extended (Vargo & Lusch 2008, 2016).

Grönroos (2011b) argued that a number of these FPs lacked clarity and precision, were too broad and sweeping and stripped S-D logic of some of its potential utility. His main objection was that it is not the

case that customers always are co-creators, but rather that, in certain circumstances when the firm and customer meet face-to-face, there is opportunity for the two entities to co-create value. Grönroos proposed alternatives to a number of the FPs as proposed by Vargo and Lusch (see Table 7).

While Vargo & Lusch argued that co-creation is a ubiquitous phenomenon, Grönroos took the view that it is limited in both time and place. He argues that it only takes place at the intersection of customer and provider spheres – the ‘joint sphere’. He also argued that it only takes place during actual interaction between customer and firm, at which time the firm plays a part in the customer’s value creation process (Viljakainen, Toivonen & Aikala 2013).

The core of interaction is a physical, virtual, or mental contact, such that the provider creates opportunities to engage with its customers’ experiences and practices and thereby influences their flow and outcomes. (Grönroos & Voima 2013, p. 7)

Table 7: Foundational premises of S-D logic and their alternatives

	Vargo & Lusch (2004)	Vargo & Lusch (2008)	Grönroos (2011b)
FP1	The application of specialised skills and knowledge is the fundamental unit of exchange	Service is the fundamental basis of exchange	Reciprocal value creation is the fundamental basis of business, with service as a mediating factor
FP2	Indirect exchange masks the fundamental unit of exchange	Indirect exchange masks the fundamental basis of exchange	
FP3	Goods are a distribution mechanism for service provision	Goods are a distribution mechanism for service provision	All resources and processes are distribution mechanisms for service provision, however without including value in themselves
FP4	Knowledge is the fundamental source of competitive advantage	Operant resources are the fundamental source of competitive advantage	
FP5	All economies are services economies	All economies are service economies	
FP6	The customer is always a co-producer	The customer is always a co-creator of value	Fundamentally, the customer is always a value creator
FP7	The enterprise can only make value propositions	The enterprise cannot deliver value, but only offer value propositions	1) Fundamentally, the firm is a facilitator of value for the customer 2) Provided that the firm can engage with its customers’ value-creating processes during direct interactions, it has opportunities to co-create value jointly with them as well The firm is not restricted to offering value propositions only, but has an opportunity to directly and actively influence its customers’ value creation as well
FP8	A service-centred view is customer oriented and relational	A service-centred view is inherently customer oriented and relational	
FP9		All social and economic actors are resource integrators	

FP10		Value is always uniquely and phenomenologically determined by the beneficiary	1) Value is accumulating throughout the customer's value-creating process 2) Value is always uniquely and both experientially and contextually perceived and determined by the customer
------	--	---	--

3.3.5.3 Value and value creation

The concept of value has many aspects (Sánchez-Fernández, Angeles Iniesta-Bonillo & Holbrook 2009; Woodall 2003) and has been examined in a number of different ways in the literature. Before examining some of the different ways value has been modelled, it is worth noting that there is general agreement in the literature that value is determined by the customer, not by the firm. The firm may attempt to encapsulate value in its products but ultimately value is determined by what the customer realised from those offerings (Khalifa 2004, p. 647).

3.3.5.4 Value models

The basic premise of the value component model of value is that in order to make a purchase, the customer must perceive that the product on offer in some way meets one or more of her needs. She must also perceive that the product is worth the asking price. The customer assesses the suitability of products on offer in terms of dissatisfiers, satisfiers and delighters (Khalifa 2004). Dissatisfiers are features of a product that cause dissatisfaction if they are absent from a product. Satisfiers are features which cause satisfaction if they are present and delighters are features which are unexpected but cause delight when present.

Another model for customer value is the cost / benefit model (Huber, Herrmann & Morgan 2001). In this model the customer assesses the product on offer in terms of the benefits he is likely to receive should he decide to purchase. These benefits can be changes in the customer's wellbeing such as an increase in pleasure, avoidance of pain, reduction in effort, time saving or an increase in social status. The customer also assesses the costs involved with acquiring the purchase under consideration. Costs involved in making a purchase can include: financial costs, search costs and time costs (possibly including time to learn to use the newly acquired product). The customer value is the difference between the perceived benefits and perceived costs of acquiring the product or service.

Woodall (2003, p. 2) expresses a cost / benefit conception of value as 'any demand-side, personal perception of advantage arising out of a customer's association with an organisation's offerings, and can occur as reduction in sacrifice; presence of benefit (perceived as either attributes or outcomes); the resultant of any weighted combination or sacrifice and benefit (determined and expressed either rationally or intuitively); or an aggregation, over time, of any or all of these.'

An alternative model for customer value is the means-ends model which proposes that the prospective customer has an end in mind and that he will purchase the product under consideration if he believes it will act as a means to him coming closer to achieving the end he has in mind (Huber, Herrmann & Morgan 2001). Ends may be personal values important to consumers and ends-means theory seeks to understand how choices made by consumers enable them to achieve those ends (Khalifa 2004). The consequences of a purchasing decision may be immediate or delayed and may be caused in part by the reactions of others to the consumer's purchasing decision. These consequences may be positive or negative in relation to the consumer's goals and, according to means-ends theory, the consumer seeks to maximise positive consequences and minimise negative consequences.

Holbrook (1996) defined value as 'an interactive relativistic preference experience' and developed a framework of three dimensions to describe customer value: extrinsic vs intrinsic, self vs other oriented and active vs re-active. According to Holbrook, a customer values a product extrinsically when he values it as a means to an end, while he values it intrinsically when he values it as an end in itself. He values it in a self-oriented manner when he values it according to his own preferences and responses, while he values it in an 'other-orientated' manner when he values it for its worth to another entity such as family, friends, community or nation. An individual values a product in an active mode when he plays an active role in realising the value, while he values the product in a passive mode when the product imparts its value with no active participation by the customer.

Service systems have been defined as 'dynamic value co-creation configurations of resources, including people, technology, organizations, and shared information, connected to each other internally and externally by value propositions' (Sitaloppi & Vargo 2014, p. 1279). The well-known customer / firm perspective on co-creation has been expanded by taking a service systems view of the matter, in which value has been defined as 'an improvement in system well-being' where well-being is measured as 'a system's adaptiveness or ability to fit in its environment' (Vargo & Akaka 2012, p. 149).

Clearly value is a difficult concept to define partly because the nature of value is fluid (Lusch & Webster 2011), complex, multifaceted, dynamic and subjective (Sánchez-Fernández & Iniesta-Bonillo 2007). The term is often misused in the literature (Khalifa 2004, p. 646).

Therefore, it is important of clarify what we mean by the term 'value'. In this review 'value' means 'customer value' – the value created for customers by a firm or value co-created by the firm and the customer together. It means value-in-use rather than value-in-exchange (see section 3.3.5.5). It is value as perceived by the customer (Sánchez-Fernández & Iniesta-Bonillo 2007), an interpretation of value generally agreed upon in the literature (Khalifa 2004, p. 647). Customer value is distinguished from shareholder value, stakeholder value and resource value (Bowman & Ambrosini 2000). Value is taken to be multi-dimensional rather than uni-dimensional, a distinction drawn by Sánchez-Fernández & Iniesta-Bonillo (2007).

3.3.5.5 Aspects of value

In G-D logic, the value-in-exchange of product is measured by the price paid at the time of exchange. The value-in-exchange has been created by the firm, is embedded in the product and is purchased by the customer for a price. 'Value-in-exchange is the negotiated measurement offered and received (e.g., money and value proposition) among exchange partners' (Vargo, Maglio & Akaka 2008, p. 150).

As opposed to value-in-exchange which in some ways is a measurement of the potential or latent value of a product, value-in-use, sometimes known as value-actualisation (Gummesson 2007), takes place after the purchase of a product or service, and is the value extracted from the product or service (while in use) by the customer. This value-in-use is contextualised by the situation in which the provider resources (service) are integrated by the customer to improve the wellbeing of the customer.

Although the concept of customer experience is hard to define precisely (Carù & Cova 2003), value-in-use is related to the experience the customer has as he or she uses the purchased goods or receives the purchased service (Lessard 2014, p. 37). If the customer feels better off as she uses the product the value-in-use is positive, if she feels worse off, the value-in-use is negative (Grönroos & Voima 2013).

Thus when considering their offerings in terms of value-in-use, firms take into consideration the ability their customers will have to extract value from their products. The value extracted could be economic, but can also be social (Edvardsson, Tronvoll & Gruber 2011; Wetter-Edman et al. 2014) emotional, ethical and environmental (Nordin & Kowalkowski 2010).

Value-in-use differs from value-in-exchange in terms of the time taken to realise the value. Value-in-exchange is an instantaneous snapshot of the value a customer places on a product, while value-in-use is realised over time as the products is used (Grönroos & Voima 2013).

In one formulation of the foundation principles of S-D logic, the tenth principle states that 'Value is always uniquely and phenomenologically determined by the beneficiary' (Vargo et al. 2010, p. 375) which is said to imply that 'Value is idiosyncratic, experiential, contextual, and meaning-laden' (Vargo et al. 2010, p. 375).

Chandler & Vargo (2011) take this examination of context a step further. They define the context in which exchange takes place as 'a set of unique actors with unique reciprocal links among them' (Chandler & Vargo 2011, p. 40). They characterise context in terms of three levels: micro, meso and macro as well as a meta layer which represents the evolution of these three layers over time. They argue the complex interplay of these three levels of context is central to value creation.

Holttinen (2010) nuanced the idea of the context in which value is created by arguing that value-in-use is best viewed through the lens of a value creating practice, which she defines as 'a context-laden arena for value creation, integrating a specific socio-cultural, spatial, and temporal context; mental states and bodily activities of consumers.' (Holttinen 2010, p. 102). She argues that consumers' participation in

different practices are liable to change as circumstances change, but that the practices themselves are more stable and therefore offer a more natural basis upon which to segment a firm's offerings. Korkman (2006) also argued that understanding how customer value is created in practice was an important aspect of designing superior service offerings.

The items comprising a product bundle (Harris & Blair 2006) and the interrelationship between newly purchased and previously owned products (Solomon & Buchanan 1991) have also been seen as an aspect of the context in which the value of goods and services are created.

Helkkula, Kelleher & Pihlström (2012) describe the value creation process of the customer as value-in-experience. They argue that both lived and imaginary experiences pertaining to a purchased good or service play a part in value creation and that past and anticipated experiences also influence the value extracted from a purchased product. This understanding, that experience is central to the creation of value (Ramaswamy 2011), adds a further dimension to the concept of value-in-use and value-in-context. Drawing on the concept of the customer-dominant logic (C-D logic) Voima, Heinonen & Strandvik (2010) go even further in claiming that, as well as taking place in realm of customer experience, it may also take place subconsciously as the customer re-interprets his or her experiences in the light of recently acquired products.

3.3.6 Co-creation

3.3.6.1 Definition of co-creation

Just as we saw that the concept of value is difficult to define (see section 3.3.5.3), it should come as no surprise that a definition of value co-creation is equally elusive. McColl-Kennedy et al. (2012) identify no less than twenty-seven different definitions of co-creation in the literature.

Customer participation has been defined as 'the degree to which the customer is involved in producing and delivering the service' quoted in (Matzler et al. 2005). Others emphasise the interaction (Grönroos & Voima 2013, p. 8) and the pro-sumption aspects of co-creation (Xie, Bagozzi & Troye 2008).

Customers may have hedonic, utilitarian (Stock, Oliveira & Von Hippel 2013) or intrinsic (Raasch & Von Hippel 2013) motives to co-create.

3.3.6.2 Taxonomy of co-creation modes

Zwass (2010) proposed a taxonomy of different modes of co-creation and co-production. The main divisions he identifies are *sponsored co-creation* in which individuals co-create in cooperation with firms and *autonomous co-creation* in which individuals co-operate with each other to create products of value. Customer value co-creation patterns such as independent bargain-hunter, comprehensive help

seeker, engaged problem solver, technology-savvy networker, and self-reliant customiser have also been identified (Moeller et al. 2013).

3.3.6.3 Prerequisites for co-creation

Prahalad & Ramaswamy (2004) and Ahrar & Rahman (2014) identify four prerequisites for successful co-creation: Dialogue, Access, Risk Assessment, Transparency (DART). They argue that the firm and the individual need to be able to communicate effectively in order to co-create, they must have access to the necessary resources, they both must be aware of and accept the risks involved with co-creation and have sufficient transparency in order to trust each other enough to co-create. Engagement (Jaakkola & Alexander 2014), interaction, self-service, and experience have also been identified as important elements of co-creation (Ranjan & Read 2014).

3.3.7 Other aspects of co-creation

3.3.7.1 Service eco-system

In further developing S-D logic, it was proposed that the market can be conceived as a service eco-system (Akaka, Vargo & Lusch 2013; Chew 2013a; Vargo & Lusch 2011) or a series of service-to-service transactions. The concept of a separate producer of value and consumer of value was supplanted by the view that all actors in the market are engaged in the process of (co) creating value for themselves and each other (Vargo & Akaka 2012; Vargo, Maglio & Akaka 2008; Wieland et al. 2012).

3.3.7.2 Co-production

Co-production is one of the main modes in which individuals and firms co-operate to produce value (Ranjan & Read 2014). Co-production is the process of involving the customer in the creation of value which is traditionally been done by the firm (Lengnick-Hall, Claycomb & Inks 2000). It involves encouraging customers to take on some of the value creation tasks which precede the usage of the products and which were previously performed by the firm. Tasks such as self-service at a supermarket checkout or construction of flat pack furniture fall into this category. Co-production has been modelled as a five stage process (Etgar 2008) in which the customer gains motivation to co-produce, estimates the costs and benefits of co-producing, engages in the co-productive act and evaluates the output from the co-production engagement. Co-production can also involve the customer co-designing the value proposition with the provider (Ranjan & Read 2014).

3.3.7.3 Co-creation in OOTF

Recent research (Chew 2013a) into value co-creation in organisations of the future (OOTF) conceptualise how OOTF could co-create inter-organisation value with both customers and collaborating firms by leveraging IT affordances.

3.3.7.4 Co-creation in new product development

In addition to co-creating value through the use of existing products, customers can also be engaged in co-creating value through participating in the development of new goods (Hoyer et al. 2010; von Hippel, Ogawa & PJ de Jong 2011) and services (Chew 2013b). In many cases, customer self-create new services which are then adopted by the firm (Oliveira & von Hippel 2011). Recent research has examined the connection between S-D logic and service science to explore how S-D logic can influence design of services (Wetter-Edman et al. 2014).

3.3.8 Conclusion

3.3.8.1 Summary

This dissertation adopts the conception of S-D logic outlined by Grönroos & Voima (2013). I find their account of value creation by both firm and customer to be logical, explanatory and meaningful. Their observation that treating co-creation as an all-encompassing, all pervasive activity between firm and customer obfuscates the locus and nature of co-creation is valid.

Having reviewed the literatures, my understanding of value co-creation and its components is as follows.

Value creation is the act of producing a good or service which, when used by the customer, will influence his or her state of mind. Positive value leaves the customer better off in some way. Negative value does the opposite. 'Better off' can mean an increase of peace of mind, making life easier, solving a customer's problems, letting the customer achieve more, satisfying customer needs or relieving the customer of some responsibility (Grönroos & Voima 2013, p. 3).

Firms, acting on their own, can create *potential value*. They produce products which they propose, through *value propositions*, to the customer. The firm controls phase of (potential) value production, and may, if it chooses, invite the customer to participate as co-producer. Products when created by a firm have the potential to leave the customer better off. The degree to which the customer accepts the value proposition is quantified as the price paid at the point of exchange. This is known as *value-in-exchange*.

That potential is realised as the customer uses the goods or consumes the services, at which time the customer experiences the value as *value-in-use*. The customer controls the value-in-use phase of value production (especially in the case of goods) and may invite the firm to participate as value co-creator.

The fact that the customer realises the potential value of the product as value-in-use is important because it emphasises that firms must make their offerings easily integrated into the customer's daily life and also into their future needs.

At the point of interaction between firm and customer there is the opportunity for co-creation of further potential value. The firm may learn more about the customer's preferences, which can then be incorporated into future goods – thereby making the products more valuable to the customer.

The firm and customer co-create value-in-use at the point of interaction, particularly in services. The co-creation often takes place through resource integration, 'the integration of resources from several actors in accordance with their expectations, needs and capabilities' (Mele, Spina & Colurcio 2010, p. 62). As the customer realises the potential value as value-in-use the firm may be invited by the customer to participate in that experience and has the opportunity to increase value-in-use through well managed interactions with customers.

S-D logic has placed more importance on the context in which value is created than G-D logic (Vargo 2009). Value can be considered as a standalone entity belong to one particular product or can be thought of value in relation to the value of the products the buyer already owns, sometimes known as a 'resource integration view of the service-dominant logic' (Grönroos & Voima 2013, p. 3).

3.3.8.2 Implications

In the context of the commodification of personal data, value co-creation is important in a number of ways.

First we note that in the literature it is the firm which creates the potential value in the form of a product and which offers it to individuals as a value proposition to be realised as value-in-use. However, when creating a personal data product, it is the individuals who will produce the raw materials (in the form of their data exhausts), it will be the infomediary that aggregates these raw materials into value propositions which will be offered to companies to use in their marketing and product development. So while we traditionally read about B2B and B2C companies, the personal infomediary will be a C2B company.

Value co-creation in new product development (section 3.3.7.4) will be particularly important in developing personal data products. Both individuals and companies will need to have input into new personal data products. Individuals will need to agree to share their data to create the product, and companies will need to indicate which data products will be useful to them. Companies will integrate these products (resources) with their own resources when they consume the data products to co-create value for themselves.

3.4 Personal Information Markets

There is a wide range of papers that touch upon elements which are relevant to a PPIM such as: data ownership (Hoeren 2014; Mashhadi, Kawsar & Acer 2014), privacy (Nissim, Vadhan & Xiao 2014;

Schunter 2014), personal informatics (Li, Dey & Forlizzi 2010; Li et al. 2012), collection and modelling of personal information (Kyriacou & Davis 2008; Shimojo et al. 2010), data security (Sugumaran, Murugan & Kamalraj 2014; Thapliyal, Mandoria & Garg 2014) and theories of co-creation (Siltaloppi & Vargo 2014; Solli-Sæther & Flak 2014). But there is a paucity of research that directly addresses the question of a need for and design of a PPIM. Here we outline the research that addresses issues closely related to the subject of this chapter.

Spiekermann & Novotny (2014) studied the research question of how individuals can share more personal information with companies while maintaining their personal privacy. They proposed a four-space model of personal information management by focusing on the type of relationship which exists between the individual and the company. Their focus was on increasing transparency and efficiency in the personal information market rather than allowing individuals to obtain financial compensation by selling access to their personal information.

Focusing on issues of privacy preservation, Schwartz (2003) proposed five elements which he claimed are necessary for a workable model of propertised personal information. Rather than focusing solely on a market mechanism, Schwartz also considered the legal and legislative frameworks needed to enable such a model. Aperjis & Huberman (2012) proposed compensating individuals for allowing their data to be included in a sample from which the buyer sought to infer a statistic of interest about a population and Laudon (1996) argued that while the collection and use of personal information has many benefits, there is a need for a market mechanism to protect individuals' privacy and allow them to gain a financial benefit from the use of their data. He proposed a National Information Market and a National Information Accounts Clearinghouse to facilitate this but claimed these would primarily apply to secondary and tertiary uses of personal information.

This review of the literature shows that while scholars have addressed the issue of compensating individuals for access to their personal information in a variety of ways, very little research has been done on the design and implementation of a primary market in personal information.

4 Why personal infomediaries failed

4.1 Introduction

This chapter examines each of the five types of personal infomediaries identified by Farrelly (2014). Drawing on transaction cost theory, the value proposition of each type of infomediary is analysed and then this analysis is used to develop the reason behind the failure of each of the personal infomediary types. The chapter concludes by identifying the type of personal infomediary which, with necessary changes to take in account recent developments in technology and societal attitudes towards privacy, is most likely to be successful.

On-going developments in mobile and wearable computing increasingly facilitates the collection of personal behavioural information (Gualtieri 2013; Li, Dey & Forlizzi 2010; McFedries 2013). It is well known that these data are commercially valuable (Porter & Millar 1985) but currently there is no market mechanism whereby individuals can be financially compensated for the data they generate as they go about their daily lives. Building on a taxonomy of personal infomediaries as outlined by Farrelly (2014) this chapter aims to take another step towards addressing the question as to why this is the case, and what can be done to remedy the situation.

This chapter examines the value proposition of each type of personal infomediary, first from the point of view of the customer and then from that of the vendor. It describes the means by which market transactions are facilitated by each type of personal infomediary and explains why each type has failed to materialise successfully.

In section 4.2 the theoretical framework is outlined within which the value proposition of each of the five types of personal infomediaries is analysed. Having introduced personal infomediaries in section 4.3.1, I analyse the value proposition of each of the five types of infomediaries and explain the reasons they failed in sections 4.3.2 through 4.3.5 (treating brokers and bonus brokers in the same section).

4.2 Theoretical framework

Drawing on the relevant aspects of transaction cost theory (Coase 1937; Williamson 1981) I propose the five types of personal infomediaries in terms of the primary role of intermediaries in general and infomediaries in particular – namely to facilitate market transactions by reducing external transaction costs.

Transaction cost theory analyses businesses in terms of the cost of making economic exchanges or transactions. It proposes that a business will expand to include those transactions that can be executed more efficiently within the company than between companies. Applying transaction cost theory to determine the legitimate function of an intermediary, Sarkar, Butler & Steinfield (1995, p. 7) state that:

Intermediary functions that benefit consumers include assistance in search and evaluation, needs assessment and product matching, risk reduction, and product distribution/delivery. Intermediary functions that benefit producers include creating and disseminating product information and creating product awareness, influencing consumer purchases, providing customer information, reducing exposure to risk, and reducing costs of distribution through transaction scale economies.

We will use each of these functions to assess the feasibility of the five types of personal infomediaries.

4.3 Analysis of personal infomediaries

4.3.1 Introduction

This section introduces five types of infomediaries and discusses possible reasons why each of these five types of personal infomediaries have failed to successfully commoditise private data on a large scale.

Before examining each of the five types of infomediary in turn, I first address questions that apply to all five types of infomediary. The so called ‘tech wreck’ of 1999 – 2000 has been well documented (Lowenstein 2004). It is possible that the tech wreck is the sole cause of the failure of personal infomediaries. However, a few infomediaries such as Amazon.com (admittedly not personal infomediaries), did survive and so it is not possible to entirely attribute the failure of these business to the tech wreck. However, I am interested here in discovering if there was something inherent in the business model of the personal infomediaries that contributed to their failure. Thus I aim to theorise not on the cause of failure per se, but to theorise on those aspects in their business models that contributed to their failure.

4.3.2 Guardians

A Guardian Personal Infomediary is a personal infomediary which has the primary aim of protecting individuals’ personal information and identity from the view of companies and other individuals. Examples of such Guardian Personal Infomediaries are @YourCommand, PrivacyBank, Enonymous, Jotter, Popular Demand and Privada.

Sarkar, Butler & Steinfield (1995, p. 7) list five services that intermediaries often provide to consumers: assistance in search and evaluation, needs assessment, product matching, risk reduction and product distribution/delivery.

Of these five services Guardians primarily only assist with one: risk reduction. They do so by protecting the individual’s privacy while engaged in e-commerce. Privacy is crucial to people in a number of ways. It is good in itself, it is inherently valuable for its own sake and many people do not wish to give it up, even in exchange for the convenience of online shopping or other online activities. People also want to keep their personal information private because they fear identity theft or online fraud. So an infomediary such as a Guardian which can protect individuals’ privacy can provide a genuine and desirable service to individuals and thereby encourage individuals to conduct more of their transactions online. In this sense the Guardian removes a psychological barrier to making transactions, the fear of losing one’s privacy.

From the vendor’s point of view, Guardian provided very little benefit. Of the five services which intermediaries often provide to vendors as listed in Sarkar, Butler & Sarkar, Butler & Steinfield (1995,

p. 7) none are performed by a personal infomediary which primarily focuses on protecting an individual's private data. In fact, Guardians made it harder for vendors to know their customers and to drive repeat business with them.

Furthermore, Leickly (2004, p. 10) claims that the lack of incentive for vendors contributed to the failure of the Guardians:

In controlling private information ... (the infomediaries) ... sought to provide support for the customer alone. In so doing, the model failed to create enough incentives for businesses to buy into the solution.

Another reason was that individuals did not trust that their personal information would be securely stored and safeguarded. This was not just a matter of the discovery of a number of minor technological flaws which could be easily remedied. A number of serious security breaches occurred in infomediaries (Lemos 2001; Slemko 2001). Featherman & Pavlou (2003, p. 456) found that concerns about privacy risk had a negative impact on the adoption of online e-services, so it is plausible that privacy breaches contributed to the failure of the Guardian personal infomediaries.

Furthermore, overcoming security breaches is not simply a technical matter of better firewalls, improved encryption and more vigilant intrusion detection. Anderson (2001) has argued that there are strong micro-economic forces such as network externalities and asymmetric information, which work against the securing of information and that information breaches will almost certainly continue to occur. This is an additional factor which further undermines the Guardian's business model.

4.3.3 Facilitators

Facilitators aimed to reduce friction in e-commerce transactions by providing the convenience of auto login and automated filling of online forms.

From the customer's point of view the benefit was supposed to be more efficient online shopping and sign-ins to secure websites. However, in general, intermediaries do not provide a service which increases the efficiency of the acquisition process (Sarkar, Butler & Steinfield 1995).

From the individuals' side, clearly the incentive of auto login between sites was not sufficient to induce individuals to share their personal information, especially when serious security flaws emerged in Passport (Lemos 2001; Slemko 2001). Concerns were also raised about the security of DigitalMe (Hancock 1999).

From the vendor's point of view, the facilitator aimed to make it easier for individuals to use their sites and thereby increase online transactions. They also aimed to increase customer loyalty by making e-commerce transactions more straightforward.

However, there was a fairly complicated and protracted sign up procedure for companies to join the Passport program. For example, to register for passport a vendor needed to sign a contract, pay a US\$1500 compliance testing fee and a US\$10,000 provision fee (Kim-Kwang Raymond 2006, p. 219). There was also widespread concern about handing over identity management to a company. Besides the possibility of security breaches, there were concerns about how the company may use individuals' digital personae (Luening 1998; Sundsted 2002).

4.3.4 Brokers and bonus brokers

These are the personal infomediaries for which the main aim is to commoditise personal information by matching vendor offerings to individuals based on their personal profiles and online behaviour. Infomediaries such as Privaseek / Persona, Lumeria and InterOmni fall into this category.

The benefit these infomediaries provide to the customer is twofold. They provide financial compensation for loss of privacy involved with sharing one's data with a company, and they provide a more relevant stream of advertisements to the individuals. They filter out irrelevant ads and allow the individuals to make better use of his or her search time. One of these functions (product matching) is traditionally fulfilled by intermediaries (Sarkar, Butler & Steinfield 1995).

From the vendor's point of view, brokers provide a stream of high quality leads. The theory is that by matching the behaviour and preferences of individuals to the products offered by the vendor, the broker can identify individuals who are likely to purchase certain goods and services at certain times.

However, there are three ways in which we can conceive the provision of financial compensation for access to private data. We can view it as an infomediary role, in which the reluctance for individuals to provide access is an impediment to online transactions and the role of the infomediary is to overcome this by providing financial compensation.

Another way to conceive of this service is in terms of the commoditisation of the individual's privacy. The individual's privacy can be viewed as a possession of value which is owned by the individual. By giving access to one's personal data, one's privacy is diminished and the individual receives financial compensation for that diminution in privacy.

A third way to understand the provision of financial compensation for access to private data is to focus on the manufacture of new data products from the raw materials of the individual's personal data. In this conception we put the emphasis on the value added by the infomediary by creating a new data product, rather than the bringing together of buyer and seller.

One of the reasons the brokers failed is that there was not enough value in the service they provided to the companies. Companies sought not simply access to personal information but business value in the

form of actionable insights which could be derived from those data. Mere access to individuals' personal data did not justify the effort required by companies to engage with the infomediary.

Technology and society have changed in important way since the late 1990s. Individuals now generate far more personal data, and the data they generate is of a much more detailed, granular nature. See for example (Facebook Inc 2012; International Telecommunication Union 2017). Society is also a lot more comfortable with sharing their lives online (Debatin et al. 2009). It could well be that in light of these changes, applying the manufacturing conception of the personal infomediary broker is now a viable business model. This possibility is beyond the scope of the present paper but could be a fruitful line of enquiry to pursue in further research.

4.3.5 Incentivisers

This group of personal infomediaries also have as their main focus the commoditisation of personal information but they differ from the brokers and bonus brokers in that they offer financial incentive in the form of pay-per-click, pay-per-hour or entry to a lottery to encourage individuals to spend more time online and thereby generate additional personal behavioural information which can be collected and leveraged.

From the individual's point of view these companies offer financial compensation and a targeted stream of advertisement. From the vendor's point of view, the incentivisers offer a greater number of customers who click through on certain ads and who spend more time viewing these ads.

Incentivisers failed because clicking on or viewing ads did not convert into sufficient sales to generate enough cash flow to pay meaningful compensation to the individuals. Incentivisers soon resorted to offering entry to lotteries instead of cash inducements as a way to recue outgoings, but this was not successful (Saunders 2000, 2001).

4.4 Concluding remarks on the viability of Infomediaries

In this chapter we have seen that, despite providing risk reduction services to individuals, the Guardians failed because they did not provide enough value to vendors. Facilitators failed because they did not provide one or more of the main services which the market expects from an intermediary. Incentivisers failed because of a flaw in their underlying business model while brokers and bonus brokers failed because they focused too much on providing benefits to consumers while paying too little attention to the needs of companies. They also focused too much on the bringing together of buyer and seller and too little on the manufacture of new and useful data products from the raw data collected from individuals.

This implies the one aspect of successful PPIM could be the commoditisation of personal information by matching vendor offerings to individuals based on their personal profiles and online behaviour.

However the design of the market must be such that it ensures sufficient value to all groups of participants on the market.

An initial design of a two-sided market in personal information is the subject of the next chapter.

5 PPIM conceived as a two-sided market

5.1 Introduction

In order to define the objectives of a PPIM we will conceive of the entity which facilitates the PPIM as an *information intermediary*. This entity will be known as the platform owner. Rose (1999) defines an intermediary as: ‘an independent, profit-maximizing economic agent mediating between two markets sides in presence of market imperfections’ (Rose 1999, p. 51). An information intermediary, or infomediary, is therefore an entity that mediates between two sides in an information market.

However, most researchers have studied infomediaries in the context of companies selling their goods to individuals – a setting which I will term ‘traditional’. In this chapter, the aim of the intermediary is to collect and aggregate individuals’ personal information and allows companies to purchase access to it. So in this situation the individuals are the vendors and the corporation is the consumer.

Table 8: Comparison of entities which fulfil vendor and consumer roles in traditional and PPIM contexts.

	<i>traditional</i>	<i>PPIM</i>
<i>vendor</i>	corporation	individuals
<i>consumer</i>	individuals	corporation

Here we summarise the objectives for successful infomediaries which have been put forward by a number of authors in traditional settings and then apply them in the context of a PPIM.

Hagel & Singer (1999) outline a wide range of possible activities which an infomediary may undertake. They describe an infomediary as a company which captures, manages and maximises the value of personal information on behalf of individuals (pp. 33ff.). They list the key functions of an infomediary as the collecting, aggregating, selling and protecting of personal information, matching of individuals’ needs with vendors and negotiating with vendors on behalf of individuals, via reverse auctions and using combined market power to obtain better offers.

Leickly (2004) lists the prerequisites for a successful infomediary as follows:

- provides value for both businesses and consumers
- works in a specialised industry which has high search cost and has high information asymmetry due to complex products

- deals in information commodities
- does not compromise consumer information and privacy.

Rose (1999) lists a number of critical success factors for information intermediaries: provision of personalised service, provision of value-adding services in addition to the services of information mediation and a focus on a thematic domain, i.e. providing a service in a specialised sector of the market (Rose 1999, p. 162).

Grover & Teng (2001) argue that infomediaries are set to play an increasingly important role in e-commerce and that there are seven main objectives for a successful infomediaries should perform (Grover & Teng 2001, p. 80): search facilitation, matching consumers with vendor products, content services such as independent reviews of products and community services such as allowing customers to share and discuss opinions.

Bakos & Bailey (1997) conducted a survey of thirteen e-commerce firms in order to examine the changing roles of intermediaries. They identify the main roles played by successful intermediary as: aggregating buyer demand and/or supplier offerings, matching suppliers and customers, building trust (between buyer and seller) and providing inter-organisational market information.

Table 9: Summary of objectives of traditional infomediaries

	Hagel & Singer (1999)	Rose (1999)	Leickly (2004)	Grover & Teng (2001)	Bakos & Bailey (1997)
Collecting and storing personal information	✓				
Selling access to personal information	✓				
Protecting privacy of individual	✓			✓	
Negotiating better deals on behalf of individuals	✓				
Matching requirements of individuals with products offered by vendors	✓	✓		✓ ⁵	✓
Aggregating buyer purchasing power	✓				✓
Reduce search cost by aggregating vendor offerings	✓	✓	✓	✓	✓ ⁶
Reduce uncertainty		✓	✓	✓ ⁷	✓ ⁸

Table 10 explains how the objecting of a traditional infomediary would apply in a PPIM setting.

Table 10: Objectives of infomediaries in a traditional and PPIM context

Traditional setting	PPIM
---------------------	------

⁵ through informational services

⁶ through inter-organisational market information

⁷ through reviews and community comments

⁸ by providing trust

Collecting and storing personal information	1	Collecting and storing personal information
Selling access to personal information	2	Selling access to personal information
Protecting privacy of individual	3	Protecting privacy of individual
Negotiating better deals on behalf of individuals		N/A
Matching requirements of individuals with products offered by vendors	4	Matching requirements of companies with (data) products offered by individuals
Aggregating buyer purchasing power		N/A
Reduce search cost by aggregating vendor offerings	5	Reduce search cost by aggregating individuals offerings
Reduce uncertainty	6	Reduce uncertainty

The objectives listed in Table 10 are now applied in the context of a PPIM.

5.2 Design and development

In designing the PPIM, we will draw on the theory of market engineering which has been described as ‘the process of consciously setting up or re-structuring a market in order to make it an effective and efficient means for carrying out exchange transactions’ (Weinhardt & Gimpel 2006, p. 6). Here I will discuss three aspects of a market which need to be designed in order to meet the objectives in Table 9: the transaction object, the market microstructure and the IT infrastructure (Weinhardt & Gimpel 2006) that we will design.

5.3 The transaction object

It is not proposed that an individual generates personal information and then sells that data to another party who then owns the data in its entirety and the individual who previously owned it relinquishes control over those data.

It is proposed that the transaction object in a PPIM is the permission to use a subset of a data contributing individual’s personal information for specified purposes for a specified time.

The permitted usage of personal information allowed on a PPIM will exclude the transfer of raw data from the central data repository to the company’s databases. Only the query that had been bid for and for which permission has been granted will be able to be run against the customer data.

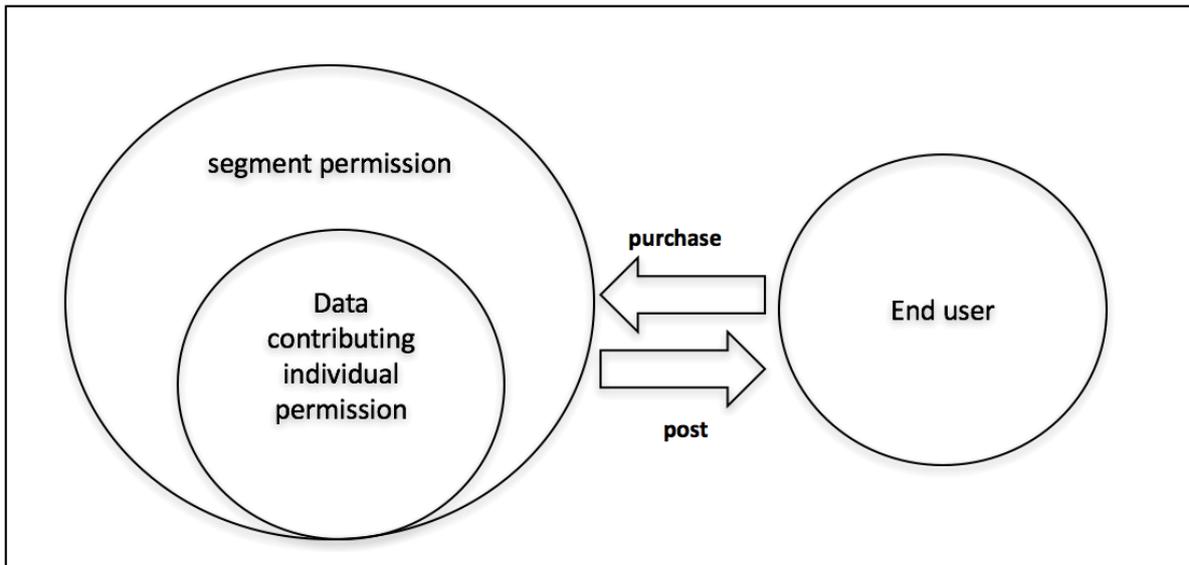


Figure 4: Data contributing individual initiated transactions

The transaction object

The PPIM will accommodate two type of market transaction: data contributing individual initiated and end user initiated. The data contributing individual initiated transaction will be a simple catalogue type transaction (Jennings, Ockenfels & Weinhardt 2008, p. 2). Data contributing individuals can post the fixed price at which they are willing to allow access to certain aspects of their personal information. End users can search the database for a market segment to which they wish to have access, ascertain the price and purchase access should they wish to do so. See Figure 4.

The end user initiated transaction will be a negotiation type transaction. An end user may initiate a request for access to data generated by a market segment of interest at a given price. Data contributing individuals may accept or reject a bid or negotiate a mutually agreeable price. See Figure 5.

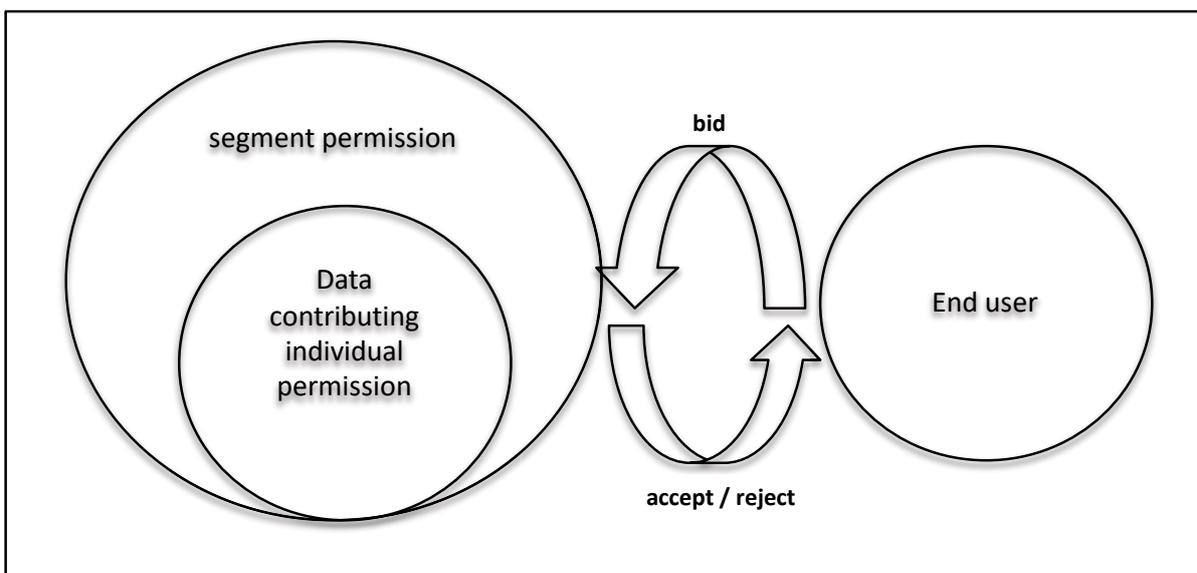


Figure 5: End user initiated transactions

5.4 The IT infrastructure

The IT infrastructure to facilitate a PPIM will be based on a capture-at-source distributed synchronisation system as portrayed in Figure 6. As a data contributing individual creates personal information on a given device, a locally installed app will capture and store the personal information on the device (1).

Once the device is connected to the central data repository, personal information will be synchronised (2) and the locally stored copy of the data will be deleted.

The central data source will consist of a massive parallel processing (MPP) distributed computing platform (3), comprising both a database and a distributed file system for the storage of non-text data.

The database will be structured around a personal information logical data model (4) which will be extensible in order to accommodate new streams of personal information as they emerge.

Management of the personal information permissions will be executed via two dashboards – one for data contributing individuals and one for companies (end users). Data contributing individuals will manage permission they have granted, receive payments, and receive bids for access to their personal information via the data contributing individual PI dashboard (5).

Companies will view aggregated views of the available personal information which is available for them to bid for access to, make payments, submit bids, view the permissions they have purchased (and the expiry dates and renewal costs of those permissions) via the company PI dashboard (6).

In effect, the company managing the centralised data store (the PPIM owner) will provide a personal information accounting system, which will accurately account for the personal information which

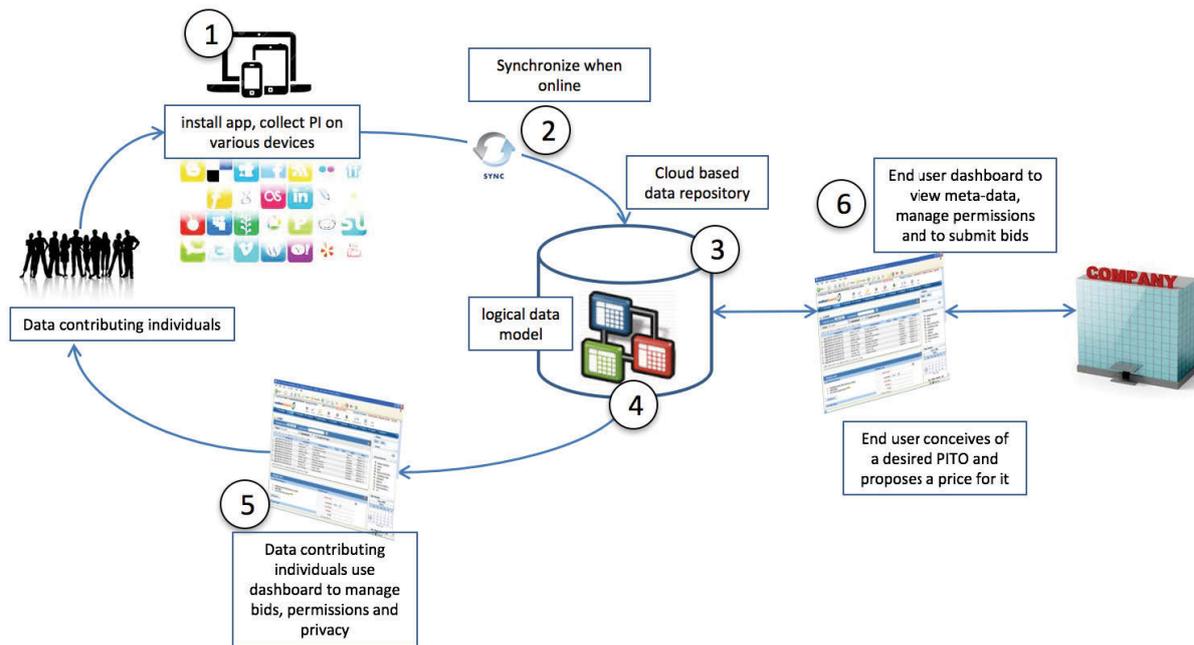


Figure 6: The IT infrastructure to support a PPIM

has been accumulated by each data contributing individual and the permissions which have been bought and sold for access to that data.

In the market described in this section, the locally installed app and the central data repository will fulfil the objective of collecting and storing personal information. The transactions shown in Figure 4 and Figure 5 will fulfil the objective of selling access to personal information. The transparent trading of permissions will fulfil the objective of protecting the privacy of data contributing individuals. The personalised nature of the data contributing individual PI dashboard will fulfil the objective of matching requirements of companies with data products offered by data contributing individuals by allowing them to only offer up those permissions they are willing to sell. The end user PI dashboard will fulfil the objective of reducing search cost by aggregating data contributing individuals' offerings. The company PI dashboard will also fulfil the objective of reducing uncertainty by providing transparency around which permission have been purchased by the end user.

5.5 Demonstration

In this section we demonstrate how the PPIM model would operate in three application scenarios.

Application scenario 1: Buyer initiated purchase for marketing to new customers

Joe's Jazz Café in Manly wishes to increase patronage on Tuesday and Wednesday nights. Joe identifies his market segment of interest as adults who live in Manly, who like jazz, eat out more than once a week, and regularly use Facebook. Joe bids for and buys access to that market segment. Having

purchased permission, Joe markets live jazz and meal deals on Tuesday and Wednesday nights to this segment. A month later, Joe gets a report on how many of the individuals who received the ads subsequently purchased a meal and how much they paid. This report is based on mobile phone location data and credit card transaction data.

The data product bought and sold in this scenario comprises the following elements:

- Permission to query the database to identify data contributing individuals who fall into the specified market segment
- Permission to push ads to those individuals' Facebook pages
- Permission to query those data contributing individuals' credit card records for transactions which took place at Joe's Café

Application scenario 2: End user initiated transaction to gain insight into existing customers

The ABC Bank of Australia wishes to know which of its existing home loan customers are putting their homes on the market and may be seeking to re-mortgage their homes with other banks. They bid to access to the desktop and mobile web browsing history of a list of existing customers. The bank receives a report showing which customers are spending a significant amount of time visiting real estate sites. It contacts these customers in an effort to retain their business.

Application scenario 3: Data contributing individual initiated transaction - ex-ante data collection and co-production

A number of customers have posted an acceptable price at which they are willing to sell access to the fitness data as captured by their Fitbit wristbands and other similar devices. The Runwell Shoe Company identifies a market opportunity to design customised running shoes. It purchases permission to access to data contributing individuals' fitness data. Runwell then collects and analyses detailed data about the running habits of consumers and designs shoes to meet the requirements of customers in certain micro-sectors of the market.

5.6 Evaluation

In the absence of a working PPIM on which to conduct experiments, the *informed argument and scenario method* of evaluation (Hevner et al. 2004, p. 86) is the most applicable method of evaluation.

In each of the scenarios listed above data contributing individuals' personal information has been collected and stored (objective 1), various permission were traded on the PPIM (objective 2), the identity of the data contributing individual was protected (objective 3), the end users were able to conduct a transaction with individuals who were willing to share the personal information they sought

(objective 4), the search cost to the end user was reduced by the PPIM (objective 5) and the end user could monitor the permission they have purchased via the business PI dashboard thereby reducing the uncertainty regarding the personal information access they could legitimately utilise (objective 6).

5.7 Conclusion

The management of personal information presents both one of the great challenges and great opportunities of the 21st century. In this chapter I have argued that advances in technology and data collection capabilities have created an opportunity to develop a Primary Personal Information Market (PPIM). It is argued that an increasing volume of valuable personal information is being generated and collected and that the individual at the centre of that personal information generation process is both the owner of those data and the party best placed to augment its value by joining multiple personal information streams into a centralised personal information repository. Drawing on the principles of design science and market engineering, this chapter then outlined a vision of a workable permission based PPIM.

It is proposed that the permission to query a subset of a data contributing individual's personal information be the tradable asset in a PPIM and I have outlined a market structure and an IT infrastructure which would facilitate this market.

The exact nature of that transaction object which would embody that permission is the subject of the next chapter.

6 The Transaction Object on a PPIM

One important element of any market design project is the design of the *transaction object* (*Negotiation, Auctions, and Market Engineering* 2006, p. 6) – the object which is to be bought and sold on the market. A market cannot operate if there is no transaction object. Furthermore, defining the transaction object is critically important because for a market to operate there must be market thickness, defined as the ability of a market to 'attract a large enough proportion of the potential participants' (Roth 2008, p. 79: 79) and thickness in a market is influenced by the design of the transaction object (*Negotiation, Auctions, and Market Engineering* 2006, p. 6).

There is an extensive body of research into areas which are relevant to the design of a workable PPIM such as: data ownership (Hoeren 2014; Mashhadi, Kawsar & Acer 2014) and privacy (Nissim, Vadhan & Xiao 2014; Schunter 2014), personal informatics (Li, Dey & Forlizzi 2010; Li et al. 2012), the collection and modelling of personal information (Kyriacou & Davis 2008; Shimojo et al. 2010) and data security (Sugumaran, Murugan & Kamalraj 2014; Thapliyal, Mandoria & Garg 2014). However

‘there is a substantial lack of empirical research on ... markets for online personal data’ (Milyaeva & Neyland Forthcoming, p. 4) and most market design research is qualitative (Wang et al. 2008).

Spiekermann & Novotny (2015) studied the research question of how individuals can share more personal information with companies while maintaining their personal privacy. They proposed a four-space model of personal information management by focusing on the type of relationship which exists between the individual and the company.

Aperjis & Huberman (2012) proposed compensating individuals for allowing their data to be included in a sample from which the buyer sought to infer a statistic of interest about a population and Mashhadi, Kawsar & Acer (2014) propose a data market for the trading of sensor data generated by the Internet of Things.

A close reading of these papers shows that very little justification is given for the design choices made in the personal information markets which have been proposed by other researchers.

Other researchers have alluded to the need for a marketplace for personal information – without explicitly calling for a PPIM. For instance de Montjoye et al. (2014, p. 2) argue for the need for individuals to retain ownership of their personal information and for the establishment of a ‘fair and efficient market’ in personal information. de Montjoye et al. (2012, p. 2) argue that personal information stores will facilitate an ‘economically efficient’ personal information market.

Zannier (2013) tested the concept of selling his personal information. He collected detailed data on many aspects of his personal life and sold those data in exchange for contributions to his kick-starter project. He raised \$2,734 USD from 213 backers.

Sevignani (2013) points out the need for individuals to have more control over their privacy while Milyaeva & Neyland (Forthcoming) take a Science and Technology Studies approach and identify a growing trend in a number of recent start-ups which focus on the value of data usage rather than the value of data itself.

In summary, there is a growing body of research which testifies to the value of personal information, to the abundance and ubiquity of personal information, to the need for protecting privacy and to shortcomings in current personal information market practices. Research into the design of a Personal Information Transaction Object (PITO) which can be traded on a PPIM as a means to addressing these issues is not yet available. This chapter proposes and evaluates such a new PITO design.

6.1 Method

Since the design of PITO, a new artefact, is an integral part of the design of PPIM, another new (parent) artefact, we follow the Design Science methodology (Hevner et al. 2004; Peffers et al. 2006; Peffers et al. 2007-8). The atomisation stage of the Design Science process (Peffers et al. 2007-8, p. 52) is

achieved by drawing on Market Engineering theory (Weinhardt, Holtmann & Neumann 2003), a well-recognised research approach to the design of markets – see (Dunnebeil et al. 2012) for example.

This chapter seeks to clarify the nature of the transaction object which is bought and sold in the PPIM. To do so it identifies key drivers of market efficiency and assess the impact that the alternative PITO design decision will have on these drivers.

The market-engineering based PITO design is evaluated qualitatively (Yin 2013) using semi-structured interviews. These interviews were conducted with senior leaders in data related business roles, who were selected on the basis of their subject matter expertise. This is consistent with other studies; see for example (Karunakaran & Puroo 2012; Knol, Sol & Van Wamelen 2012).

The aim of the evaluation was to gather opinions from interviewees on the following key implementation factors: the feasibility of building a PPIM for trading permission-based PITO, the viability of such a PPIM if it were to be built, the relative merits of a permission-based PITO versus an ownership-based PITO, and the appetite end users and data contributing individuals might have to participate in such a PPIM.

The selection of interviewees was guided by theoretical sampling (Eisenhardt 1989) in which experts were sampled who could evaluate the constructs and structure of the proposed designs based on their domain expertise and years of experience (as indicated by the subjects' LinkedIn profiles). Some interviewees recommended other suitable subjects with applicable knowledge which led to clustered samples (Merkens 2004).

Interviewees included representatives in senior analytical roles from academia, government funded research organisations, a large federal government department, an Australian home loan brokerage company, and a major political campaign advisory company. On average, interviewees had 9.7 years of work experience in data-related roles. The interviews were digitally recorded and transcribed. The interviews lasted on average 30 minutes.

An interview topic guide (King 1994) was used for interview consistency. After a brief explanation of the purpose of the research and the proposed roles of PPIM, PITOS and PITO the interviewees were asked their opinions on the sensibility and feasibility of the proposed solutions. They were asked if they saw any major flaws or issues with the proposed solution. They were asked whether or not such a market would be viable and, based on their experience, whether there was appetite for companies to purchase detailed personal information on such a market and the willingness of individuals to share personal information on such a market. The results of these interviews are summarised and analysed in section 6.6.

6.2 Design criteria

In designing a Personal Information Transaction Object (PITO) we aim to make design choices which will improve the performance of the PPIM on which it is traded. Design choices are compared against market success criteria, for which there is an extensive body of research. Market quality has been described as ‘the extent to which a market satisfies its participants’ (Granados, Gupta & Kauffman 2005, p. 3). Commonly identified characteristics of high quality markets are: thickness, low congestion, safety, simplicity, transparency, immediacy and low search and transaction costs (Granados, Gupta & Kauffman 2005; Milgrom 2011; Roth 2008). They therefore represent the success criteria for a PITO design in a PPIM.

In order to design a working PPIM, we must design each element of the market so that it satisfies the needs of both companies who wish to access personal information and individuals who want to profit from sharing their personal information.

We use these characteristics as criteria to assess a design alternative for the transaction object in a PPIM.

6.3 Design parameters

This section examines the crucial design decision of ownership versus permission to use, which must be made when designing a PITO. It then assesses the alternatives in terms of the success criteria outlined in Section 6.2 and explores the preferred option in more detail.

The fundamental design decision to be made in a PITO is the question of ownership. Is the outright ownership of the personal information to be bought and sold, or does the ownership of the personal information remain with the individual who generated it in the first place?

Allowing the ownership of the personal information to remain with the individual is likely to contribute to market thickness because it will be more beneficial to individuals. For example, Schwartz (2003, p. 2056) argues that ‘limitations on an individual’s right to alienate personal information’ is a key aspect of successfully propertising personal information.

Furthermore, it is clear that companies can extract commercially valuable insights from personal information even when the ownership remains with the individual. The end user licence agreements of the major social network sites (Facebook 2014; LinkedIn 2014; Twitter 2014) which attest to this.

Since retention of ownership of personal information is beneficial to both individuals and companies it will contribute favourably to thickness in the market compared with the alternative design choice.

Allowing individuals to retain ownership of their personal information would also enhance safety. In a permission-based PPIM, individuals would have more control of the use to which their personal information was put.

A PPIM in which ownership was transferred from individual to end user would be simpler than the alternative as the concept of ownership is a less nuanced concept than that of permission. However, allowing individuals to retain ownership of their personal information would enhance transparency. By explicitly allowing or denying use of their personal information by certain companies, individuals would have a much clearer view of what was being traded on the PPIM. Spiekermann & Novotny (2015) have shown that in the absence of these explicit permissions an opaque shadow market in personal information has emerged. A transparent regulation framework also encourages individuals to share their information, adding to thickness in the market (Mashhadi, Kawsar & Acer 2014, p. 3).

The success factor of immediacy would not be significantly influenced by the design decision regarding ownership while the success criterion of search costs (from the perspective of companies searching for suitable permissions to fulfil their needs) would be negatively influenced by the added complexity of allowing individuals to retain ownership. Since permission is a multi-dimensional concept, the potential number of different products available on a permission-based PPIM is very large. Furthermore, the lack of a consistent language for describing permission could also contribute to additional search cost. The issue of search cost would be addressed by employing a consistent taxonomy of permissions.

In order to further clarify the proposed PITO it is necessary to define the permission which will be granted by the individuals to the end user. We propose three main groups of descriptors: *content*, *use* and *nature* of the personal information.

The *content* of the personal information is described in terms of behaviour descriptors, identity descriptors and attributes (Wiese 2013). Behaviours are descriptors such as travel, exercise and sleep patterns. Identity descriptors are descriptors of the data-generating individual such as name, age, gender, etc. Attributes are annotations (Rawassizadeh et al. 2013, p. 629) associated with the personal information, such as intention and preference descriptors which can be extracted and derived from personal information. Also included in the content of the personal data is the date range over which the personal information will be retrieved. Additional content characteristics of the PITO are dependent on the way in which the personal information is modelled and the semantic extraction (Albatal et al. 2013) of meaning from that information.

The *use* to which the personal information may be put includes the channels through which the individuals may be contacted (if at all) and the advertising channels through which the individuals may be marketed to (if any).

The *nature* of the personal information to which permission has been purchased is defined in terms of the *concurrency* of the permission: the maximum number of buyers which can simultaneously purchase access to that personal information. Clearly, exclusive access to personal information is more valuable

than access which is shared with many other buyers. It also comprises the *refresh frequency*, the frequency at which the personal information will be refreshed. Personal information may be accessed at rates ranging from a one-off batch load through to a real time feed. Finally, we define the *tenure*, the period of time after which the granted permission will expire.

Privacy meta-parameters

In addition to the design parameters described above, we propose a privacy meta-parameter. This acts as a check and balance to inform the individual of the privacy implications of the permission he or she has granted. It can be used as a descriptive metric to measure the degree to which privacy has been exchanged or it can be used as a setting to indicate to an individual whether or not to exchange further permissions. The privacy setting is calculated using approaches such as k-anonymity (Sweeney 2002), t-closeness (Li, Li & Venkatasubramanian 2010), l-diversity (Machanavajjhala et al. 2007) and uniqueness of traces (de Montjoye et al. 2013).

A comparison of the relative merits of a permission based PITO vs an ownership based PITO is presented in Table 11.

Table 11: Summary of design parameters’ relative impact on market quality

	Permission-based PITO	Ownership-based PITO
thickness	higher – beneficial to individuals and also of utility to businesses	lower – more useful to businesses but less attractive to individuals
safety	higher – enhanced by transparency	lower – individuals lose control of PI
simplicity	lower more complex, but manageable by a well-designed taxonomy	higher – less nuanced concept
transparency	higher – though retaining ownership of data	lower – limited visibility on how data is used
immediacy	similar	similar
search costs	higher – as concept is more nuanced	lower – as it is a simpler concept
transaction costs	similar	similar

6.4 Proposed design solution

Personal Information Transaction Object Space (PITOS):

From the design alternatives described in Section 6.3 above, we define the Personal Information Transaction Object Space (PITOS) as the set of all possible PITOs, i.e. the space from which individual PITOs can be drawn. More formally we define the PITOS as follows:

Let $n \in \mathbb{Z}^+$ be a positive integer denoting the number of permissions, each permission being an agreement between the data contributing individual and the end user. Let $P_i, i \in \{1 \dots n\}$ be the i^{th} permission in the PITOS. Let $n_i \in \mathbb{Z}^+, i \in \{1 \dots n\}$ be the number of settings in the i^{th} permission. Each permission comprises n_i settings: $P_i = \{S_{i1}, S_{i2} \dots S_{in_i}\}$ and the PITOS comprises the n permissions $\{P_1, P_2 \dots P_n\}$.

Clearly the number of data sources and the number of uses to which the data could be put is extremely high and it would not be possible to know in advance the nature of the PITOs for which companies will have a need. To address this issue a highly developed Personal Information Logical Data Model (LDM) is required to facilitate the creation of a very wide range of PITOs. Furthermore, late-binding, schema-on-read elements to the personal data store will provide the flexibility for the system to respond by creating novel PITOs from semi-structured data such as free text, XML and JSON data.

6.5 Personal information transaction objects (PITO):

A PITO is a single instance within the PITOS, defined by the setting of the design parameters.

For example, if an individual grants an end user permission for two years to query against her Facebook data, contact her by SMS and to refresh that data every 12 hours, these permissions will comprise the PITO represented in Figure 7.

```
"permissions": {
  "content": {
    "source": "Facebook"
  }
  "use": {
    "contact channel": "SMS"
    "advertising channel": "none"
  }
  "nature": {
    "concurrency": 1
    "refresh": 12
    "tenure": 2
  }
  "privacy": {
    "level": "High"
  }
}
```

Figure 7: A Simplified PITO

The proposed permissions based PITO's would be created and traded on the PPIM as per. Figure 8.

1. End user browses the PITOS (as described in the LDM metadata).
2. End user conceives of a PITO and posts a description and price on the PPIM.
3. Individuals whose data could contribute to the PITO are identified in the data store.
4. Offers are sent to individuals.
5. Individuals accept or reject the offer.
6. If a sufficiently large number of individuals accept the offer, the PITO is created and exchanged on the PPIM.
7. The permitted query is run against the data which comprise the PITO and the results are returned to the end user.
8. The end user puts the results of the query to use as the permitted in the PITO via the PPIM.

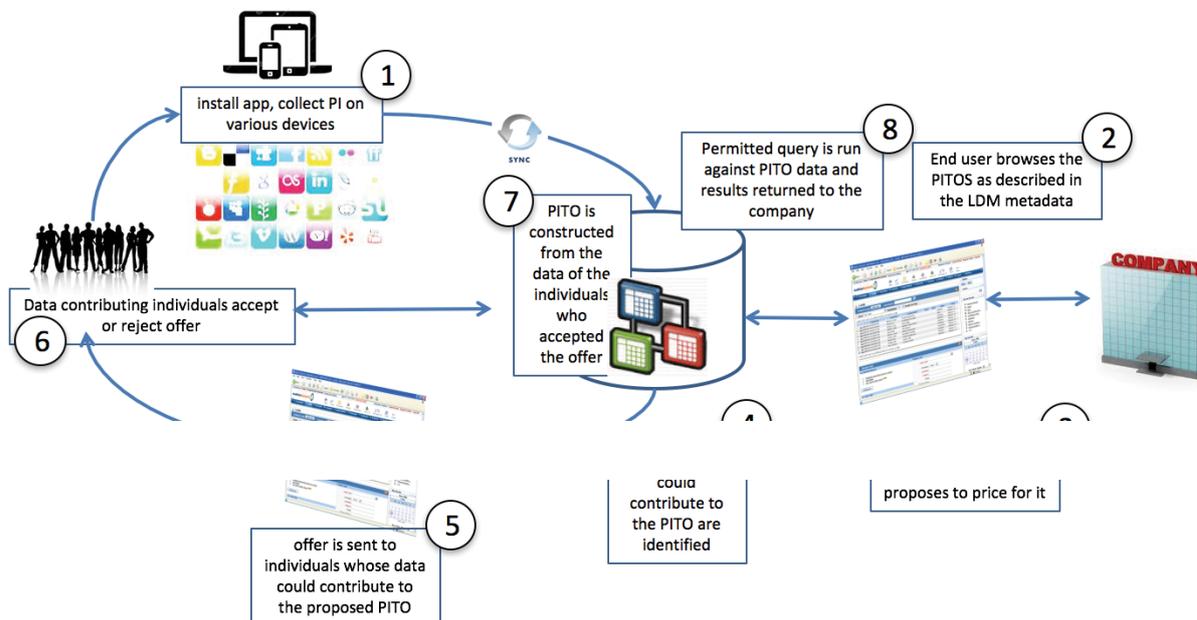


Figure 8: Creation and exchange of PITO's on a PPIM

6.6 Evaluation

Following Yin (2013), I identify, compare and contrast the main themes in the interviewees' responses. Table 14 contains a classification of the interviewee's opinions with illustrative quotations.

Of the seven experts interviewed, five were of the opinion that permissions-based PPIM was, broadly speaking, viable.

Five of the seven respondents thought that there would be significant appetite from companies to purchase permission to query detailed personal information with two interviewees mentioning the imperative companies now have to gain a 360° view of the customer. Only the expert from the political strategy advisory company was of the opinion that such a market would be of little interest to his industry. Three of the interviewees were of the opinion that individuals would be willing to sell permission to query personal information on the proposed market.

Interviewees were less emphatic on the question of a permission-based PITO versus an ownership-based PITO with only four proffering an opinion. These opinions are summarised in Table 12. Full transcripts of selected interviews can be found in section 14.1.

There were three main issues raised by interviewees (see Table 13). The question of pricing access to personal information and how best to gain a critical mass in such a market were the most commonly raised objections. The use of proxy measures to infer information about individuals beyond the permission which has been granted was also raised as a concern.

Table 12: Summary of interviewees' opinions

	Positive	Neutral	No opinion
Viability	5	1	1
Permission vs ownership	4	0	3
End user appetite	5	1	1
Individual appetite	3	1	3

Table 13: Issues and concerns raised by interviewees

Major Objections / Concerns	Count
gaining a critical mass	3

pricing	3
use of proxy measures	2
privacy	1

6.7 Conclusion

Following design science methodology (Peppers et al. 2007-8; Weinhardt & Gimpel 2006) and qualitative interview methods (Schmidt 2004) this chapter has respectively synthesised and qualitatively evaluated a design for a personal information transaction object (PITO) in accordance with the market design criteria espoused by Market Engineering theory (Weinhardt & Gimpel 2006) and extant market design practices so as to render our prior Primary Personal Information Market (PPIM) system-architectural design (Farrelly & Chew 2016b) technically feasible to implement. It proposes that a permission-based PITO is preferable to an ownership-based PITO as it maintains personal privacy, enhances data transparency and promotes market thickness – the prerequisite conditions for viable commercialisation.

This chapter reveals that the model proposed for a permission-based PITO and the resulting permission-based PPIM has the potential to be a viable approach to implementation and is therefore a worthwhile avenue of further research. A majority of the interviewed experts broadly agreed that there is a need for such a market and that if it could be designed and implemented there would be a potential high demand for such a market from a range of organisations.

This finding is significant, mainly due to the magnitude and wide-ranging implications of the problem being addressed. A viable PITO (and the resulting PPIM) would potentially affect the lives of millions of individuals in a digitalised world and dramatically change the ways business utilises the value of personal information.

Having proposed a design for both a two-sided market in personal information in chapter 5 and a design for the transaction object in this chapter, we now turn to the question of how willing individuals would be to participate on such a market. This question is addressed in the next chapter.

Table 14: Classification of opinions and illustrative quotations from interviewees' responses

Interviewee		Permission-based PPIM viability	Permission-based PITO vs ownership-based PITO	Companies' appetite to use a PPIM	Individuals' appetite to use a PPIM	Objections / concerns
Senior Director of Analytics in large Federal Government Department	opinion	positive	none offered	positive	positive	proxy measures
	illustrative quotes	'I think it's got merit'	N/A	'understanding the customer is probably the primary driver of the work in this organisation for the next so many years'	'I see a huge market for it'	'We're living in an era when you can use proxy measures to measure what it is you want to measure and therefore the public will always, not always, most of them will not know what it is exactly they are giving away or could give away'
Analytics Manager in Government Agency	opinion	positive	positive	positive	neutral	critical mass
	illustrative quotes	'sounds quite viable' and '... it's end to end possible technically'	'it seems rich, in possible use cases...'	'there's a demand for a lot more existing companies who have account holders, but they don't know enough about those customers, and they would love to be able to plug into this resource'	'I think that if you want the network effect you have to have a particular, narrow applications space where you can target people'	'I'm not sure whether the supply side can be easily done'
Chief Scientist in high tech research organisation	opinion	positive	none offered	positive	positive	pricing, critical mass, proxy measures
	illustrative quotes	'It's a truly important question ... I think this is absolutely fantastic'	N/A	'So in general my answer to your question would be yes if I was in a position to be able to play into this market'	'The more they realise the value of information ... they may actually enter the game'	'you can for example identify ... very intimate traits of people by simply looking at traits that are not actually [that] sensitive'
Senior Developer in an Australian home loan brokerage company	opinion	neutral	positive	positive	positive	pricing
	illustrative quotes	'It's in addition to your primary, you setup a secondary market'	'it ... provides transparency into what you're giving away'	'I do agree that the Holy Grail for everyone is to have that single stream [of personal information]'	'you're ... monetising what you're giving away for free'	'it's the buyer who knows more about the value [of the data] ... and the seller really is at an absolute disadvantage as they have no idea as to what ... benefits their personal information can generate'
Managing Director of major political campaign advisory company	opinion	none offered	none offered	neutral	none offered	none
	illustrative quotes	N/A	N/A	'it would be a nice to have rather than essential'	N/A	
Academic 1	opinion	positive	positive	positive	none offered	raised alternative views on privacy

	illustrative quotes	'It sounds very convincing'	'As a person, I completely agree with you. I could not agree more'	'the market for personal information is large and booming'	N/A	'[some people are] highly critical of the idea of data vaults'
Academic 2	opinion	positive	positive	none offered	none offered	pricing, critical mass
	illustrative quotes	'It's good to see ...workable platforms'	'I find your model very interesting.'	N/A	N/A	'what prices do you expect?' 'how do you get the individuals to actually contribute the data to the marketplace?'

7 Individuals' appetite to participate on a PPIM conceived as a two-sided market

7.1 Research questions

The aim of this chapter is to study whether or not a PPIM would be a viable solution to the problem outlined above in section 1.1. It is written in the context of a broader design science based research project to design and develop a Primary Personal Information Market (PPIM) (Farrelly & Chew 2016a, 2016b). One of the key requirements of a viable market is thickness, defined as the ability of a market to 'attract a large enough proportion of the potential participants' (Roth 2008, p. 79).

Following Alvesson & Sandberg (2011) who propose gap spotting and problematisation as two distinct but not mutually exclusive methods for developing research questions, I challenge the (often unstated) underlying ideological assumption that personal information is not well suited to commodification, by proposing the research questions as follows:

RQ1: Would a sufficiently large proportion of digital citizens be willing to participate in a permissions-based PPIM to make such a market viable?

RQ2: What factors determine whether or not individuals would be willing to participate on a permissions-based PPIM?

The question of whether or not a permissions-based PPIM is feasible has wide ranging implications and significance for both individuals and end users. Should such a market prove to be feasible it would enable any internet-connected individual who was so inclined to benefit financially from their personal information. The magnitude of such a market should not be underestimated. The value of personal information in the European economy alone has been estimated at €330 billion annually for organisations and €670 billion annually for individuals (Rose, Rehse & Röber 2012). A permissions-based PPIM would also affect the way in which companies acquire and use personal information of both prospective and existing customers. A working PPIM would enable companies to conceive of and purchase novel personal data products which would benefit them in many ways, some of which are discussed in section 7.5.

7.2 Literature review

This dissertation takes social media to be 'a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content' (Kaplan & Haenlein 2010, p. 61). The literature on social media is vast and varied - see for example Cao et al. (2015), Ngai et al. (2015), Ngai, Tao & Moon (2015) and Roch & Mosconi

(2016). Since this research is concerned with the design and development of an m-sided platform market in personal information it focuses on three aspects of social media. I review the literature which concerns itself with the individual's motivations to share and also the monetary value individuals place upon their social media data. I then turn to the willingness individuals have to share their social media personal information.

7.2.1 Motivation to share

Numerous motivations for sharing personal data have been identified in the literature. Drawing on Reiss's model of sixteen basic desires (Reiss 2004) and Kraut et al. (2012) argue that individuals participate on Social Media Platforms (SMPs) because it is intrinsically rewarding to do so. They argue that external incentives can have the effect of undermining intrinsic motivation. Hau & Kim (2011) found that online gamers were motivated to share innovative knowledge by intrinsic motivations as well as external motivations such as the strengthening of social capital. Wang, Min & Han (2016) identified trust and appetite for risk as factors which strongly influence individuals' behaviour on SMPs. Kietzmann et al. (2011) identifies the ability to share as functional building blocks which motivate individuals to participate on SMPs. According to reciprocity theory, individuals share personal information on SMPs in the belief that such sharing will increase the cohesiveness of their social network and, by implication, their own welfare will be enhanced (Hsu & Lin 2008). Other factors in sharing include: enjoyment, altruism and reputation and community identification (Hsu & Lin 2008).

There are a number of factors which negatively impact on an individual's motivation to share personal data on SMPs. (Hu, Poston & Kettinger 2011) found that perceived risks associated with sharing personal data was a contributing factor to individuals' decisions to not to adopt SMPs. They also found that perceptions about ease-of-use and loss of control were significant in the decision not to share on SMPs.

7.2.2 Value of personal information

There have been a number of attempts to put a price on some aspects of personal information. Researchers have attempted to estimate the worth of individual users to various social media sites (Brustein 2012; Li et al. 2013). Taylor (2002) reports on value of targeted mailing lists and Carrascal et al. (2013) estimated the monetary value of personal information such as web browsing behaviour and demographic details. Cvrcek et al. (2006) found that an individual's location data was worth around €50-100 EUR a month while Danezis, Lewis & Anderson (2005) found location data to be worth around £10. Benndorf & Normann (2014) found that, in a one-off experiment, individuals were willing to sell their contact details and Facebook details for €15 and €19 respectively. AT&T offered fibre internet subscribers the opportunity to opt out of their deep packet inspection program for a price of

\$29 per month (Schneier 2015). Hann et al. (2007) found that the disallowance of secondary use of personal information is worth between \$39.83 and \$49.78.

Using a series of laboratory experiments, Benndorf & Normann (2014) found that 80% – 90% of participants were willing to sell personal data. They found that participants were willing to accept approximately €15 for their contact details and €19 for Facebook details. Grossklags & Acquisti (2007) compared willingness to sell with willingness to protect and found that participants had a clear preference for money in exchange for data in both the protection and release scenarios, even when the amount of money received was very small.

Zannier (2013) explored the feasibility of selling his personal information. He sold detailed data on many aspects of his personal life in exchange for contributions to his kick-starter project. He raised \$2,734 USD from 213 backers.

Regardless of the exact monetary value of personal information, Katell, Mishra & Scaff (2016) found that individuals generally considered the services they received on SMPs to be fair reward for the personal information they contributed, although this was somewhat mitigated when participants learned more about current data mining and profiling practices. However, there is also a growing body of research (Bagust 2014; Doyle 2015) which examines the social media phenomenon from the perspective of the commodification of affective labour, suggesting that the time and effort individuals expend creating personal information on SMPs is undervalued.

7.2.3 Willingness to share

A number of related approaches have been taken to assess whether or not people would be willing to exchange their personal information. Spiekermann & Novotny (2014) studied the research question of how individuals can share more personal information with companies while maintaining their personal privacy. They proposed a four-space model of personal information management by focusing on the type of relationship that exists between the individual and the company. Their focus was on increasing transparency and efficiency in the personal information market rather than allowing individuals to obtain financial compensation by selling access to their personal information.

Shibchurn & Yan (2014) found that individuals are more willing to share personal information on social networking sites (SNSs) when offered a monetary reward and that a number of other factors such as gender, marital status and educational attainment are significant when predicting whether or not individuals are willing to share personal information on SNSs. Aperjis & Huberman (2012) suggested that individuals should be compensated for allowing their data to be included in a sample from which the buyer sought to infer a statistic of interest about a population.

Yassine et al. (2011) studied the willingness of individuals to share personal information in return for a payoff value that balances out their loss of privacy. They explored the use of consumer agents which

act as brokers on behalf of individuals to manage access to their personal information. Gabisch & Milne (2014) found that receiving a monetary reward reduces consumer expectations for privacy and that one of the factors influencing an individual's decision to share personal data was whether or not the personal data was sensitive in nature. de Montjoye et al. (2014) demonstrated that individuals were willing to grant access to their personal information if their privacy was preserved. They found that 81% of participants would archive personal data in a personal data store and that participants were willing to share personal metadata with health care professionals. Rose, Rehse & Röber (2012, p. 25) found that while individuals expressed concern about use of private data by companies, this did not affect their willingness to share personal information. Chang & Chuang (2011) found that altruism, identification, reciprocity, and shared language were significant factors influencing knowledge sharing on SMPs. Chen & Hung (2010) identified reciprocity, interpersonal trust, knowledge sharing self-efficacy, and perceived relative advantage as important factors determining whether or not individuals would participate on professional virtual communities.

7.2.4 Summary and critique

The literature indicates the central issue pertaining to participation on SMPs is the question of whether individuals are motivated by intrinsic motivations, extrinsic motivations or a combination of both. Intrinsic motivations such as a desire to share, altruism and enjoyment of camaraderie are cited as motivation for participation while other authors argue that extrinsic motivations such as reputational enhancement, increasing one's status and the development of the cohesiveness of one's social group are the predominant motivations. The literature also confounds motivations with inhibitors and catalyst. These are features of SMPs which facilitate or impede individuals' participation and should be treated separately from motivational impulses. Concerns about trust, loss of control and risk aversion are examples of inhibitors while ease of use should be considered a catalyst rather than a motivator.

It is also clear that there is a dearth of research investigating the effect of monetary reward on individuals' willingness to participate on SMPs. While there has been a disparate and multifaceted approach to attempting to estimate the value of personal information of various types, the nexus between perceived value and willingness to exchange permission to access that information has yet to be determined.

7.3 Method

A survey was conducted to test participants' willingness to exchange personal information on a permissions-based PPIM similar to the one proposed by Farrelly & Chew (2016b). To do so, the design science methodology was followed and a simulation used (Gregor & Hevner 2013; Peffers et al. 2006) to generate 1534 requests to access personal data. Randomly selected requests were then presented to each respondent. Comparable simulation approaches have been used in other contexts to evaluate such

phenomena as the efficacy of information retrieval systems (Borlund 2016) and online consumers' browsing patterns (Islam & Miah 2012). Each request comprised the following components: a company making the request, the use to which the data would be put, a contact channel through which the company would contact the individual (if any), the data to be accessed, the device from which the data would be collected and a price to be paid for access to the personal information. These components are described in the following sections.

7.4 Company types

The companies included in the survey questions were based on the Australian and New Zealand Standard Industrial Classification (ABS 1993). We chose representatives of the following sectors: Finance and Insurance, Education, Electricity, Gas and Water Supply, Accommodation, Cafes and Restaurants, Personal and Other Services, Government Administration and Defence, Health and Community Services, Communication Services, Property and Business Services, Retail Trade, Construction, Manufacturing, Transport and Storage and Cultural and Recreational Services.

7.5 Uses

I drew on the literature to identify end uses to which personal data could be put. I included both existing end uses and potential uses, which have been suggested by researchers but not yet realised in practice. I then matched these end uses with companies in appropriate categories. These end uses were: personalised search (Teevan, Dumais & Horvitz 2005), predictive marketing (Gualtieri 2013; Rose, Rehse & Röber 2012, p. 25), predictive customer service (Gualtieri 2013; World Economic Forum and Bain & Company 2011), customer self-service (Rose, Rehse & Röber 2012, pp. 11, 25), demographic studies (de Montjoye et al. 2014, p. 1), process automation (Rose, Rehse & Röber 2012, p. 11), personalised medicine (Rose, Rehse & Röber 2012, p. 81), personalised marketing (Rose, Rehse & Röber 2012, p. 25), personalised products (Rose, Rehse & Röber 2012, p. 10), provide critical public services more efficiently and effectively (World Economic Forum and Bain & Company 2011, p. 5), personalised education (Holdren & Lander 2014; Kyriacou & Davis 2008; Van Kleek & O'Hara 2014, p. 30), service innovation (Laney, Buytendijk & Linden 2013; Lee, Kao & Yang 2014), personalised pricing (Holdren & Lander 2014, p. 11; Swan 2013; Taylor 2002), personalised health (Holdren & Lander 2014; Shah 2015; Shilton et al. 2009, p. 10; Swan 2013), personalised services (Holdren & Lander 2014, p. 41), personalised recommendations (de Montjoye et al. 2014, p. 1) and research and development (Rose, Rehse & Röber 2012, p. 9; World Economic Forum and Bain & Company 2011, p. 5).

7.6 Price

Estimating a reasonable price for access to personal information is an inexact science since we are proposing a novel permissions-based PPIM which has not been implemented previously. However, from the literature (Benndorf & Normann 2014; Carrascal et al. 2013; Cvrcek et al. 2006; Danezis, Lewis & Anderson 2005; Hann et al. 2007; Li et al. 2013; Muschalle et al. 2013; Schneier 2015; Taylor 2002; Yassine et al. 2011) estimates of the potential worth of a number of different personal information streams were derived. By converting these estimates into 2015 Australian dollar equivalents I took as a working hypodissertation that access to a single stream of personal information could be worth as little as a few cents and as much as \$3 per month. A conservative estimate of the worth of personal information streams was taken in order to get a realistic estimate of the number of potential participants.

7.7 Requests

Having created database tables of organisations, uses, contact channels and devices, three intermediate tables were created: `uses_data`, `data_devices` and `org_uses`. `uses_data` matched the uses to which personal data could be put with the appropriate personal data sources. `data_devices` matched personal data sources to the devices from which they could be collected. `org_uses` matched organisation types with relevant uses of personal data. These tables were then used to generate 1534 distinct requests.

7.8 The IT infrastructure

The IT infrastructure that supports this survey was as follows. A Postgres database was constructed comprising five initial tables. These tables contained the names of the contact channels, the data feeds, the uses, devices and organisations to be used to construct the requests. These tables were joined into three secondary tables which contained combinations of uses and data, data and devices and organisations and uses respectively. These combinations were based on pairs of the relevant data field which made sense. For example, ('Warshals Insurance', 'offer you personalised premiums') is an organisation/use combination which does make sense so it was included in the `org_uses` table while ('Burger Chef', 'offer you personalised premiums') does not make sense and is therefore not included in the table.

Rows from the `uses/data`, `data/devices` and `organisations/uses` tables were then combined with the contact channels to create the requests. The data model used to facilitate these combinations is shown in Figure 9.

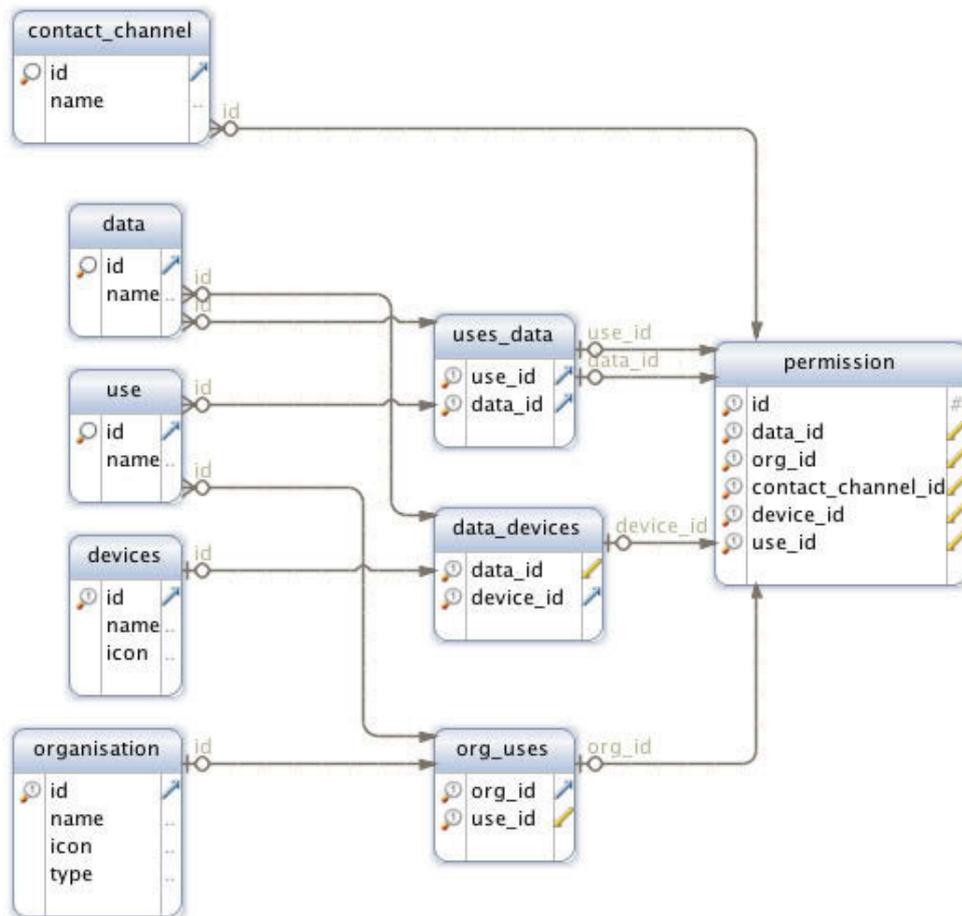


Figure 9: Permissions logical data model

7.9 The survey

An online survey was used to collect data for this chapter because it has been shown to be highly effective with regards to both cost of data collection and the time taken to collect the data (Wright 2005). Respondents were recruited by the chain-referral sampling technique via social media (Roth 2008, p. 79) which has been shown to be a cost effective and accurate way to sample a population (Baltar & Brunet 2012; Brickman-Bhutta 2012). The survey was conducted by presenting five webpages to the respondents. The first page was an overview of the research (Figure 10).

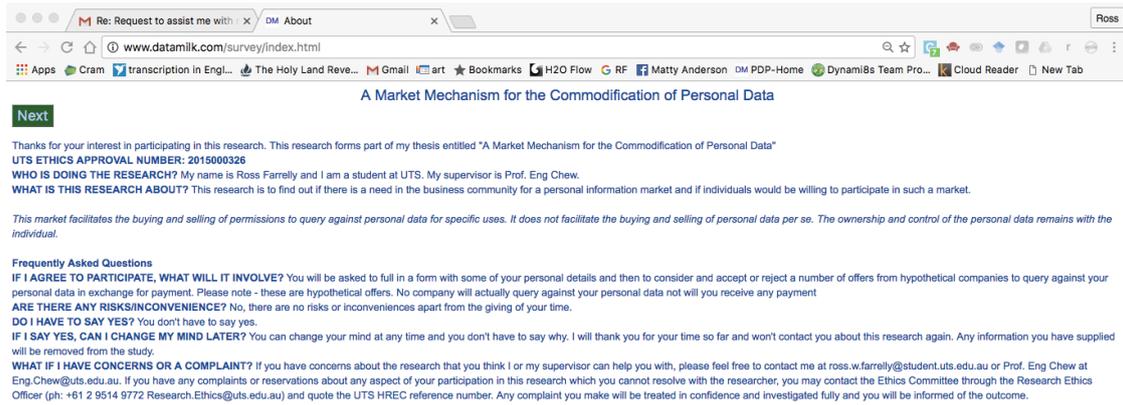


Figure 10: About screen from survey

The second page presented a description of the permissions-based PPIM and how it would work in practice (Figure 11).

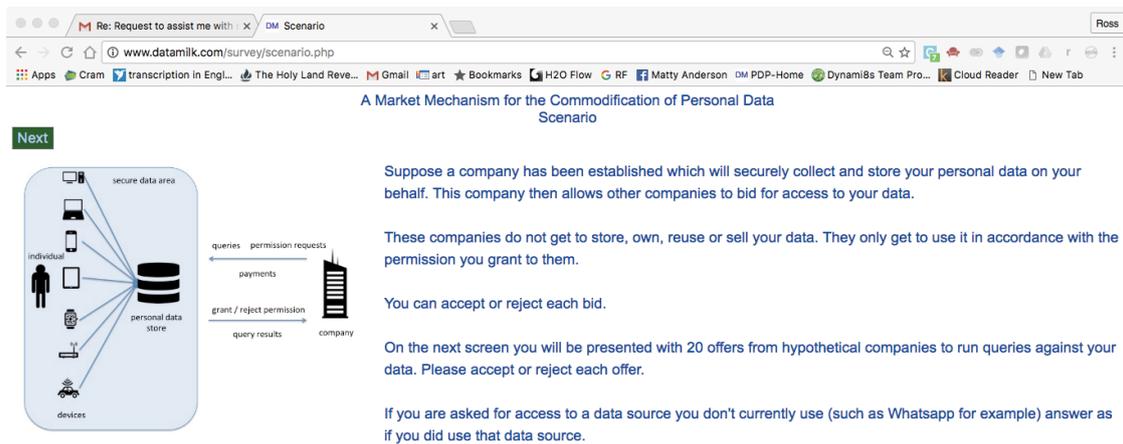


Figure 11: Scenario screen from survey

The third page asks the participant to enter personal details such as age, gender etc. Name and email address were optional for those who wished to be contacted with the results of the research.

Step 1 - Sign In
Please complete the following form. In the next screen you will be presented with permission to accept or reject.

About You	
age	28
gender	<input checked="" type="radio"/> Male <input type="radio"/> Female
highest educational attainment	Postgraduate Degree
ethnic background	Non-Hispanic White or Euro
State	NSW
postcode	
If you would like to be sent the results of this research or would be willing to participate in further research, please enter your details below.	
first name	
last name	
email	
<input type="button" value="Continue"/>	

Figure 12: Sign in screen from survey

The fourth page presented twenty randomly selected requests from the requests table and presented each one consecutively to the respondent. The respondent either accepted or rejected each offer by clicking on the appropriate button. Figure 13 and Figure 14 show examples of such requests.

Permission number: 1

Monk Home Improvements is a house renovation company
 Monk Home Improvements requests permission to run queries against your emails collected from your PC so they can develop more relevant personalized products and services.
 Monk Home Improvements would like to send you relevant messages via linkedin ads.
 Monk Home Improvements will pay you \$2.85 per month if you accept this request.

Permission summary	
Organisation Monk Home Improvements	Data emails
Device PC	Use develop more relevant personalized products and services

Compensation offered: \$2.85 per month

Figure 13: Screen shot 1 of example permissions as presented to respondents



Figure 14: Screen shot 2 of example permissions as presented to respondents

The fifth and final page thanked participants for their participation and invited them to share a link to the research via email and social media Figure 15.

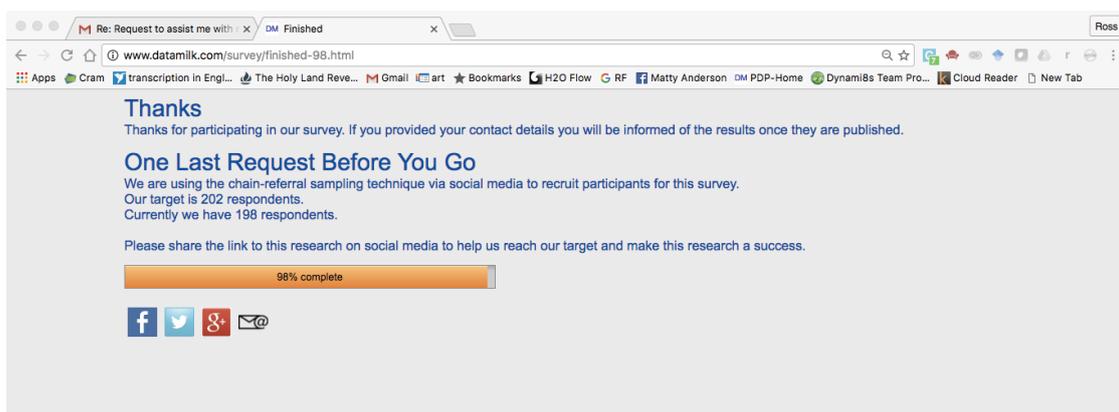


Figure 15: Finish and request to share page from survey

Since the wish is to obtain an initial indication of the proportion of digital citizens who would be willing to participate in permissions-based PPIM, sample size calculations based on random sampling guidelines to calculate the sample size are used. With a margin of error of 0.05, a confidence interval

of 0.9 and a conservative estimate of the proportion of 0.8 based on prior work by Benndorf & Normann (2014), the required sample size is:

$$0.25(1 - 0.25) \left(\frac{1.645}{0.05} \right)^2 = 203 \text{ (Hinkle, Oliver \& Hinkle 1985)}$$

7.10 Results

204 individuals responded to the survey, 62 females and 141 males with ages ranging from 10 to 72. A total of 2060 requests were presented to respondents of which 767 were accepted (37.2%). Of the 204 individuals who participated, 163 accepted one or more request (79.9%).

RQ1:

Would a sufficiently large proportion of digital citizens be willing to participate in a permissions-based PPIM to make such a market viable?

Using the 1 sample proportion test in R (R Core Team 2015) we estimate the proportion of digital citizens who would participate in a permissions-based PPIM to be 79.9% and we can be 90% confident that the proportion of digital citizens who would participate in a permissions-based PPIM lies between 74.6% and 84.3%.

This is the 90% confidence interval for the proportion of digital citizens who accept one or more request was $\chi^2 = 56.0111$, $df = 1$, $p\text{-value} = 7.206e-14$.

RQ2:

What factors determine whether or not individuals would be willing to participate on a permissions-based PPIM?

To address this RQ I first examine individual factors and then examine their combined effect in a multivariate logistic regression model to identify which explanatory variables are significant in explaining the response variables (accept or reject). Multivariate logistic regression is a commonly used technique to estimate the effect of multiple explanatory variables on a single binary response variable. The covariates we have are relate to the both the respondent and the permission being sought namely: age, gender, highest educational attainment, ethnic background, type of data (to which permission will be granted), device (from which the data will be collected), contact channel (by which the individuals would be contacted), price offered to access the personal information, type of use (to which the information will be put) and type of organisation (which is seeking access to the information).

The multivariate logistic regression model was built using forward selection using R package glmulti (Calcagno 2013). I test goodness of fit using Pearson chisquare statistic ($p > 0.05$), and significance of the included variables ($p < 0.05$)

Multivariate logistic regression can be expressed in the following model:

$$\text{logit}(p) = \beta_0x_1 + \beta_1x_2 + \dots + \beta_kx_k$$

where p is the probability the request will be accepted, β_i is the i th coefficient, x_i is the i th variable and logit is defined as the natural log of the odds of the outcome or $\ln\left(\frac{p}{1-p}\right)$.

The variables together with the statistically significant levels ($p < 0.05$) are shown in Table 15.

Table 15: Logistic regression model

variable / level	Estimate	exp(estimate)	Std. Error	z value	Pr(> z)
Intercept	0.850087		0.327931	2.592	0.00953
age	-0.013061		0.003997	-3.267	0.00109
price	0.290256		0.055372	5.242	1.59e-07
contact_channel_email	-0.465816	0.627622756	0.203564	-2.288	0.02212
contact_channel_Four Square	-0.446096	0.640122317	0.206553	-2.160	0.03079
contact_channel_Google plus ads	-0.439082	0.644627918	0.205076	-2.141	0.03227
contact_channel_phone	-0.439653	0.64425994	0.205017	-2.144	0.03200
contact_channel_sms	-1.026743	0.358171628	0.219081	-4.687	2.78e-06
use_type_personalized services	-2.013606	0.133506382	0.837175	-2.405	0.01616
use_type_provide critical public services more efficiently and effectively	0.474013	1.60642787	-2.490	0.01277	0.01277
ethnic_background_Asian	0.445742	1.561648509	0.165421	2.695	0.00705
ethnic_background_Black, Afro-Caribbean, or African American	1.297544	3.660295932	0.503999	2.574	0.01004
org_type_Finance and Insurance	-1.308181	0.270311306	0.649671	-2.014	0.04405

The coefficients of the model are interpreted as follows:

Age: Holding all other variables fixed, for every one year increase in age of the recipient, the odds of permission request being accepted decreases by a factor of $e^{-0.013} = 0.987$, a decrease of 1.3%.

Price: Holding all other variables fixed, for every one dollar increase in price offered, the odds of permission request being accepted increase by a factor of $e^{0.290} = 1.337$, an increase of 33.7%.

contact_channel_email: Holding all other variables fixed, compared to the reference level of will_not_contact, including contact_channel_email in a permission increases the odds of the permission request being accepted by a factor of $e^{0.290} = 0.628$, an increase of 33.64%.

The interpretation of the other categorical variables follows a similar pattern and is summarised in Table 16. The full code used to create analyse these data can be found in Appendix 3.

Table 16: Statistically significant variables and associated changes in odds

variable / level	Estimate	reference category	exp(estimate)	% change in odds
age	-0.013	n/a	0.987	-1.30%
price	0.290	n/a	1.337	33.68%
contact_channel_email	-0.466	will_not_contact	0.628	-37.24%
contact_channel_Four Square	-0.446	will_not_contact	0.640	-35.99%
contact_channel_Google plus ads	-0.439	will_not_contact	0.645	-35.54%
contact_channel_phone	-0.440	will_not_contact	0.644	-35.57%
contact_channel_sms	-1.027	will_not_contact	0.358	-64.18%
use_type_personalized services	-2.014	personalized_medicine	0.134	-86.65%
use_type_provide critical public services more efficiently and effectively	-2.014	personalized_medicine	1.606	60.64%
ethnic_background_Asian	0.474	non_hispanic_white_euro	1.562	56.16%
ethnic_background_Black, Afro-Caribbean, or African American	0.446	non_hispanic_white_euro	3.660	266.03%
org_type_Finance and Insurance	1.298	health_and_community_services	0.270	-72.97%

7.11 Conclusions

As discussed in section 7.2, extant literature shows that individuals are willing to share their social media and other personal information for reasons such as altruism, intrinsic reward and strengthening of social capital. This research extends these finding by showing that a large proportion of digital citizens would be willing to grant permission to use their social media data while retaining ownership of those data. Specifically, this research shows that younger citizens are more willing to participate in such a market and that offering higher prices for access to personal information increase the odds of the request being accepted. It was also found that the contact channel had a significant impact on a citizen's willingness to be contacted. Being contacted by email was more popular than not being contacted at all, while being contacted by phone, SMS and via ads was less popular. Collecting personal information in order to personalise services and to inform the provision of public services were less favourably received than other uses to which personal data could be put.

7.12 Discussion

These findings are significant because they lend weight to the arguments in favour of a permissions-based PPIM and they will guide the developers of such a market as to which factors are significant in individuals' decisions about participating in a permissions-based PPIM. The findings in this chapter represent important information for researchers concerned with the viability and design of a PPIM. First, the overall estimated participation rate (86.5%) indicates that the concept of a PPIM is worth pursuing. There seems to be significant interest among the general population in exchanging personal information in exchange for financial compensation. Secondly, the significant factors in the logistic regression model inform researchers who are looking to develop a working PPIM about requests which are most likely to be accepted by individuals and therefore mostly likely to be traded successful on a PPIM. For example, it tells us that younger individuals are more likely than older individuals to accept request to access their personal data. It shows that requests that include a component which allows the

end user to contact individuals by email are more likely to be accepted than those that seek permission to contact individuals by SMS. Each coefficient in Table 16 interpreted in a similar manner informs researchers about how a PPIM might be expected to operate.

The research journey thus far has involved the design for both a two-sided market in personal information in chapter 5 and a design for the transaction object in chapter 6. In this chapter we established that individuals would be fairly likely to participate on a permissions-based PPIM. However, in the course of this research a number of additional design questions have arisen. How could we rigorously design an innovative service to commercialise personal information? How can individuals and companies co-create value from personal information? Are additional groups of market participants required to ensure innovation, scalability and evolution of a PPIM? These questions are addressed in the next chapter through the novel combination of *iSIM* and industry platform theory.

8 PPIM conceived as a four-sided industry platform

8.1 Introduction

This chapter is based largely on Farrelly & Chew (2017). It represents a second iteration in the design science process (Hevner et al. 2004; Peffers et al. 2006; Peffers et al. 2007-8). This chapter explores the potential for commercialising a PPIM through the provisioning of a variety of possible personal data products (PDPs). We employ the integrated Service Innovation Method (*iSIM*) as a framework by which to design PPIM as an industry platform for innovative personal data service delivery (Krueger et al. 2015).

iSIM is ‘an integrated cross-disciplinary, holistic method to design and commercialize service innovation’ (Chew 2015:1). It is a method to simultaneously design an innovative service offering and an associated business model with the aim of engendering sustained value co-creation between customer and firm.

The research question we seek to address is: How can a PPIM be conceptualised and prototyped as an industry platform using *iSIM*?

This chapter is relevant because it simultaneously deepens and improves the design of a PPIM and also demonstrates a novel application of the *iSIM*. It will be of relevance to researchers in a plurality of disciplines including design science, service innovation and platform development because it applies a novel design approach in an industry platform setting and develops a novel conception of a large-scale market for personal information.

This chapter makes a contribution to the literature by improving and refining the design of a PPIM in the light of the evaluation from previous research. It further reconceptualises a PPIM as an industry

platform and explains how value can be co-created between PPIM stakeholders with a view towards PPIM commercialisation. It also demonstrates the efficacy of *iSIM* as an innovation method for industry platform design.

This chapter is organised as follows. Section 8.2 outlines the background to the research including both platform theory and markets in personal information. Section 8.3 describes the method utilised for the research while section 8.4 presents the proposed design solution, organised in sub-sections according to the *iSIM*. Section 8.5 describes the prototype which instantiates the design. Section 8.6 presents the results of the evaluation of the prototype. Section 8.7 concludes the chapter.

8.2 Background

The references in this section were sourced using a process similar to Kitchenham's systematic literature review methodology (Kitchenham et al. 2009) by searching Business Source Complete, ACM Digital Library, IEEE Explore, ProQuest and Scopus for the phrases 'platform theory', 'industry platform', 'product platform', 'personal information markets' and 'online marketplace design'.

The concept of the platform originated in manufacturing where product platforms were developed as a core base to which features could be added to meet customer needs (Wheelwright & Clark 1992). These platforms evolved into systems along with their associated interfaces, which were used as a standardised substratum upon which complex and customised products could be efficiently produced (Lehnerd 1987; Muffatto & Roveda 2002). The concept of the industry platform (often associated with the concept of the 'keystone firm' (Iansiti & Levien 2004) further evolved to encompass software-based products and services (Tiwana, Konsynski & Bush 2010). While Thomas, Autio & Gann (2014) identified four main types of platform, a binary distinction can be drawn between internal (company or product) platforms and external (or industry) platforms.

An industry platform is defined to be 'products, services, or technologies that ... provide the foundation upon which outside firms ... can develop their own complementary products, technologies, or services' (Gawer & Cusumano 2014:418). An m -sided ($m \geq 2$) platform facilitates a market by connecting m groups of users which provide each other with network effects (Muzellec, Ronteau & Lambkin 2015).

There are three key aspects to an industry platform: a relatively stable core component, versatile, complementary peripheral components and a stable interface component which links the two (Baldwin & Woodard 2008).

Industry platforms are open (Gawer & Cusumano 2014), in the sense that third party companies can interact with an industry platform in order to develop and create their own products. Industry platforms are also scalable (Toppenberg, Henningson & Eaton 2016), able to accommodate a large volume of both actors who wish to interact with them. Two other related characteristics of industry platforms are their ability to adapt and evolve (Gawer & Cusumano 2014), i.e. they are able to adjust and

accommodate unanticipated changes in the external business environment. Industry platforms are also modular and non-deterministic or emergently generative (Gawer & Cusumano 2014; Thomas, Autio & Gann 2014; Tiwana, Konsynski & Bush 2010) in the sense that they do not determine the end use of products which are built upon them.

Recent developments in platform theory focus on linking two previously separate conceptions of the nature of a platform: the market perspective and the engineering perspective. From the market perspective Gawer (2014, p. 421) argues that industry platforms must ‘perform a function that is essential to a broader technological system’ and ‘solve a business problem for many firms and users in the industry’. Gawer & Cusumano (2014) also emphasises that industry platforms must act as multi-sided markets, facilitating the co-creation of value from multiple contributors. Such multi-sided markets provide ‘goods or services to several distinct groups of customers, all of whom need each other in some way and rely on the platform to mediate their transactions’ (Gawer & Cusumano 2014, p. 422).

Research into markets in personal information is multi-dimensional and cross-disciplinary and touches on numerous areas of related research such as: data ownership (Hoeren 2014; Mashhadi, Kawsar & Acer 2014), privacy (Nissim, Vadhan & Xiao 2014; Schunter 2014), personal informatics (Li, Dey & Forlizzi 2010; Li et al. 2012), the collection and modelling of personal information (Kyriacou & Davis 2008; Shimojo et al. 2010) and data security (Sugumaran, Murugan & Kamalraj 2014; Thapliyal, Mandoria & Garg 2014). However, as noted by e-market scholars, ‘there is a substantial lack of empirical research on ... markets for online personal data’ (Milyaeva & Neyland Forthcoming, p. 4) and most market design research is qualitative (Wang et al. 2008).

Spiekermann & Novotny (2015) explored the tension between sharing personal information with companies while maintaining personal privacy. By focusing on the type of relationship that exists between the individual and the company, they developed a four-space model of personal information management. It has also been proposed that individuals receive compensation for allowing their data to be included in a sample from which the buyer sought to infer a statistic of interest about a population (Aperjis & Huberman 2012) and for allowing access to sensor data generated by the Internet of Things (Mashhadi, Kawsar & Acer 2014).

Other researchers (de Montjoye et al. 2014; de Montjoye et al. 2012, p. 2) have alluded to the need for a marketplace for personal information without explicitly calling for a PPIM. They have identified the need for individuals to retain ownership of their personal information, to have more control over their privacy (Sevignani 2013) and have identified a growing number of start-ups which focus on the value of data usage rather than the value of data itself (Milyaeva & Neyland Forthcoming).

Farrelly & Chew (2016b) propose a permissions-based PPIM on which individuals receive compensation for granting access to their personal data. However, these markets only involve two

actors, individuals supplying the personal information (the raw materials) and companies purchasing the PDPs.

In summary we see that, while the issue of personal information and its inherent value has been approached by numerous scholars from many perspectives, there is little (if any) research which examines a personal information market, both from an economic and engineering perspective, as an industry platform. This represents a significant gap in the literature as a scalable multi-sided PPIM which incorporates input from multiple contributors, has many of the characteristics of an industry platform and would therefore benefit from being conceptualised as such.

8.3 Method

Following the *iSIM* (Chew 2015), the initial design of a PPIM is explored in terms of the service business strategy, customer type and customer value proposition, the service concept, the service system, customer experience, service architecture and monetisation. In practice, the application of *iSIM* will be non-linear and agile – in the sense that entrepreneurs implementing a PPIM may follow different (iterative) paths through the design process. However, the design process is described sequentially for the sake of logical design information flow. It is outlined how a PPIM would incorporate these perspectives in a multi-sided market (see section 8.4) and then instantiate the design in a prototype (described in section 8.5).

Using semi-structured interviews, in accordance with design science, we evaluate the PPIM industry platform qualitatively (Yin 2013). These interviews were conducted with reflective practitioners (Gummesson 2002) some of whom worked in data related roles. They were selected on the basis of their subject matter expertise, both data related and business related. This is consistent with other studies; see for example (Karunakaran & Puroo 2012; Knol, Sol & Van Wamelen 2012; Nenonen & Storbacka 2010).

Representatives were interviewed (who are considered to be potential target PPIM product/service consumers) from eight businesses, one from each of Education, Cultural and Recreational Services; Accommodation, Cafes and Restaurants; Government Administration and Defence; and Communication Services and two from each of Finance and Insurance and Personal and Other Services (ABS 1993). There was one micro, two small, three medium and two large businesses.

The aim of the evaluation was to gather opinions from interviewees on the following key implementation factors: the feasibility of building a PPIM as an industry platform, the viability of such a PPIM if it were to be built, and the appetite that companies might exhibit to participate in such a PPIM, additional PDPs which could also be created on the PPIM so described and any issues which may prevent or impede the workability of the PPIM.

The selection of interviewees was guided by theoretical sampling (Eisenhardt 1989) in which we sampled experts who could evaluate the constructs and structure of the proposed designs based on their domain expertise and years of experience. Interviewees had an average of 22.25 years' experience (as indicated by their LinkedIn profiles).

An interview topic guide (King 1994) was used for interview consistency. After a brief explanation of the purpose of the research and the proposed PPIM industry platform the interviewees were presented with a demonstration of the prototype. They were asked their opinions on the feasibility of the PPIM industry platform and if they identified any major flaws or issues with the proposed solution. They were asked whether or not such a market would be viable, and based on their experience whether there was appetite for companies to purchase PDPs on a PPIM and about the willingness of individuals to share personal information on such a market. They were asked to conceive of any novel personal data product ideas of value to their businesses which could be provisioned on the proposed PPIM. They were asked which departments (if any) within their businesses would use the proposed PPIM on a daily basis. The average duration of the interviews was 44.6 minutes.

8.4 Design solution

8.4.1 Service business strategy design

The business strategy envisioned for the proposed PPIM is an instantaneous 'build-to-order' logic (McGrath 2010) associated with a virtual marketplace. The raw materials (personal data) are ingested via data feed connections. The output from models that have extracted inferred information from the raw data are also present in the centralised data store. The end user co-creates his or her customised PDP with PPIM by selecting an appropriate PPIM-provisioned template and populating it with the requisite components to customise the PDP. After previewing the PDP, the end user creates the PDP which is instantaneously built-to-order at the completion of the customisation process. When the PDP is executed, the platform facilitates the revenue distribution from the purchaser to the contributors to the PDP.

While the prototype presented in this chapter is a four-sided market the PPIM industry platform's strategic intent is towards becoming an m-sided market ($m > 4$). Additional stakeholders who could participate on a PPIM include (but are not limited to) PDP template developers, plug in developers to interface the PDS API with company enterprise resource planning (ERP) or customer relations management (CRM) systems, experimental developers to facilitate experiments and testing on PDP usage, personal data brokers to manage access to personal data on behalf of individuals, pricing model developers to test alternative pricing models such as auctions, reverse auctions and pricing for exclusive use.

8.4.2 Customer type and customer value proposition design

Designing the customer type and customer value proposition (CVP) is an iterative process (Chew 2015). Following the first iteration of the design process (Farrelly & Chew 2016a) it is proposed that four distinct types of participants could participate on a PPIM to co-create value from personal data: data contributors, model developers, data feed developers and personal data product consumers (see Figure 16 below).

Data contributors have the role of granting permission for their data to be collected on their behalf, stored in the centralised data repository and for granting permission for their data to be included in the PDPs to which they give their consent. The value proposition to data contributors is the full control of their personal data and the financial compensation they receive in exchange for limited and transparent use of their personal data on which they have granted permission.

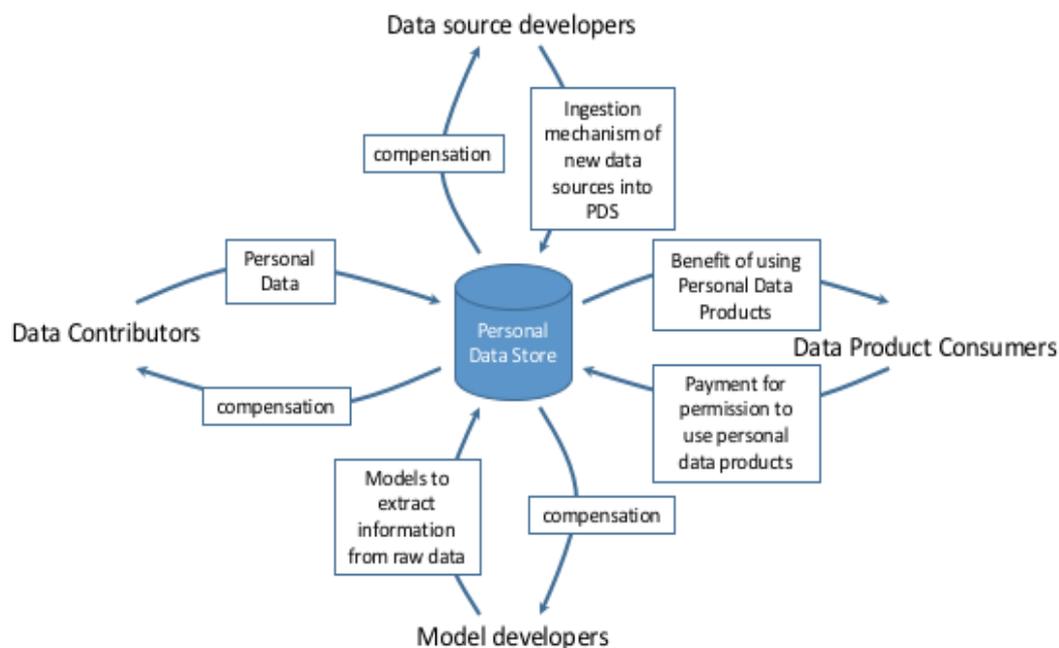


Figure 16: Four participant types and value propositions on a PPIM

Model developers participate on a PPIM by co-producing models with the data contributing individuals. These models extract insights of value from the raw personal data. A model may be as simple as extracting an individual's age from his or her Facebook data or as advanced as predicting an individual's preferred holiday destination using clustering techniques. PDPs are customised by selecting the output from one or more models together with an associated permitted usage and then instantaneously built-to-order at the completion of the customisation process. The value proposition to model developers is the access right to the permitted personal data repository and the financial compensation they receive

whenever a personal data product which comprises the results of one or more of their models is executed.

Data feed developers participate on a PPIM by developing feeds from new devices or services that generate or capture personal data into the centralised personalised data repository. The value proposition to data feed developers is the financial compensation they receive whenever a personal data product which comprises the raw data collected via one or more of their feeds is executed.

End users benefit from participating on a PPIM by specifying the PDPs they require and purchasing access to them.

8.4.3 Service concept design

Here the design of the service logic of a PPIM is outlined to fulfil the business logic and customer value propositions outlined above – from the perspective of four actor types: data contributing individuals, data feed and model developers and end users.

Data contributing individuals receive access to a centralised personal data store which captures (via a data feed / device interface) and securely stores their data. They receive requests for access to their personal data which, if accepted, result in financial compensation. Individuals receive access to a personal information dashboard which enables them to manage and control access to their personal data.

Data feed and model developers benefit from participating on a PPIM by gaining access to a market to promote and sell the fruit of their labours. The data feeds and models they develop are made available on the PPIM and are available to be incorporated into PDPs at build time on demand. When a PDP comprising a developer's data feed or model is executed the developer receives financial compensation. Developers also receive the ability to transparently track the use to which their components are being put and the financial compensation they are receiving.

End users benefit from participating on a PPIM by gaining the ability to co-design, co-build, then purchase and execute PDPs with the PPIM. It is well known that personal data is of great value to a wide range of businesses (World Economic Forum and Bain & Company 2011). End users receive the ability to browse personal data assets such as templates and information which has been inferred and extracted from the raw materials of personal data by model developers. Via the end user facing PDP dashboard, they receive a means to transparently co-create, preview, purchase, execute, delete, renew and monitor the PDPs. In line with best practice in a sharing economy (Thierer et al. 2015), end users also receive the ability to validate the quality of the data to which they are buying access and the validity of the models they are incorporating into their products via fair and transparent rating systems whereby the market can give feedback on the validity and utility of the models.

On a PPIM, PDPs can be co-created from four types of components: templates, data feeds, extracted information and permissions. The end user can utilise a template for a given type of PDP (targeted marketing, personal response, customer identification, etc.). The template is then populated with the data feeds (made possible by the data feed developers) and the information extracted from the data feeds (by the model developers). Finally, permission to execute a PDP, in terms of time duration, frequency of use and mode of contact is incorporated into the PDP.

This service concept design fulfils individuals' expectations of transparency (Spiekermann & Novotny 2015) (via the individuals' dashboard), opt in and opt out ability (by the ability to withhold and withdraw permission), fair compensation (via the payments system), low transaction costs and security of data. It also fulfils model developers' and data feed developers' expectations that end users can search for, locate and compare their offerings, and that they receive fair compensation for the use of them.

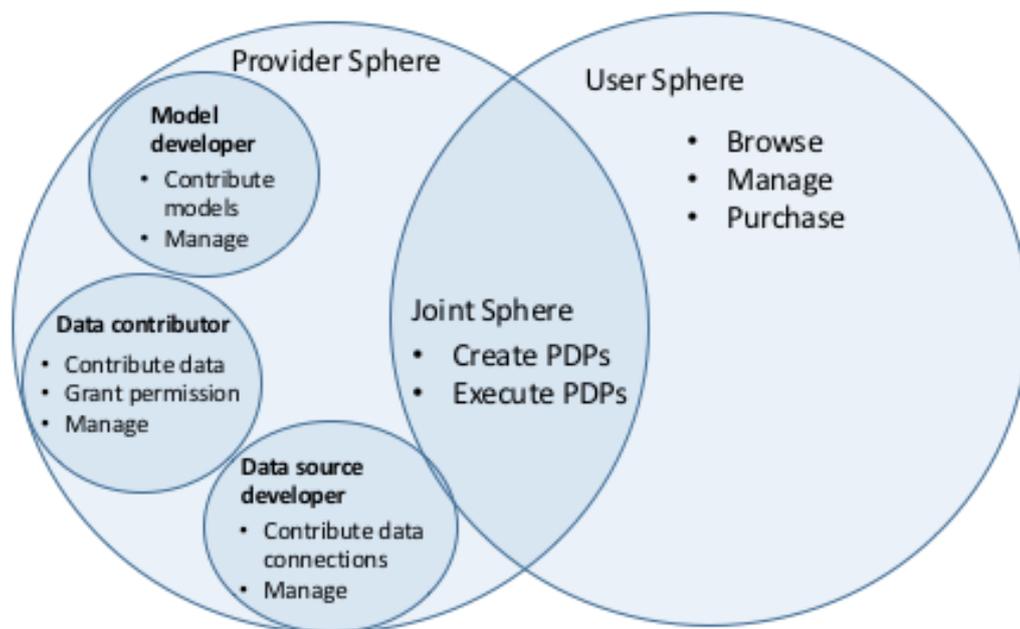


Figure 17: Provider, user and joint spheres of value creation on a PPIM

Figure 17 shows how, at the point in time the PDP is executed, its value-in-use (Vargo & Lusch 2004) (value co-creation) is experienced by the end user in the joint sphere at the intersection of the user and providers spheres (Grönroos & Voima 2013). The PDP has been co-created by the data contributing individuals granting permission and the end user assembling, purchasing and executing the PDP (using a template co-produced by the model developer).

8.4.4 Service architecture design

The purpose of service architecture is to systematise and capture the aspects of the service enabled by a PPIM as an industry platform. It does so by identifying the components, interfaces, degree of coupling, and commonality of sharing between components in the PPIM industry platform.

In the PPIM service architecture, there is a close relationship between the four stakeholders: data feed developers, model developers, data contributing individuals and end users in various PPIM value co-creation processes. Data feed developers co-create value by facilitating the ingestion of new data feeds (in turn from the individuals' raw personal data). Model developers co-create value by inferring, predicting and extracting additional knowledge from the raw data. Data contributing individuals co-create value by granting permission for their personal data to be used, and end user co-creates value by constructing the PDP, paying for the right to use it and co-creating value-in-use by executing the PDP and integrating the inherent valuable PDP knowledge into their own value-creating business activities.

The PDP templates and the data feed and model outputs are designed as modular components which enable reuse and adaptability to co-create customised PDPs. The four interfaces (APIs) share many common standardised elements as they interface to a common PDS. The dashboards share common elements such as PDPs, usage, payments and status of PDPs. The underlying PDS can be adapted to support PDPs for various uses including purely personal use, creation of PDPs and comparison between individuals.

The commonality between the API interfaces will be designed for extensibility to incorporate additional stakeholders, thereby facilitating the expansion of the four-sided market instantiated in the platform prototype to an m-sided market ($m > 4$).

8.4.5 Service system design

Here the service system design is outlined which defines how the service value co-creation will be performed. The PPIM service system design is shown in Figure 18. A centralised cloud based Personal Data Store (PDS) comprising a distributed file system and a massively parallel processing database stores the individuals' personal data. Apps which can be installed on physical devices capture at source individuals' personal data. A data model in the PDS tracks which data feeds, models and individuals' raw data have been used by the end users to co-create various PDPs.

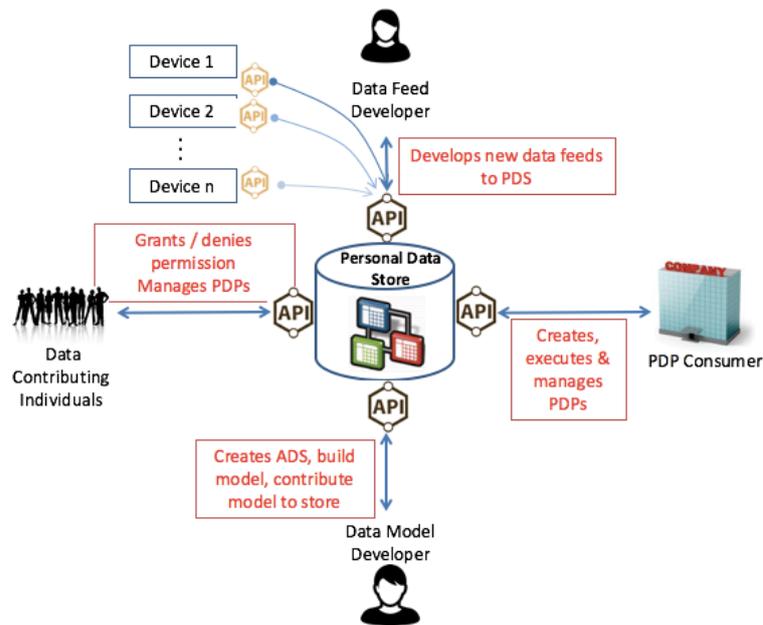


Figure 18: The PPIM service system design

The data captured at source on individuals' devices are transferred to the PDS. End users co-create PDPs using raw data captured via data feeds and the outputs from models that extract and infer additional information from the raw data. The data model is utilised to track which data feeds, models and individuals contributed to the PDP. When the PDP is executed, the contributors' accounts are credited appropriately (automatically).

The system requires database and file system administrators to maintain the PDS platform. It requires database file system data modellers (as opposed to personal data modellers) to maintain and extend the data model to incorporate new data streams. There is also a need for back-end admin staff to resolve disputes over billing and associated issues.

8.4.6 Customer experience design

Figure 19 shows the service experience blueprint for participation on a PPIM and its attendant service process flow for each customer type on the PPIM industry platform. PPIM platform customer experience is designed and described below in accordance with the cognitive, emotive and behavioural lenses of customer engagement (Frow et al. 2015, p. 473).

Data contributing individuals interact with the PPIM by installing the custom-built apps to collect their personal data on their behalf and to ingest and store them in the PDS.

Data contributing individuals' cognitive requirements are supported via a customer facing dashboard showing in real time which companies have purchased access to their personal data, the uses to which their personal data are being put, the PDPs in which their personal data are included and the revenue they have accrued through the use of their personal data. Data contributing individuals also receive

requests to access their personal data via their chosen channel (email, SMS, IM, etc.). The data contributing individuals' emotive experience of this service is designed to be one of trust engendered by the transparency and immediacy of the service which grants them visibility of and control over the uses to which their personal data are being put (Spiekermann & Novotny 2015). Their emotive needs of fairness and equity are addressed as they are compensated for the use of the valuable personal data they have generated.

Data feed developers and model developers are supported with modeller facing dashboards, which allow them to monitor the utilisation of their models and data feed in PDPs and to monitor the revenue generated by the use of those PDPs. Their behavioural interaction with the system entails them searching and browsing existing models and data feeds to avoid replication and the contribution of new models and data feeds to the PDS. The modellers' cognitive requirements for visibility and control of the uses to which their models and data feeds are being put are catered for via the modeller dashboards. Their emotive requirements for fairness and equity are fulfilled as they are compensated for the use of models and data feeds they have programmed and a means to expose their products to a wider audience.

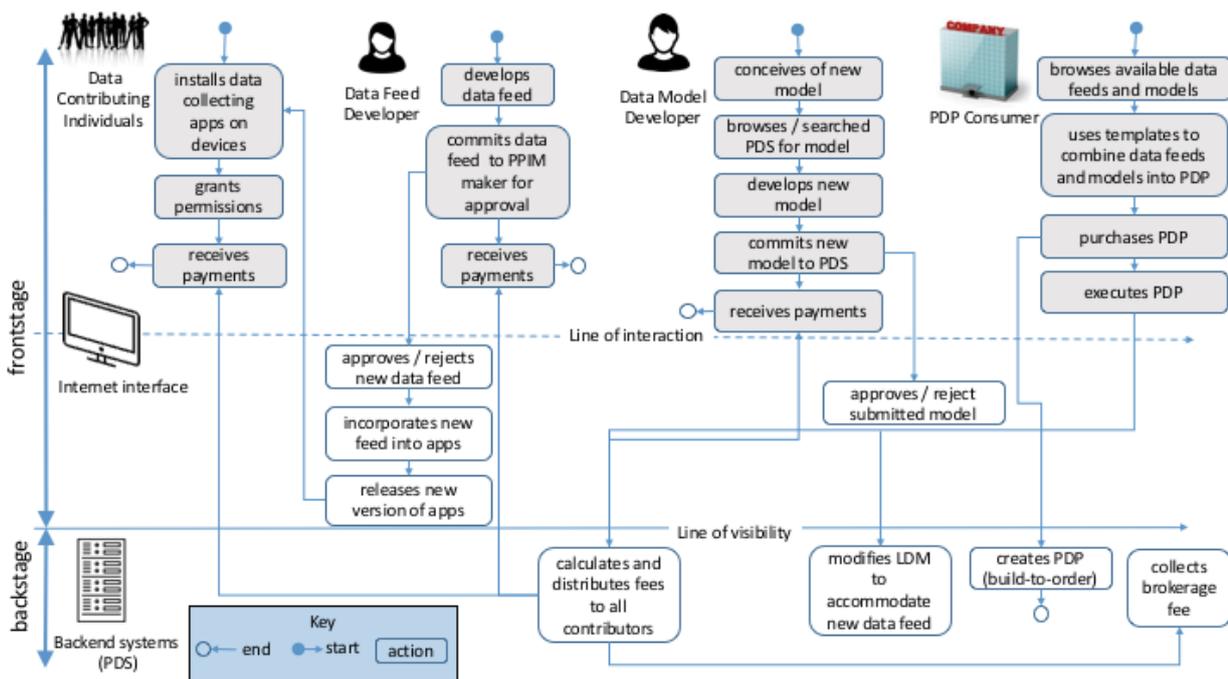


Figure 19: Service experience blueprint for participation on a PPIM

Data feed developers' behavioural interaction with the PDS involves developing and contributing ingestion mechanisms for new devices or applications into the PDS via the PDS API. They match data fields between the device API and the PDS API and contribute to the PDS data model to model how the new data feed will be ingested and stored in the PDS and how it will relate to the other personal data in the PDS. Once the new data feed has been established the developer's ID is associated with that data feed and whenever a PDP comprising data from the developer's data feed is executed credit is

associated with the developer’s account. Using the developer facing dashboard the developer can view, manage and track funds in his or her account.

Model developers’ behavioural interaction with the system involves developing and contributing models that extract or infer information from the raw personal data stored in the PDS. Once the new model has been contributed to the PDS the developer’s ID is associated with that model and whenever a PDP comprising information extracted or inferred using the developer’s model is executed, credit is associated with the developer’s account. Using the developer facing dashboard the developer can view, manage and track funds in his or her account.

End users’ cognitive requirements are catered for via a consumer facing dashboard which connects to the PDS via an API. The end user utilises the dashboard to browse, co-create, preview, execute and renew PDPs – thereby gaining knowledge of the products he or she has purchased and used. The emotive requirement for trust in the quality of the PDP is enhanced through the rating and customer feedback mechanisms.

Figure 20 shows a screen shot of the end user facing dashboard. End users experience this service as the means by which they can interact (co-create) with the personal information of individuals in a novel, legal, transparent and fair way.

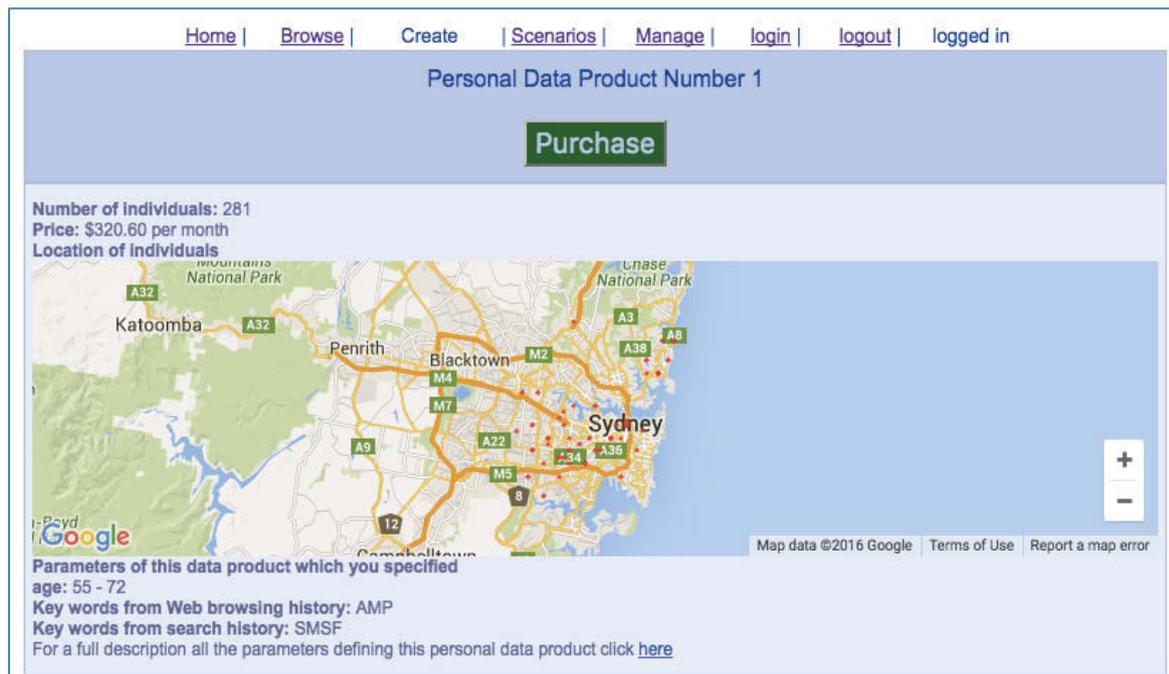


Figure 20: Screen shot of prototype end user facing dashboard

8.4.7 Monetisation design

The PPIM is monetised on a percentage pay-per-use model (de Reuver et al. 2015) (see Figure 19 above). Whenever a PDP is executed funds are transferred from the end user's account to the PDP contributors accounts (model developer, data feed developer and data contributing individuals). A brokerage fee is calculated as a percentage of the transaction and transferred to the PPIM maker's account.

8.5 Prototype

Drawing on previous research (Farrelly & Chew 2016c), which examines the kinds of personal information individuals would be willing to share on a PPIM, a database of individuals and their associated permission granting activity was simulated. Using 3,603 permissions granted by 203 participants in a previous study (Farrelly & Chew 2016c) 43,200 permissions using stratified sampling were simulated. The full code used to generate this simulated data can be found in section 14.4.

A web interface was then developed in line with the above *iSIM* design to present the PPIM as a service system. This interface facilitated the provisioning of two popular end uses to which companies are likely to put personal data: personalised customer service (Gualtieri 2013; World Economic Forum and Bain & Company 2011), and targeted marketing (Rose, Rehse & Röber 2012, p. 25). This facility was combined with tools, which allowed the end user to preview, purchase, execute and manage these PDPs to complete the prototype (see figures Figure 21 to Figure 34).

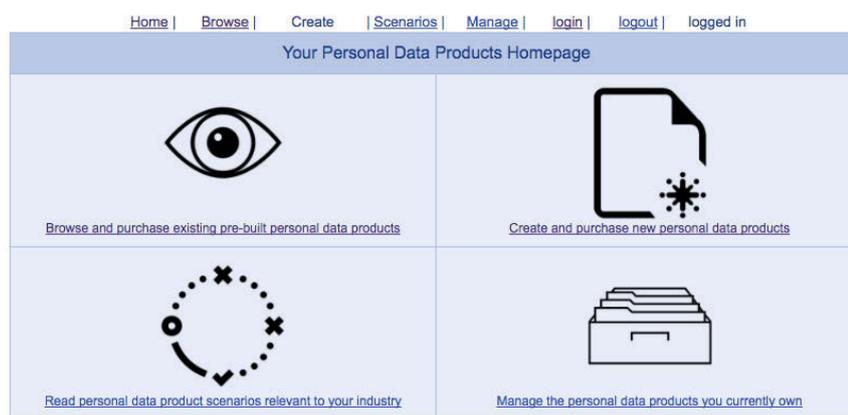


Figure 21: Home page of simulated PPIM

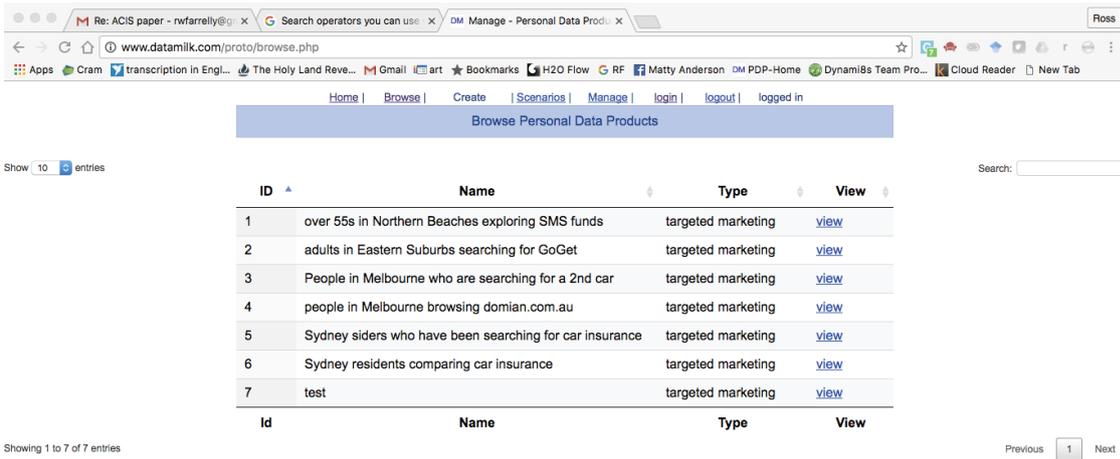


Figure 22: Browse page of simulated PPIM

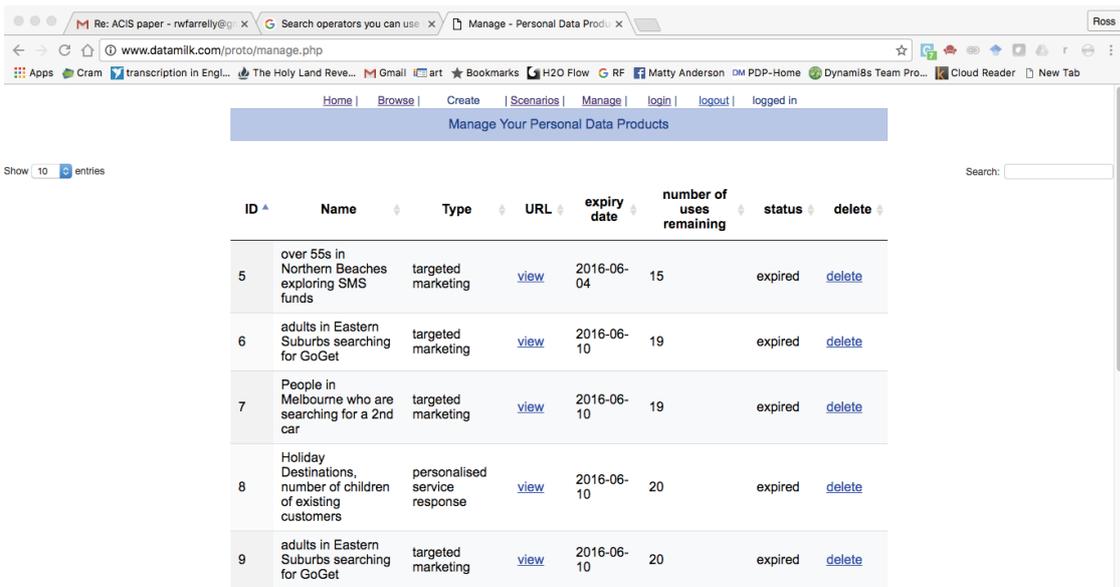


Figure 23: Manage page of simulated PPIM

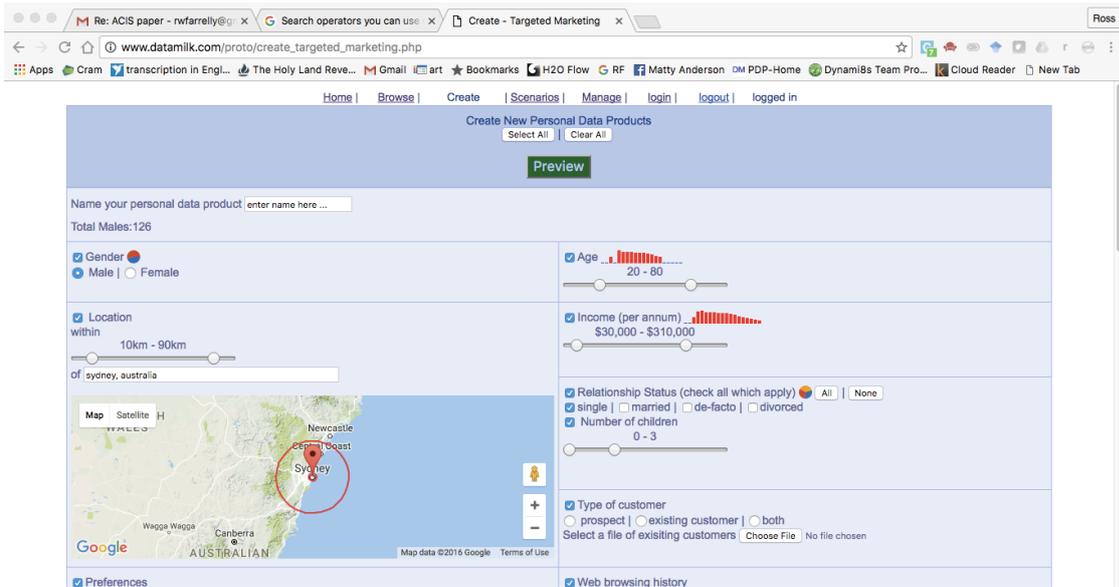


Figure 24: First part of create targeted marketing PDP of simulated PPIM

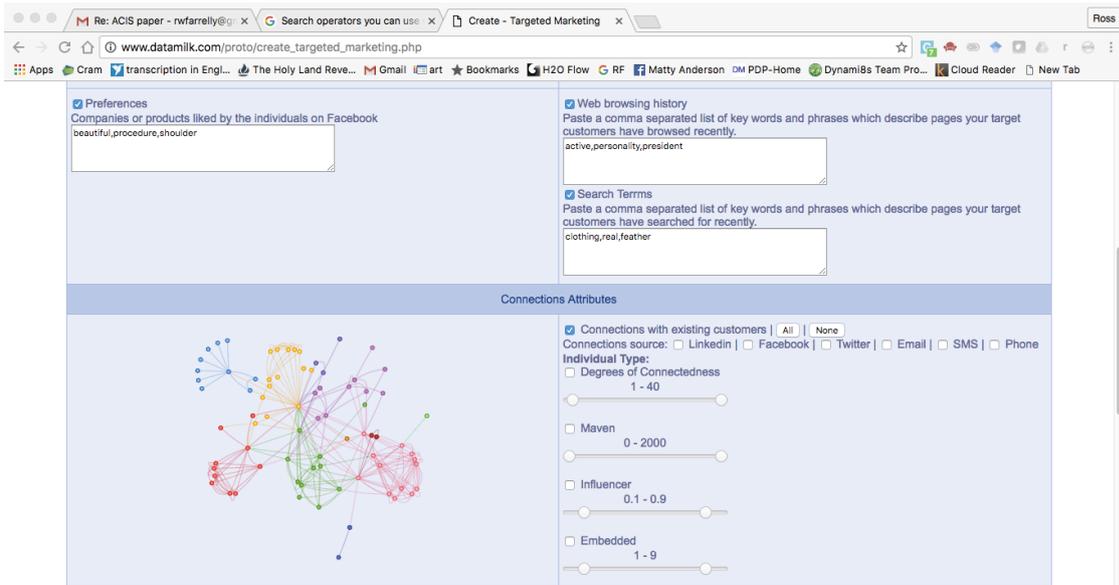


Figure 25: Second part of create targeted marketing PDP of simulated PPIM

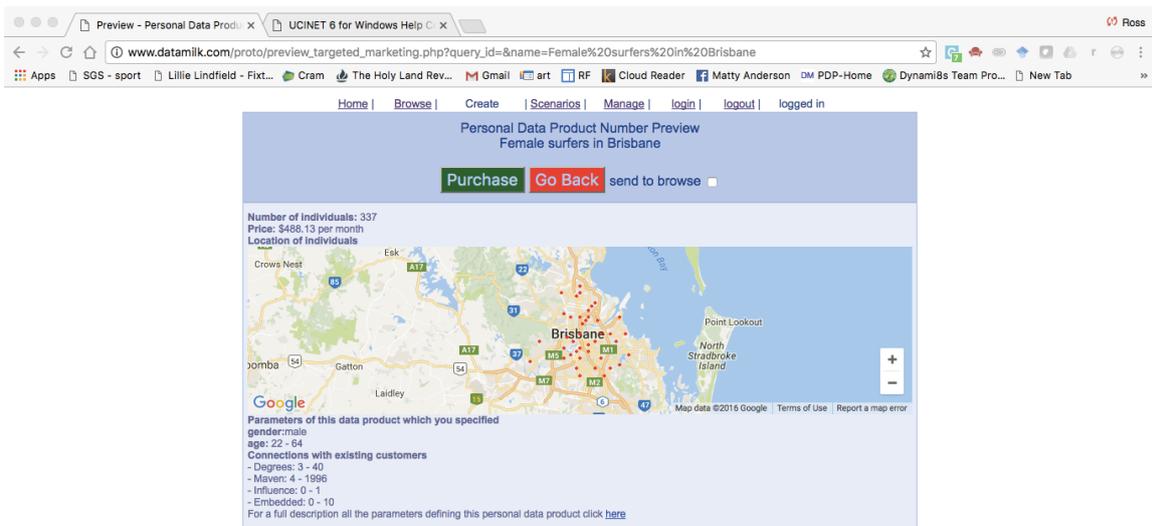


Figure 26: Preview page of targeted marketing PDP on simulated PPIM

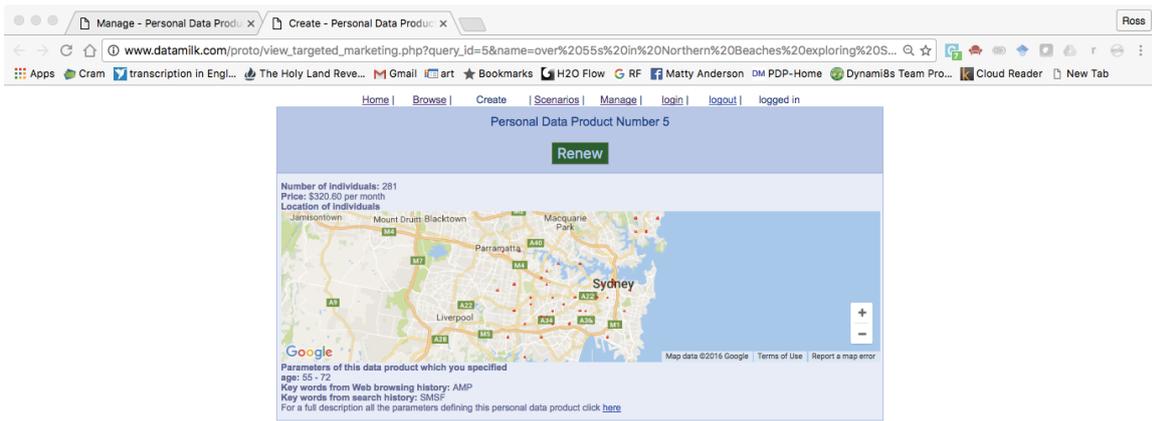


Figure 27: View and renew page of targeted marketing PDP on simulated PPIM

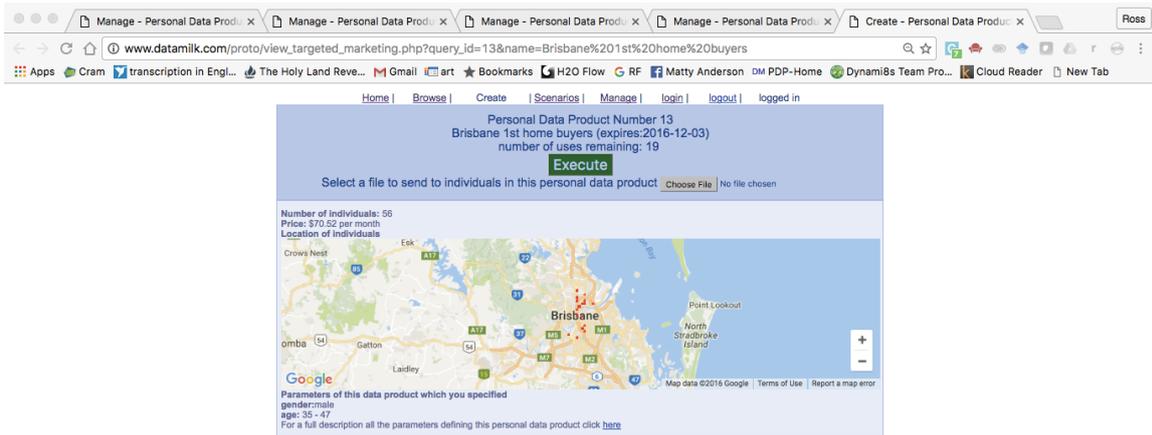


Figure 28: View and execute page of targeted marketing PDP on simulated PPIM

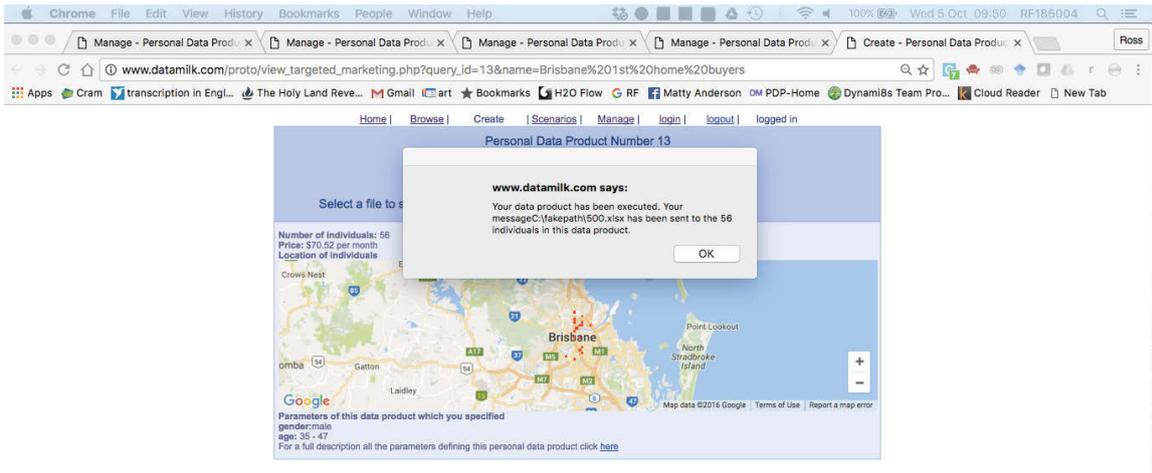


Figure 29: Success notification page targeted marketing PDP on simulated PPIM

The screenshot shows the 'Create New Personalised Service Response Data Product' form. It includes a 'Preview' button and instructions to select information for a dashboard. The form is divided into several sections with selection options:

- Individual:** Includes checkboxes for Age, Gender, Ethnic Background, Relationship Status, Income, Religion, Current Employment, Employment History, Education, Home Postcode, Work/Study Postcode, Accommodation Type, Holiday Destinations, and Mode of Transport to work/study/passtimes, sports and hobbies.
- Family Relationships:** Includes checkboxes for Number of parents (living), Number of siblings, and Number of children.
- Text Based Information:** Includes a 'Keywords' section with a text input field containing 'mobile phone, mobile plan', and checkboxes for Recent searches, Recent internet browsing, Recent Facebook Likes, and Recent Facebook Posts. It also has a slider for 'Number of months history to take into account'.
- Social Network Information:** Includes a checkbox for Connections to existing customers.
- Life Events:** Includes checkboxes for recent, forthcoming, and no longer possible.
- Inferred Information:** Includes checkboxes for political leanings, issues which are important to individual such as price, attitude, promptness etc, and sexual orientation.

Figure 30: First part of create personalised service response PDP of simulated PPIM

<input type="checkbox"/> All <input type="checkbox"/> None <input type="checkbox"/> Age <input type="checkbox"/> Gender <input type="checkbox"/> Ethnic Background <input type="checkbox"/> Relationship Status <input type="checkbox"/> Income <input type="checkbox"/> Religion <input type="checkbox"/> Current Employment <input type="checkbox"/> Employment History <input type="checkbox"/> Education <input type="checkbox"/> Home Postcode <input type="checkbox"/> Work/Study Postcode <input type="checkbox"/> Accommodation Type <input type="checkbox"/> Holiday Destinations <input type="checkbox"/> Mode of Transport to work / study <input type="checkbox"/> passtimes, sports and hobbies	<input type="checkbox"/> All <input type="checkbox"/> None <input type="checkbox"/> Number of parents (living) <input type="checkbox"/> Number of siblings <input type="checkbox"/> Number of children
Text Based Information <input type="checkbox"/> All <input type="checkbox"/> None <input type="checkbox"/> Keywords Key words you are interested in: mobile phone, mobile plan <input type="checkbox"/> Recent searches <input type="checkbox"/> Recent internet browsing <input type="checkbox"/> Recent Facebook Likes <input type="checkbox"/> Recent Facebook Posts Number of months history to take into account <input type="range"/>	Social Network Information <input type="checkbox"/> All <input type="checkbox"/> None <input type="checkbox"/> Connections to existing customers
Life Events <input type="checkbox"/> All <input type="checkbox"/> None <input type="checkbox"/> recent <input type="checkbox"/> forthcoming <input type="checkbox"/> no longer possible	Inferred Information <input type="checkbox"/> All <input type="checkbox"/> None <input type="checkbox"/> political leanings <input type="checkbox"/> Issues which are important to individual such as price, attitude, promptness etc <input type="checkbox"/> sexual orientation
Possessions / Suppliers <input type="checkbox"/> All <input type="checkbox"/> None <input type="checkbox"/> vehicles <input type="checkbox"/> Insurances <input type="checkbox"/> Financial Institutions <input type="checkbox"/> Superannuation <input type="checkbox"/> technology (smart phones, PCs etc) <input type="checkbox"/> Utilities <input type="checkbox"/> Financial advisor <input type="checkbox"/> Magazine / media subscriptions <input type="checkbox"/> Real Estate <input type="checkbox"/> Pay TV <input type="checkbox"/> Share portfolio <input type="checkbox"/> Telcos	
Custom <input type="checkbox"/> Custom Describe the information you would like to know about your customers and/or prospects. We will match your requirements to existing information we have or if there is no suitable match, we will submit your request to our developers and get back to you if that information can be extracted from the information we have.	

Figure 31: Second part of create personalised service response PDP of simulated PPIM

<input type="checkbox"/> All <input type="checkbox"/> None <input type="checkbox"/> Age <input type="checkbox"/> Gender <input type="checkbox"/> Ethnic Background <input type="checkbox"/> Relationship Status <input type="checkbox"/> Income <input type="checkbox"/> Religion <input type="checkbox"/> Current Employment <input type="checkbox"/> Employment History <input type="checkbox"/> Education <input type="checkbox"/> Home Postcode <input type="checkbox"/> Work/Study Postcode <input type="checkbox"/> Accommodation Type <input type="checkbox"/> Holiday Destinations <input type="checkbox"/> Mode of Transport to work / study <input type="checkbox"/> passtimes, sports and hobbies	<input type="checkbox"/> All <input type="checkbox"/> None <input type="checkbox"/> Number of parents (living) <input type="checkbox"/> Number of siblings <input type="checkbox"/> Number of children
Text Based Information <input type="checkbox"/> All <input type="checkbox"/> None <input type="checkbox"/> Keywords Key words you are interested in: mobile phone, mobile plan <input type="checkbox"/> Recent searches <input type="checkbox"/> Recent internet browsing <input type="checkbox"/> Recent Facebook Likes <input type="checkbox"/> Recent Facebook Posts Number of months history to take into account <input type="range"/>	Social Network Information <input type="checkbox"/> All <input type="checkbox"/> None <input type="checkbox"/> Connections to existing customers
Life Events <input type="checkbox"/> All <input type="checkbox"/> None <input type="checkbox"/> recent <input type="checkbox"/> forthcoming <input type="checkbox"/> no longer possible	Inferred Information <input type="checkbox"/> All <input type="checkbox"/> None <input type="checkbox"/> political leanings <input type="checkbox"/> Issues which are important to individual such as price, attitude, promptness etc <input type="checkbox"/> sexual orientation
Possessions / Suppliers <input type="checkbox"/> All <input type="checkbox"/> None <input type="checkbox"/> vehicles <input type="checkbox"/> Insurances <input type="checkbox"/> Financial Institutions <input type="checkbox"/> Superannuation <input type="checkbox"/> technology (smart phones, PCs etc) <input type="checkbox"/> Utilities <input type="checkbox"/> Financial advisor <input type="checkbox"/> Magazine / media subscriptions <input type="checkbox"/> Real Estate <input type="checkbox"/> Pay TV <input type="checkbox"/> Share portfolio <input type="checkbox"/> Telcos	
Custom <input type="checkbox"/> Custom Describe the information you would like to know about your customers and/or prospects. We will match your requirements to existing information we have or if there is no suitable match, we will submit your request to our developers and get back to you if that information can be extracted from the information we have.	

Figure 32: Third part of create personalised service response PDP of simulated PPIM

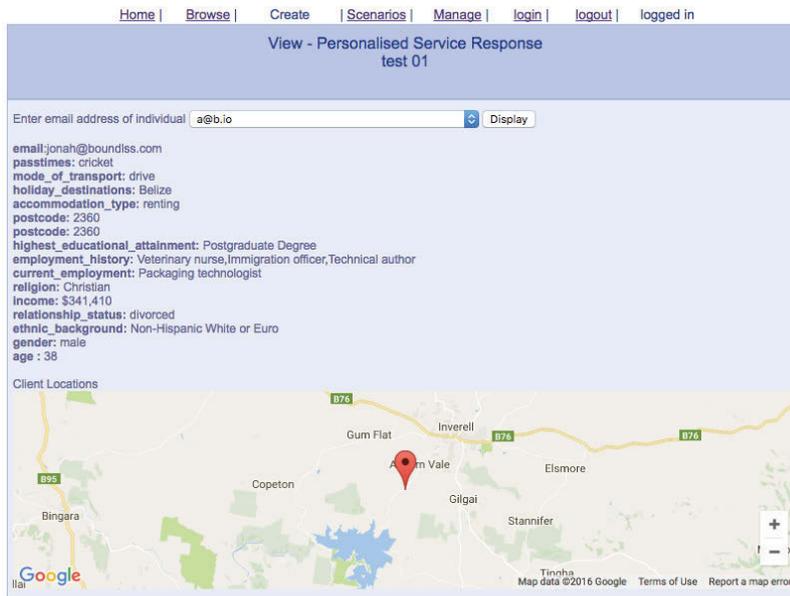


Figure 33: Preview of personalised service response PDP of simulated PPIM

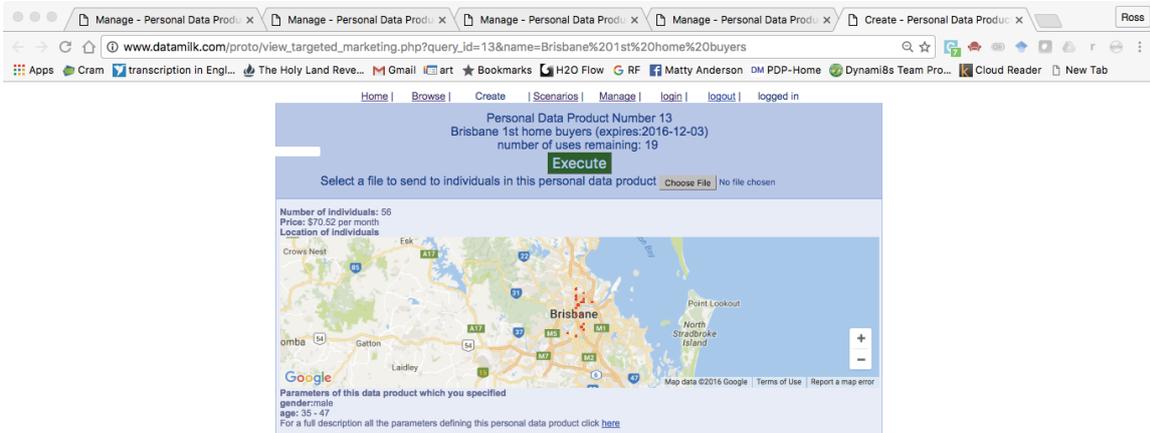


Figure 34: Screen display when executing personalised service response PDP of simulated PPIM

8.6 Evaluation

Following Yin (Yin 2013), the main themes in the interviewees' responses are identified, compared and contrasted. Of the eight experts interviewed, all expressed broadly the opinion that the PPIM would have utility for the businesses in which they worked. Table 17 contains indicative supportive quotations from each of the interviewees. One interviewee offered to be part of a pilot program to test the proposed PPIM in practice.

One notable aspect of the interviews was the ability for interviewees to identify and propose novel ideas for PDPs which would be of relevance to their industries. Novel data products proposed include: an auto-form filling PDP which would reduce friction between end user and customer, personalised risk estimation for insurance, personalised policy offers based on expiry of current policies and a recruitment PDP to allow companies to go direct to potential new hires, by-passing recruitment firms possibly implemented via a candidate search tool, a customer identification and authentication PDP, a driver classification PDP based on driving behaviour data and PDPs which allow companies to infer customer intention based on geo-location data.

In addition to in-principle agreement that the proposed PPIM would be of utility, interviewees also proposed numerous additional features which would improve its viability. Such suggestions include the ability to define a segment of interest and be alerted when a pre-set number of individuals join the PPIM in that segment. Another suggestion made by multiple interviewees was the need for a mechanism whereby the return on investment could be measured by the end user using the PPIM.

Table 17: Indicative supportive quotations from interviewees

Industry class	size	Indicative quotation
Finance and Insurance (1)	small	'If it works and works well it would an incredibly powerful tool.'
Personal and Other Services (1)	large	'No question about it. Incredibly valuable.'
Accommodation, Cafes and Restaurants	micro	'It could, there would be a huge number of uses [to which it could be put].'
Personal and Other Services (2)	med	'Yes, this system would have benefits for companies, - it would reduce cost of acquisition.'
Government Administration and Defence	med	'I really like it ... [it's a very] interesting concept, I like the prototype.'
Finance and Insurance (2)	large	'It's almost limitless the things you could do ... [with such a platform].'
Communication Services	small	'It would be gold for B2C companies.'
Education	med	'We would certainly use this to reach new clients.'

8.7 Conclusion

Using the *iSIM* (Chew 2015) and consistent with the Design Science methodology (Hevner et al. 2004; Peffers et al. 2006; Peffers et al. 2007-8) this chapter comprises a design for a PPIM as an industry platform. The proposed design is instantiated as a prototype which was evaluated using qualitative interview methods.

This chapter reveals that the proposed industry platform model for the PPIM incorporating open access to model developers and data feed developers resolves the scalability concerns raised in prior research by obviating the need for a single organisation to have to develop all data ingestion and data modelling aspects of the PPIM in house.

All the interviewed experts broadly agreed that the proposed PPIM as an industry platform represents a service that would have multiple applications to a multiplicity of business types. They also demonstrated an ability to devise and envision novel data products (of value to the end users' businesses) which would be facilitated by such a market.

This finding is significant as it lends weight to the argument that the industry platform conceptualisation of a PPIM has broad appeal to business of varying sizes and industry types. It also confirms the efficacy of *iSIM* as an effective method to design and commercialise a novel industry platform service.

The research journey thus far has taken us from a two-sided PPIM to a four-sided PPIM designed as an industry platform. Following the evaluation the PPIM designed as a four-sided industry platform it was clear that a more detailed design was required to progress towards the realisation of a workable PPIM. When searching for a suitable paradigm within which to provide this additional level of detail it became clear that there was a more suitable alternative to industry platform theory.

Industry platforms have been defined as 'products, services, or technologies that ... provide the foundation upon which outside firms ... can develop their own complementary products, technologies, or services' (Gawer & Cusumano 2014, p. 418). Most (but not all) of today's industry platforms are digital and many have physical elements included in the product offering (Evans & Gawer 2016). Companies such as 'Microsoft, Apple, Google, Intel, Cisco, ARM, Qualcomm, EMC, and many other firms, small and large, build hardware and software products for computers, cell phones, and consumer electronics devices that in one form or another serve as what we can call industry platforms.'

On the other hand, digital platforms exist purely in the digital realm. They have been defined from a technical viewpoint as 'an extensible codebase to which complementary third-party modules can be added' and from a sociotechnical view as 'Technical elements (of software and hardware) and associated organisational processes and standards' (de Reuver, Sørensen & Basole 2017).

Developing the work of Tiwana & Konsynski (2010), Ghazawneh & Henfridsson (2015) defined digital platforms as: 'software-based external platforms consisting of the extensible codebase of a

software-based system that provides core functionality shared by the modules that interoperate with it and the interfaces through which they interoperate' (Ghazawneh & Henfridsson 2015, p. 199).

Since the purpose of a PPIM is the collection, analysis and monetisation of personal information that takes place in a purely digital realm, I argue that a PPIM can be better classified as a digital platform. Furthermore, following Evans & Gawer (2016, p. 7), it is claimed that a PPIM falls under the heading of 'innovation platforms' which 'consist of technological building blocks that are used as a foundation on top of which a large number of innovators can develop complementary services or products'. In the context of a PPIM, these innovators are the developers who build the models and data feeds which facilitate the creation of Personal Data Products (PDPs) by end users.

A more detailed third iteration of the PPIM design, this time conceived as a digital platform is the subject of the next chapter.

9 PPIM conceived as a digital platform

In this chapter a third iteration of the design process is presented in which we conceive of the PPIM as a digital platform (as opposed to an industry platform).

This chapter is largely based on the material in Farrelly & Chew (forthcoming).

The rise of digital platforms shows that they have the potential to transform existing industries (Constantinides, Henfridsson & Parker 2018; de Reuver, Sørensen & Basole 2017; Evans & Gawer 2016; Van Alstyne & Parker 2017). But do they also have the potential to create new markets where none currently exist? Do they have the potential to address the growing discontent with the one-sided relationship between corporations that gather and exploit the value of individuals' personal information, and the relatively powerless individuals who originally created the data? As Constantinides, Henfridsson & Parker (2018, p. 13) put it:

The question ... everyone ... might reasonably ask is, how much value is being created from their data, and is there some way they can benefit as well?

The research agenda on digital platforms has evolved from the study of intermediaries (Rose 1999), infomediaries (Leickly 2004), two-sided markets (Evans 2009; Evans & Schmalensee 2005) and industry platforms (Gawer & Cusumano 2014). Prior to the technological advances such as parallel computing, cloud computing and the widespread adoption of the internet and mobile devices which have fuelled platform growth (Basole & Karla 2011), it was recognised that intermediaries had a part to play in bringing together two sides of certain markets (Bakos & Bailey 1997; Rose 1999; Sarkar, Butler & Steinfield 1995). More recently, the role of platforms has been studied extensively from both the business perspective (Gawer & Cusumano 2008), the economic perspective (Parker & Van Alstyne

2014; Parker, Van Alstyne & Jiang 2016) and an industrial architectural perspective (Tee & Gawer 2009).

Tiwana (2013, p. 7) defines a platform as ‘The extensible codebase of a software-based system that provides core functionality shared by apps that interoperate with it, and the interfaces through which they interoperate’. A digital platform has been defined from a technical viewpoint as ‘an extensible codebase to which complementary third-party modules can be added’ and from a sociotechnical view as ‘Technical elements (of software and hardware) and associated organisational processes and standards’ (de Reuver, Sørensen & Basole 2017). Evans & Gawer (2016) define a digital platform as a platform that captures, transmits and monetises data, including personal data, over the internet.

Digital platforms have changed the way value can be created by companies. Rather than controlling the value chain, platforms create an ecosystem which attracts value-adding participants (Van Alstyne, Parker & Choudary 2016). Rather than using command-and-control top down decision-making to control the company, platforms seek to incentivise and influence participants to undertake value creating activities, shifting focus of value creation to external contributors (Parker, Van Alstyne & Jiang 2016).

Digital platforms can also foster online communities of consumers and contributors through network effects (Spagnoletti, Resca & Lee 2015) thereby building a clearing house and innovation locus which facilitates third-party participants such as app developers to develop and market products of their own (Parker, Van Alstyne & Jiang 2016). Furthermore, designing for a modular rather than a monolithic architecture of the platform (Tiwana & Konsynski 2010) allows for an alignment between architecture and strategy and created conditions conducive to platform evolution (Tiwana 2013).

Since the purpose of a PPIM is the collection, analysis and monetisation of personal information, which takes place in a purely digital realm, it is claimed that a PPIM can be classified as a digital platform. Furthermore, following (Evans & Gawer 2016, p. 7), it is asserted that a PPIM falls under the heading of ‘innovation platforms’ which ‘consist of technological building blocks that are used as a foundation on top of which a large number of innovators can develop complementary services or products’. In the context of a PPIM, these innovators are the developers who build the models and data feeds that facilitate the creation of Personal Data Products (PDPs) by end users. Previous research (Farrelly & Chew 2017) has shown that business owners (i.e. end users of a PPIM), selected from multiple industries for a prototype PPIM case study, were capable of utilising the PPIM prototype to conceive, devise and construct multiple innovative personal data products which were previously unavailable to them.

9.1 A service architecture for a PPIM

In this section, a service architecture for the PPIM platform is developed which is required to execute the PPIM service strategy. Drawing on service architecture theories (Broekhuis, van Offenbeek & van der Laan 2017; Eissens-van der Laan et al. 2016) and the extant literature on platform growth and scalability (Gawer & Cusumano 2014; Tiwana 2013; Van Alstyne, Parker & Choudary 2016) we outline a set of service architecture principles which, when applied to a PPIM, will ensure innovation and growth. The aim is not to develop a perfectly modular architecture because modularity has both benefits and drawbacks and, in fact, designing a perfectly modular platform architecture is almost impossible (Tiwana 2013). We aim for an architecture that has an intermediate level of modularity which has been found to produce the most useful level of innovation (Ethiraj & Levinthal 2004).

Since service architecture is an active and evolving field of research (Brax et al. 2017), a number of key terms are defined here for the sake of clarity and to avoid ambiguity.

Service architecture: ‘the way that the functionalities of the service system are decomposed into individual functional elements to provide the overall services delivered by the system. At each level of decomposition the architecture can be either integral or modular’ (Voss & Hsuan 2009, p. 546)

Module: ‘a system of components that offers a well-defined functionality via a precisely described interface and with which a modular service is composed, tailored, customized, and personalized.’ (Tuunanen, Bask & Merisalo-Rantanen 2012, p. 101)

Interface: ‘the set of rules and guidelines governing the flexible arrangement, interconnections, and interdependence of service components and service providers’ (de Blok et al. 2014, p. 186)

In designing a service architecture for a PPIM three core modularity principles are adopted which have been found to be implemented by design professionals (Broekhuis, van Offenbeek & van der Laan 2017) namely: that modules should have a specific function, that modules should be relatively independent and that modules should have standardised interfaces for integration purposes (Baldwin & Clark 2000; Ulrich 1995).

Furthermore, when making a design choice concerning how the modules will be arranged, a *multilevel combined orientation architecture* comprising one layer of service oriented modules and a second layer of sub-modules modules is adopted. This design choice provides flexibility since it allows the service offering to be decomposed into multiple phases on multiple decomposition levels. It is also capable of accommodating both pooled and sequential dependencies. That is to say dependencies that need to occur one after another (sequentially) and those that need to occur simultaneously (pooled).

This design choice is applicable in a start-up setting in which it is likely that there will be a high variation in both demand and throughput uncertainty (Eissens-van der Laan et al. 2016).

9.2 Design of base service modules

Using as a starting point the Service experience blueprint for participation on a PPIM (Figure 19Figure 19), which has been validated by external practitioners (see section 8.6), by examining the core processes involved in a PPIM and correlating these with the requirements of each of the four participant types, the following standardised base service modules are identified (Table 18) which are consistent with the modular design principles of specificity, independence and standardisation. These standardised base service modules constitute a modular service design (Broekhuis, van Offenbeek & van der Laan 2017) for a PPIM.

Table 18: Process modules of a PPIM service architecture

Module	Standardised base service modules
approve module	Approves or rejects an asset which has been proposed to be included on the PPIM platform
browse module	Enables participants to browse assets on the PPIM platform
create module	Enables participants to create assets on the PPIM platform
communications module	Enables communications between platform participants
distribute module	Distributes funds to participants on the PPIM platform
execute module	Executes a PDP on the PPIM platform
monitor module	Monitors usage of assets on the PPIM platform
order module	Enables participants to place an order on the PPIM platform
payment module	Enables participants to pay or receive funds on the platform
preview module	Enables participants to preview assets on the PPIM platform
rank module	Ranks assets (such as models) on the PPIM platform
renew module	Renews an asset on the PPIM platform
report module	Enables participants to report issues and change requests to the platform owner
request module	Enables participants to send and receive requests on the platform
resolve module	Resolves disputed transactions on the PPIM platform
search module	Enables participants to search for assets on the PPIM platform

implement module	Implements a policy on the platform
insert module	Inserts an asset into the PPIM platform
update module	Updates assets on the PPIM platform

The ability for customers to reuse standard offerings and derive variations of standard offerings are important outcomes of well-designed modules (Bask, Merisalo-Rantanen & Tuunanen 2014). Also important is the mixing and matching of the modular service units within a PPIM to provide different combinations or services to meet specific customer and market needs (de Blok et al. 2014). One example of how a module can be reused and varied is as follows: the search module is a service module that executes a search of a database and returns a set of search results sorted by relevance. Data contributing individuals will leverage this service module when searching for PDPs in which their personal information is included, while model developers would leverage this service when searching for models similar to the one they have in mind before proposing a new model for inclusion on the platform. Data feed developers would utilise the search module in a similar way but would search for data feeds rather than models. End users will utilise the search model by searching for PDP templates and models from which they can construct new PDPs. Table 19 presents a reuse matrix showing which groups of participants would utilise the available service modules.

Table 19: Reuse and variation matrix of process modules of a PPIM service architecture

Module	data contributing individual	model developer	data feed developer	end user	platform owner
approve module	✓	✓	✓	✓	✓
browse module	✓	✓	✓	✓	✓
create module	✗	✓	✓	✓	✗
communications module	✓	✓	✓	✓	✓
distribution module	✓	✓	✓	✓	✓
monitor module	✓	✓	✓	✓	✓
order module	✗	✗	✗	✓	✓
preview module	✗	✓	✓	✓	✓

rank module	✗	✓	✓	✓	✓
payment module	✓	✓	✓	✗	✓
renew module	✗	✗	✗	✓	✗
report module	✓	✓	✓	✓	✓
resolve module	✓	✓	✓	✓	✓
search module	✓	✓	✓	✓	✓
insert module	✗	✓	✓	✓	✗
implement module	✗	✗	✗	✗	✓
update module	✗	✓	✓	✓	✗

9.3 Core processes involved in a PPIM

By referring to the PPIM service experience blueprint (Figure 19) the key activities of each of the four participant classes needed to perform on a PPIM are identified. These activities are listed in the tables numbered Table 20 to Table 24. These tables also describe the reuse or variation of the base service module which is needed to facilitate the activity.

Table 20: Activities, service modules and reuse submodules for end users on a PPIM

Activity	Service module	Reuse submodule
Browses available model components to build PDP	browse module	browses models
Browses available template from which to construct PDPs	browse module	browses templates
Selects and customises a template to construct a new PDP	create module	creates PDP from template
Constructs a PDP from scratch	create module	creates from blank template
Previews PDP	preview module	previews PDPs prior to submitting them to the platform
Modifies PDP	create module	creates by modifying existing PDP
Submits PDP	report module	reports requested change to platform owner in the form of the inclusion of a new PDP
Purchases PDP	payment module	makes payments to platform owner to obtain permission to execute PDPs
Executes PDP	execute module	executes a PDP
Renews PDP	update module	renews subscription to a PDP
Update PDP	update module	updates the parameters of a PDP
Check balance	payment module	checks payments have been made
Adds funds / pays invoices	payment module	pays into end user's account
Measures ROI on PDPs	payment module, monitor module	compares usage of PDPs with payments made for those PDPs

Table 21: Activities, service modules and reuse submodules for data contributing individuals on a PPIM

Activity	Service module	Reuse submodule
Installs app	occurs off platform	
Selects which data fees to capture	occurs off platform	
Receives request to have his/her data included in PDPs	communications module	platform communicates request to data contributing individual
Accepts / rejects requests	approve module	approves or rejects request to include their personal information in the specified PDP
Receives payments	payment module	payments for usage of personal information in PDPs is received into account
Receives messages from companies	communications module	receives messages from companies on an opt in basis
Browses PDPs in which his / her data is included	browse module	browses PDPs
Browses companies which are using his / her data	browse module	browses companies
Browses uses to which his / her data is being put	browse module	browses usage

Table 22: Activities, service modules and reuse submodules for model developers on a PPIM

Activity	Service module	Reuse submodule
Browses existing models	browse module	browses existing models to identify opportunities to develop new models
Searches for existing models	search module	searches for existing models similar to new models which could be developed
Develops new models	create module	develops new models for inclusion on the platform
Commits new models	communication module	communicates to the platform owner that a new model has been developed and applied for inclusion of the model on the platform
Manages usages of models	monitor module	monitors the uses to which his or her models are being put
Views status of models	monitor module	views status and relative ranking of his or her models relative to the performance of other similar models
Receives payment for use of models	payment module	payments for usage of models is received into account
Disputes payments for use of models	report module, payment module	reports payments issues to platform owner based on model usage
Reports abuse of model usage	report module	reports unauthorised usage of models

Table 23: Activities, service modules and reuse submodules for data feed developers on a PPIM

Activity	Service module	Reuse submodule
Browses for existing data feeds	browse module	browses existing data feeds to identify opportunities to develop new data feeds
Searches for existing data feeds	search module	searches for existing data feeds similar to new data feeds which could be developed
Develops new data feeds	create module	develops new data feeds for inclusion on the platform
Commits new data feeds	report module	communicates to the platform owner that a new data feed has been developed and applied for inclusion of the model on the platform
Manages usages of data feeds	monitor module	monitors the uses to which his or her data feeds are being put
View status of data feeds	monitor module	views status of his or her data feeds
Receives payment for use of data feeds	payment module	payments for usage of data feeds is received into account
Disputes payments for use of data feeds	report module, payment module	reports payments issues to platform owner based on data feed usage
Reports abuse of model data feeds	report module	reports unauthorised usage of data feeds

Table 24: Activities, service modules and reuse submodules for platform owners on a PPIM

Activity	Service module	Reuse submodule
Develops and distributes data collecting apps	create module, communications module	creates, maintains and distributes the data collection apps which will reside on the mobile and desktop devices of the data contributing individuals
Approves models and data feeds submitted for inclusion on the platform	approve module	approves / rejects applications to include data feeds, models and PDPs on the platform
Incorporates new data feeds into data collecting apps	update module	updates and maintains the data collection app as new data feeds are developed to collected data from new devices as they enter the market
Incorporates new models into data platform	update module	makes the insights made available by the development of new models available to end users to include in their PDPs
Calculates fees due to participants	payment module	calculates fees bases on revenue sharing policies
Distributes fees due to participants	payment module	distributes fees to data contributing individuals, model developers and data feed developers
Implements pricing policies	update module, communications module	updates pricing policies to maintain competitive advantage of the platform and communicates these changes to participants
Implements governance policies	update module, communications module	updates governance policies in accordance with decision rights to maintain competitive advantage of the platform and communicates these changes to participants

In addition to the base process modules there are associated reuse submodules (where appropriate) to embody the ways in which the base module is reused. The submodules are illustrated in the following scenarios.

9.4 Scenarios enabled by the service architecture

In this section, adopting the Modular Service Architecture Framework of Bask, Merisalo-Rantanen & Tuunanen (2014), the scenarios commonly engaged in by the various participants on a PPIM are described. It is demonstrated how the service modules facilitate the execution of these modules. These scenarios are fulfilled by mixing and matching the standardised base service modules to provide different combinations to meet specific participant's needs (Bask, Merisalo-Rantanen & Tuunanen 2014). The participant can avail him or herself of the base service module by either utilising it to achieve his or her ends or receiving the output of one of the modules after it has been utilised by another participant.

In Figure 35 to Figure 39 five scenarios are presented, which demonstrate how the platform participants utilise the service modules to execute commonly performed workflows in the PPIM platform.

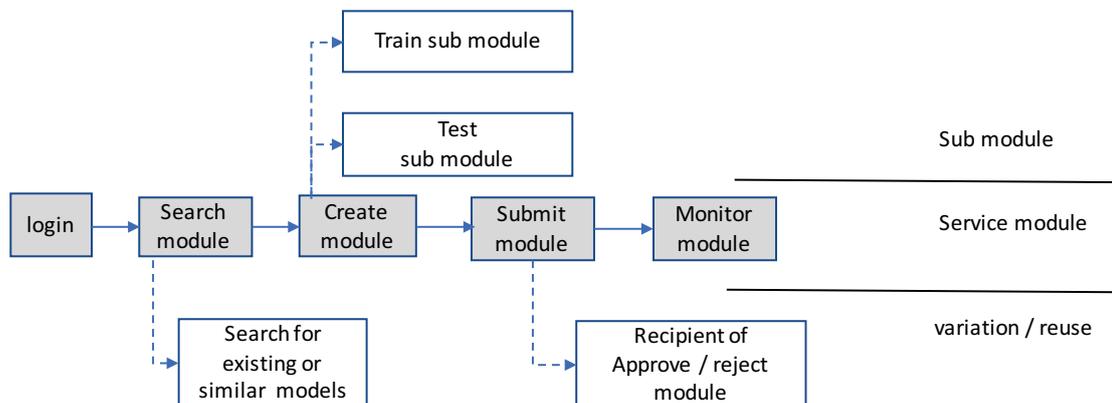


Figure 35: Model developer develops new model

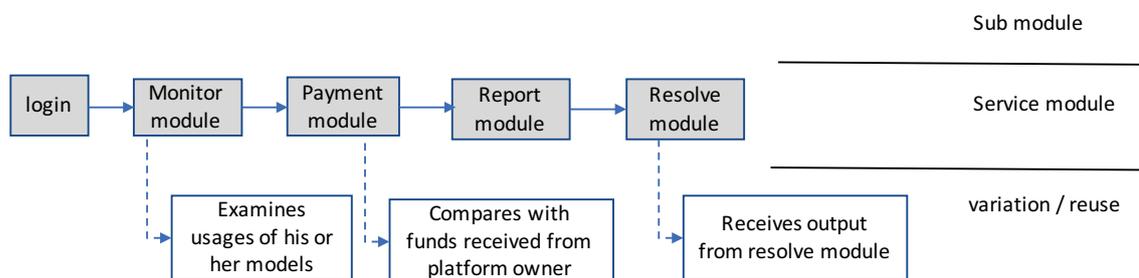


Figure 36: Model developer checks on usage of his or her models and reconciles against payments

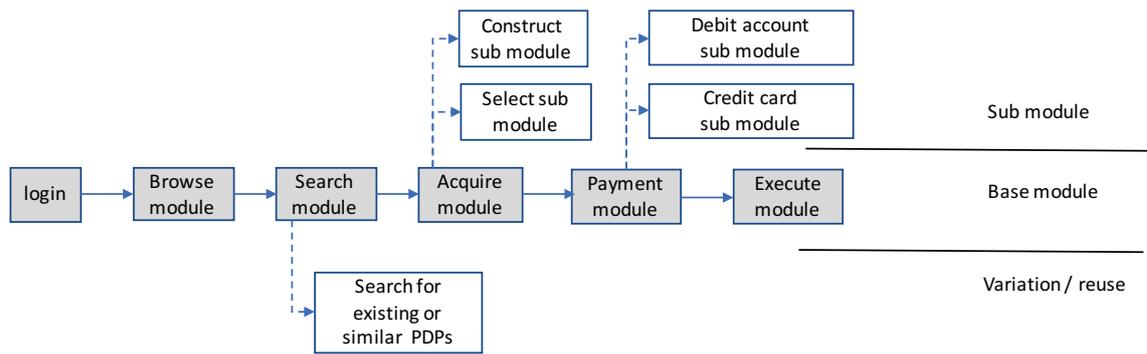


Figure 37: End user creates PDP from template

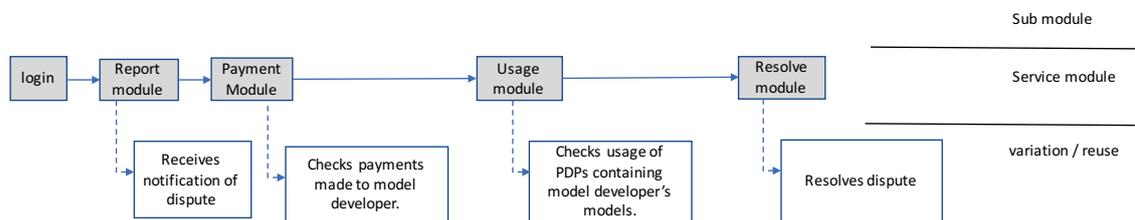


Figure 38: Platform owner receives notification of disputed payment from a model developer and resolves the dispute

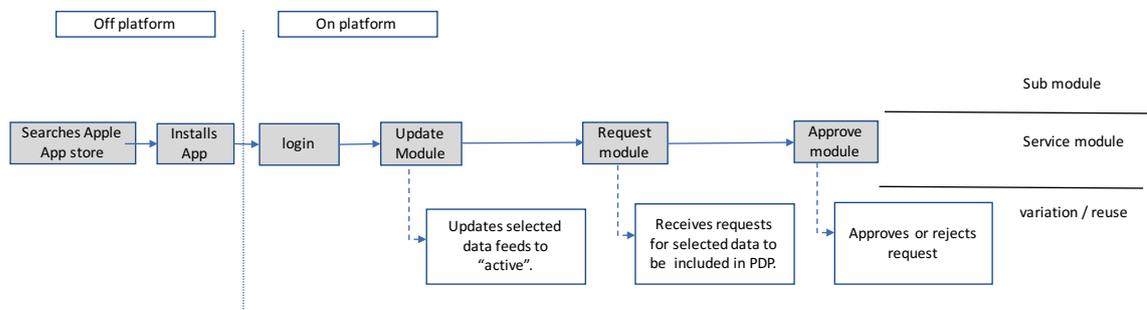


Figure 39: Data contributing individual installs data collecting app, selects data feeds to be collected and receives requested for his or her data to be include in PDPs

9.5 Interfaces

As defined above the interface between modules comprises the ‘set of rules and guidelines governing the flexible arrangement, interconnections, and interdependence of service components...’ (de Blok et al. 2014, p. 186). An interface can also be defined in terms of ‘specifications that define the inputs and outputs that cross the interfaces between interacting components’ (Voss & Hsuan 2009).

In common with many platforms (Roth & Menor 2003) the majority of service modules are delivered directly to the customer via a web-based delivery mechanism while others, such as the approval module and the resolve module, include communication with the customer through a choice of modes.

Technology facilitated interfaces are defined by Application Programming Interfaces (APIs) which set and govern the arrangement, interconnections and interdependence of the service elements in the PPIM. APIs and standardised interfaces with well-documented rules describe the inputs and outputs from each service module.

Drawing on API design principles from the literature (Blanchette 2008; Chen, Xu & Zhu 2012; Tulach 2008), the APIs required to enable the PPIM are outlined below.

9.6 Hierarchy of API

The interface between modules (both service oriented and outcome oriented), submodules, participant types and variations and reuses is governed by a hierarchical implementation of an API as shown in Figure 40.

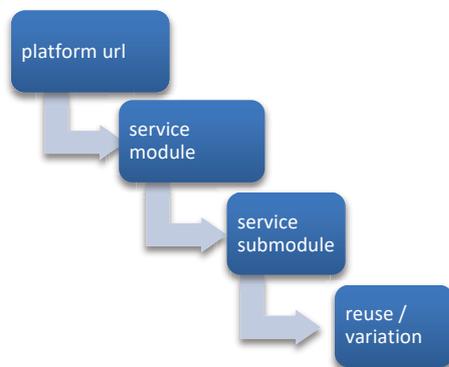


Figure 40: Hierarchy of API implementation.

This hierarchy is implemented via web-based URLs as follows:

`https://platform_url.com/<module>/<submodule>/<reuse_variation>`

In the following section, we present a number of examples from Table 20 to Table 24 with their associated API calls.

Scenario: End user searches available model components or insights in order to build a customised PDP.

API Call: `https://platform_url.com/search/model_components/<string>`

Action: Returns all model components and insights which contain the search string.

SQL call: `select * from insights where lower(insight name) like '%lower(string)%';`

Scenario: Model developer submits model for approval by platform owner.

API Call: `https://platform_url.com/approve/submit/<model_id>`

Action: Inserts a new model into the pending approval table.

SQL call: `insert model_id into models_pending_review;`

9.7 Value chain enabled by the service architecture

In the scenarios above the value of the artefact is enhanced – these diagrams present the value chain, demonstrating how artefacts of value are created as the various service modules are invoked via the API interfaces.

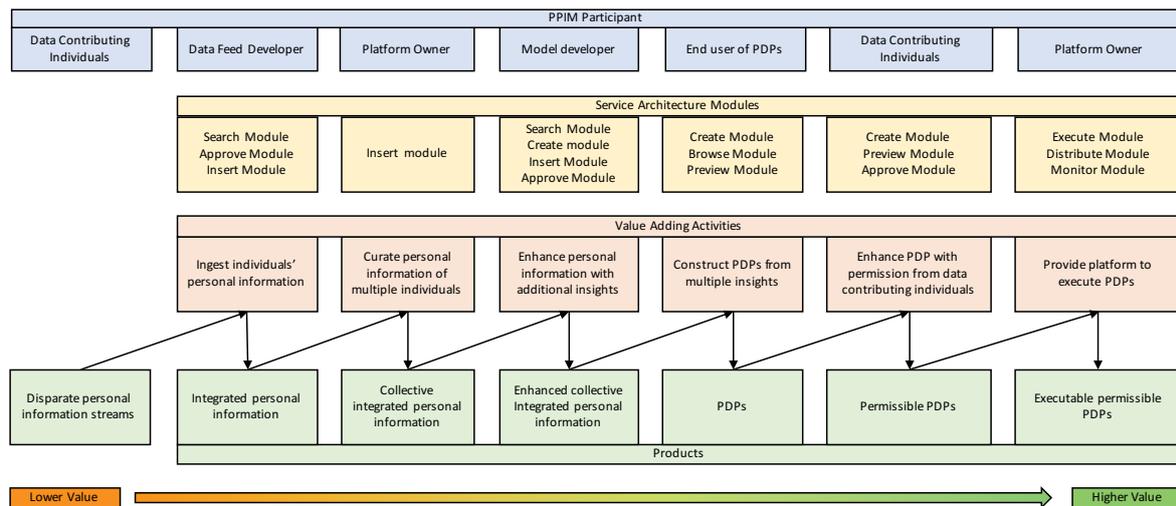


Figure 41: The value chain for the creation of a new PDP

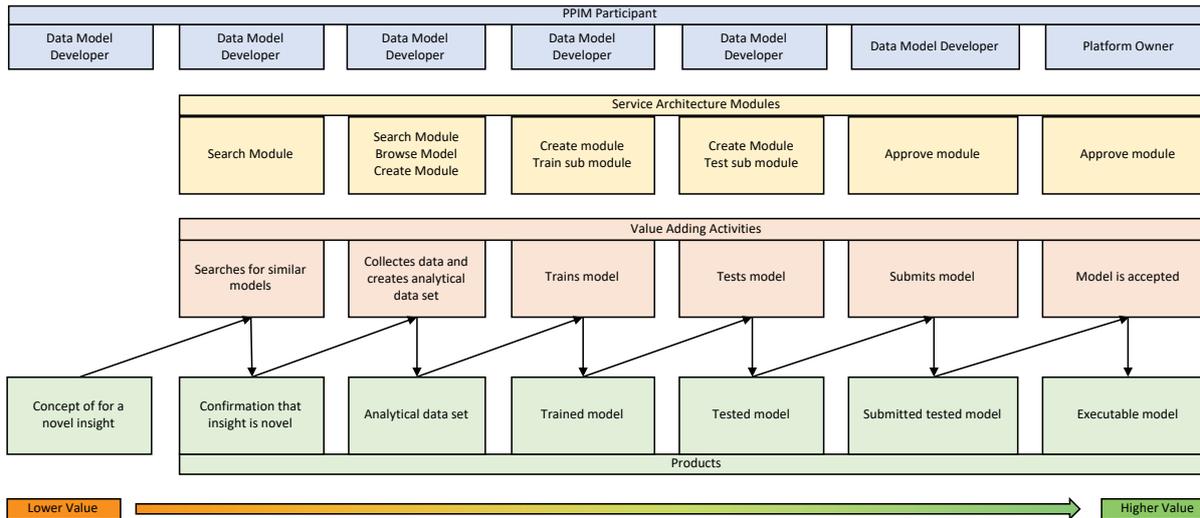


Figure 42: The value chain for the creation of a new model

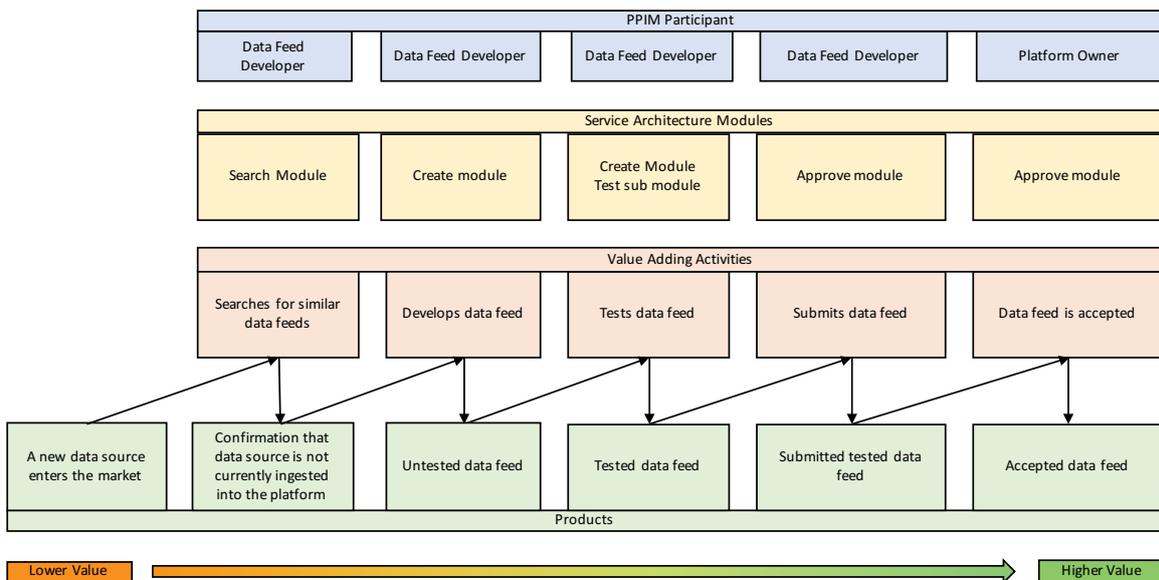


Figure 43: The value chain for the creation of a new data feed

9.8 Service architecture enabling platform evolution

Platform survival is largely contingent on its ability to create strategic disequilibrium in the market place through continuous innovation (Yoo 2010). In this section, we examine how the PPIM service architecture is aligned with the need for a PPIM to evolve in this manner and how innovations are fed into the PPIM in order to maintain its competitive advantage.

The resilience of a platform in terms of its ability to resist inroads from competitors largely depends on it having a funnel of innovation which is filled more quickly than competitors can emulate them, a concept known as the ‘Bathtub model’ (Tiwana 2013, p. 204).

Continuing this analogy, there are four streams of innovation on a PPIM which feed into the ‘bathtub’ of innovation stock.

On a PPIM, the role of App developer in Tiwana’s bathtub model is fulfilled by model developers and data feed developers. Data feed developers represent the first stream of innovation. The innovation inflow to a PPIM would be continuously supplied since, as new devices and software are developed by third party companies, they would require additional data feeds to be developed in order for those data to be captured and collected on behalf of the data contributing individuals. It is difficult to get statistics on the number of new wearables coming onto the market but, anecdotally, there were 120 wearable devices seeking funding on Kickstarter on 11th January 2018 (Kickstarter 2018).

Furthermore, there would be a continuous inflow of innovative models to extract new insights from these data. This would comprise the second stream of innovation. This inflow would be driven by advances in data mining research. For example, there are thousands of R packages available on CRAN, many of which represent new ways to extract insights from data. The number of new packages being developed is increasing rapidly (Revolutions 2018). As new R packages are developed and released to the open source community there would be an incentive for model developers to adapt these and use them to derive innovative insights from personal data on a PPIM.

A third stream of innovation comprises the templates for PDPs which would be developed by end users. The PDPs themselves, which are private to the end users who developed them, nevertheless represent innovative intellectual property which did not previously exist. End users have two avenues for creating innovating PDPs. They can create new artefacts by starting with existing templates and then customising with additional insights from models. This ability to remix and associate existing artefacts has been identified by Yoo (2010, p. 219) as one of the four essential characteristics of the digitally mediated experience. End users can alternatively create new PDPs from scratch by mixing and matching insight starting with a blank canvas.

The fourth stream of innovation is the platform owner. The platform owner will provide innovative revenue sharing models, collective decision-making models, reputational assessment tools and marketing efforts to ensure the continued evolution of the platform.

Figure 44 depicts the bathtub model as it applies to a PPIM with model developers, data feed developers and end users fulfilling the role of innovators.

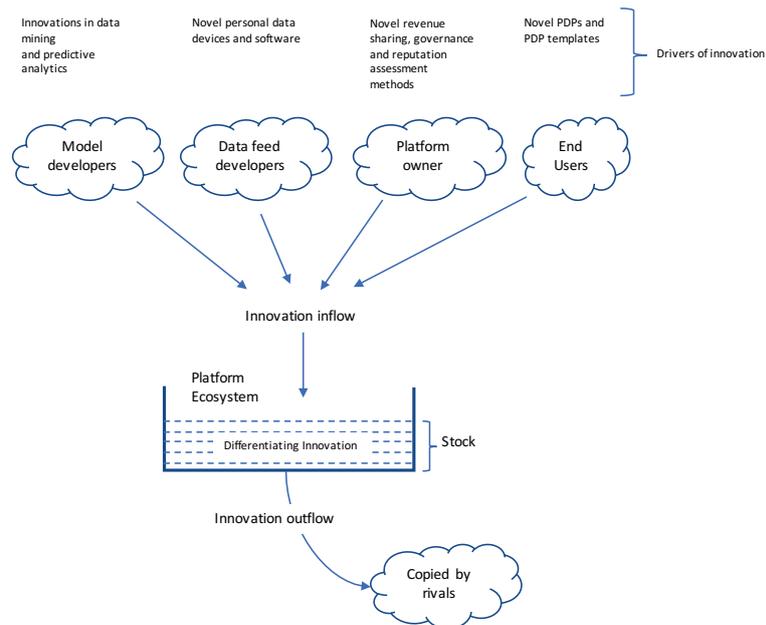


Figure 44: The Bathtub model of innovation as applied to a PPIM, adapted from (Tiwana 2013, p. 204)

The ways in which innovation can take place on a PPIM are facilitated by the ‘liquefaction’ of resources (Lusch & Nambisan 2015). This means the increased resource density brought about by the bringing together in an accessible manner the service output of the four participant groups on the PPIM. ‘Resource density’ means the phenomenon of having relevant resources (people, processes and technology) in close proximity and available to be configured quickly in order to address end user needs (Lusch, Vargo & Tanniru 2010).

Finally – the service architecture for a PPIM in Figure 45 is summarised, which shows an overarching reference architectural model diagram into which all the noted elements and their interrelationships are brought together into a single visualisation.

In Figure 46 the workflow for one scenario is indicated, that is the creation of a new model. The work flow shows how the service modules and interfaces are utilised to create a new model on a PPIM.

1. The model developer searches for models similar to the one he has in mind and examines the models returned by the search algorithm in ranked order.
2. He then browses the search results and examines them for similarity to the model he has in mind.
3. Having assured himself of the uniqueness of the model he has in mind he then uses the model development console presented in 57 to develop the new model.
4. When the model has been developed he submits it to the platform owner for inclusion on the platform.
5. The platform owner examines the model and approves it.

6. The PDP catalogue of available insight is then updated with the insights made available from the new model, allowing end users to include the new insight into their PDPs.
7. An end user includes the new insight in a PDP which is subsequently purchased.
8. The PDP is executed.
9. The model developer receives payment for the execution of a PDP which includes an insight from a model he developed.

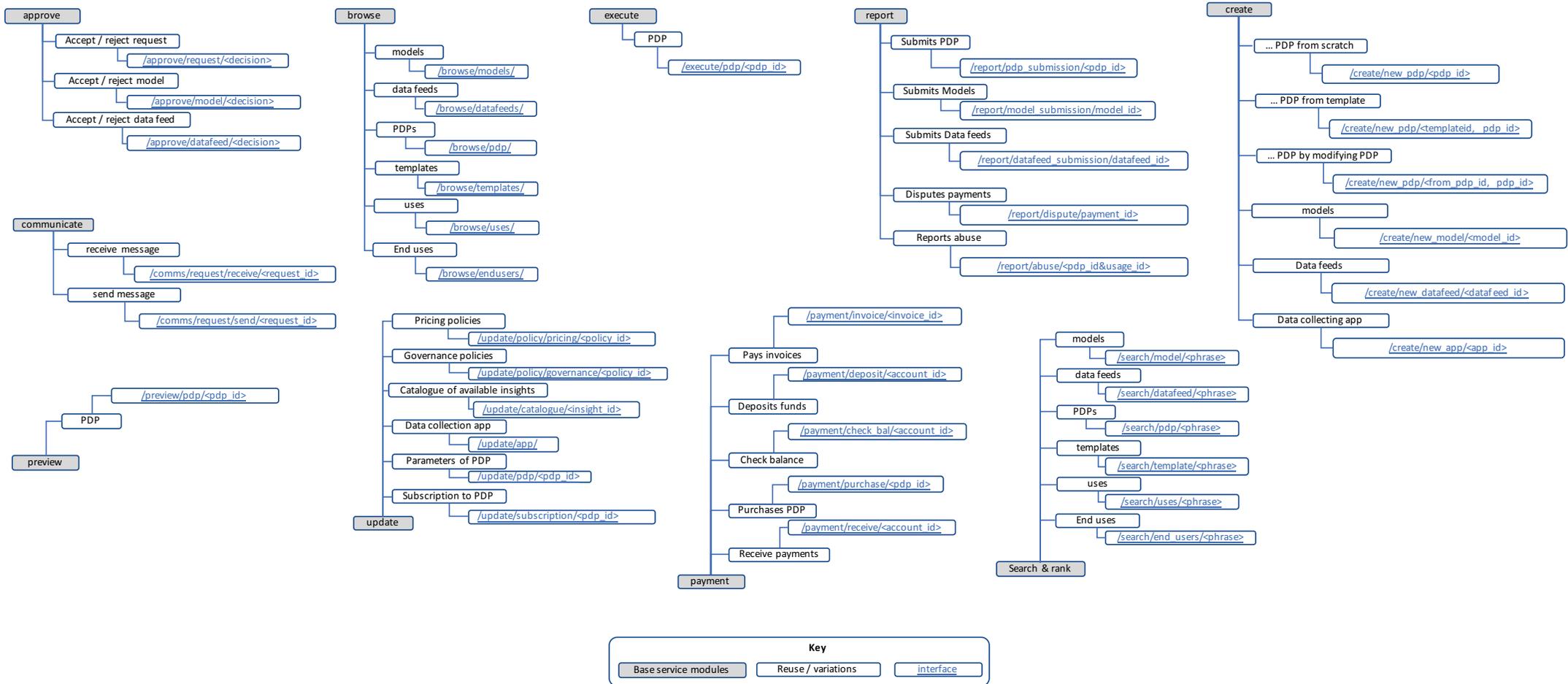


Figure 45: Service architecture for a PIM

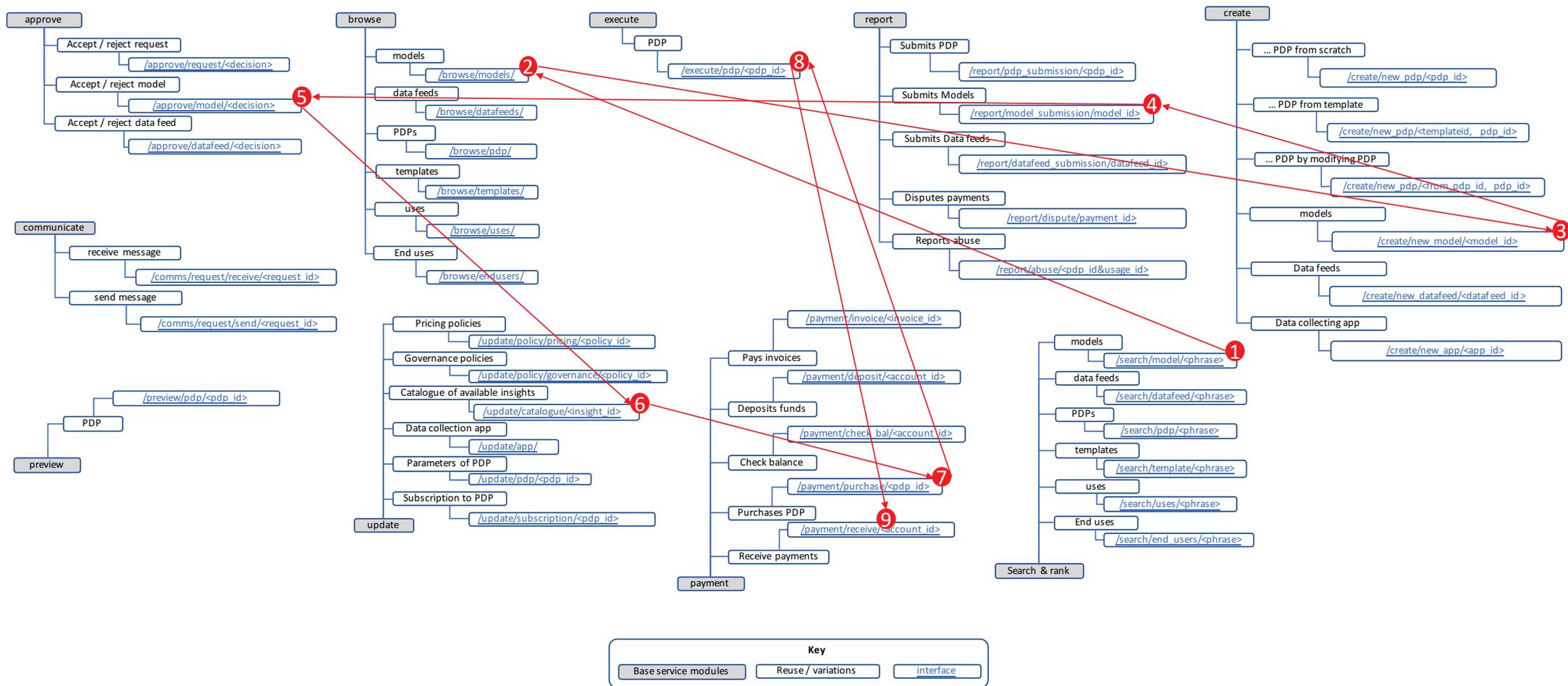


Figure 46: Create model scenario executed on a service architecture for a PPIM

10 PPIM end to end

The aim of this chapter is twofold. First, to synthesise the research presented in chapters 8 and 9 to demonstrate the end to end working of a PPIM. Second, to enhance the validation of our proposed artefact (a PPIM) with additional analysis demonstrating the alignment of our design with the integrated Service Innovation Method (*iSIM*). This multi-faceted, iterative validation approach, combining interviews, prototypes and analysis is consistent with other Design Science research. Hevner et al. (2004) lists five methods by which artefacts can be evaluated: analysis, case studies, experiments, field studies and simulation and does not argue against utilising multiple methods to evaluate an artefact in different iterations of the design process. There are examples of the employment of multiple evaluation methods in the literature. For example, McLaren et al. (2011) evaluated their artefact by both analysing a case study and assessing an example instance in a real world context, while Adomavicius et al. (2008) used both case study and interviews to evaluate their artefact. Lee, Wyner & Pentland (2008) used prototype and analysis in their evaluation.

First, the difference between a PPIM considered as a m-sided platform and a platform as usually described in the literature (Tiwana 2013; Tiwana, Konsynski & Bush 2010) is described. The end-to-end viability of the PPIM is verified by demonstrating that our design incorporates and adapts the key features necessary for a viable platform, as identified in previous research (Tiwana, Konsynski & Bush 2010).

10.1 Comparison of PPIM and standard platform

10.1.1 Roles

A platform is generally conceived as having three main roles, which are fulfilled by different groups of participants. Orchestrating the evolution of the platform is performed by the platform owner, the lead firm primarily responsible for the platform. The platform owner is referred to elsewhere as the ecosystem's keystone firm (Iansiti & Levien 2004) or economic catalyst (Evans 2009). The development of apps is fulfilled by app developers who are not employed by the platform owner. The purchasing and use of these apps is performed by end-users (Tiwana 2013, p. xvii).

On a PPIM, there is one additional role and the two of the three roles described in the previous paragraph are fulfilled by different groups of participants. The role of orchestration of platform evolution is fulfilled by the platform owner. App development is fulfilled by model developers and data feed developers. The additional role of granting permission for access to personal information is fulfilled by the data contributing individuals. The purchasing and use of apps is

fulfilled by companies that co-create personal data products (PDPs) and purchase the right to use these PDPs.

Furthermore, on a PPIM there is also the need for an app (in the traditional sense of the word) which is installed on the data contributing individual’s device to facilitate the collection and ingestion of personal information into the centralised data repository. This app is developed by the platform owner. When we refer to apps in this chapter, we are not referring to this data collection app, we are referring to the apps developed by the data model and data feed developers which are installed on and run on the platform.

The comparison between standard platform and a PPIM is summarised in Table 25.

Table 25: Comparison of roles fulfilment between standard platform and PPIM

Role	Standard platform	PPIM
orchestration of platform evolution and management of platform	platform owner	platform owner
app development	app developers	data model developers data feed developers
granting permission to access personal information		data contributing individuals
end users of apps	individuals	companies

10.1.2 Location of app execution

There is one additional important difference between a PPIM and a standard platform, the location of the execution of the apps. On a standard platform, apps are downloaded by end users and installed and executed on their devices. This necessitates careful planning of the app microarchitecture which description of how an app interacts, communicates, and interoperates with the platform (Tiwana 2013, p. 289).

On a PPIM, apps are amalgamated into PDPs where they are executed *on the platform*. This important difference impacts the way the platform owner needs to design the platform architecture and the app micro-architecture. On a PPIM, the apps are run on the central personal data store and use the resources of the platform such as CPU cycles, IO and storage to execute. Therefore, the important considerations are the efficiency of the algorithm, the amount of resources it will consume, the size of the data set it will run on, the execution time of the algorithm and the impact the app will have on other users. The platform owner therefore needs to implement a control mechanism to impose workload management, resource allocation and scheduling on the users of the platform. This control mechanism is described in section 10.4.3.1.

10.1.3 Definition of a PPIM app

10.1.3.1 Non-predictive apps

Non-predictive apps extract a feature or features such as age, gender and marital status from raw data and associates these features with an individual. An app of this kind comprises the following components:

- Metadata which describes the data feed from which the feature is extracted
- The nature of the feature

Figure 47 shows a diagrammatical representation of a non-predictive app.

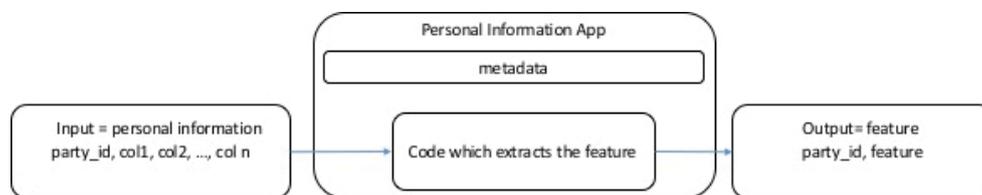


Figure 47: A non-predictive personal information app

10.1.3.2 Predictive apps

Predictive apps infer or predict a characteristic of an individual from one or more data sources. These apps have the following components.

- Metadata which describe the data feed from which the prediction is inferred and the nature of the prediction.
- Code which encapsulates the data manipulation required to build the data set on which the prediction is made.
- Code which encapsulates the algorithm which calculates the prediction.
- A metric which measures the accuracy of the prediction.
- Code which calculates the measure of accuracy.

This when we speak of an ‘app’, we refer to an entity comprising the relevant components described above. Figure 48 shows a diagrammatical representation of a predictive app.

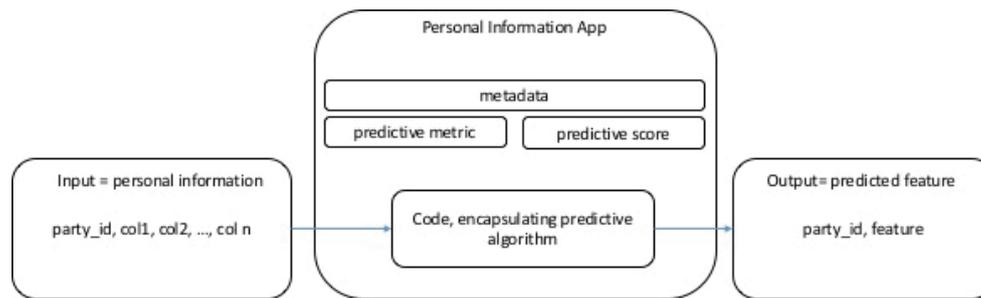


Figure 48: A predictive personal information app

10.2 Workflow to create a targeted marketing personal data product

Step 1: Data is captured in user's phone, tablet, fitness band, smart watch and PC. These data comprise Facebook, Twitter, search terms, browsing history, locations, online purchases.

Step 2: When the user's devices are online, the data is synchronised into the central data repository and stored in the user's account.

Step 3: From these data, certain features are extracted and inferred using the models that have been contributed to the PPIM by model developers. These features are: home location of users, age, gender, income, relationship status, preferences based on companies and products liked by the individuals on Facebook, current interests based on search terms and web browsing history and number of connections as inferred from number of Facebook friends, Twitter followers and LinkedIn connections. Since the personal data store contains the graph of the user's Facebook, Twitter and LinkedIn connections, certain social network metrics can be calculated. These metrics include eigen value centrality and betweenness. From the online buying history of users, certain propensity scores such as propensity to purchase and propensity to subscribe are also calculated.

Step 4: An end user, wishing to advertise to a niche market, logs onto the PPIM system and constructs a targeted marketing personal data product by specifying certain characteristics that define the subset of users they wish to target. They do so by selecting the characteristics of interest from a page similar to the one shown in Figure 24 and Figure 25 above. The end user can if so desired upload a list of existing customers and then specify that the individuals included in the target segment to be either prospects or existing customers (Figure 24). It then specifies the means by which it wishes to contact the individuals and the frequency with which it wishes to do so.

Step 5: The end user constructing the personal data products can then preview and refine the product before purchasing. He or she can iterate through the construction process by previewing

the product (Figure 26), adjusting the number of users included in the product and the price paid to execute the product by adjusting the characteristics that define it. Only individuals whose permissions coincide with the characteristics that define the product are included in the product preview and, once the product is constructed, in the product itself.

Step 6: Once the end user is satisfied with the definition of the product, the product is purchased and is added to the list of products owned by the end user (Figure 23).

Step 7: When the end user is ready to use the product, the end user opens the product, uploads the advertisement which is to be sent to the target segment and executes the product (Figure 28). Executing the product sends the uploaded advertisement to the individuals in the personal data product (PDP) via the specified contact channel. Once the product has been executed a success message is shown to the end user to indicate that the advertisement has been delivered successfully to the defined segment (Figure 29).

This workflow is represented in its entirety in Figure 49.

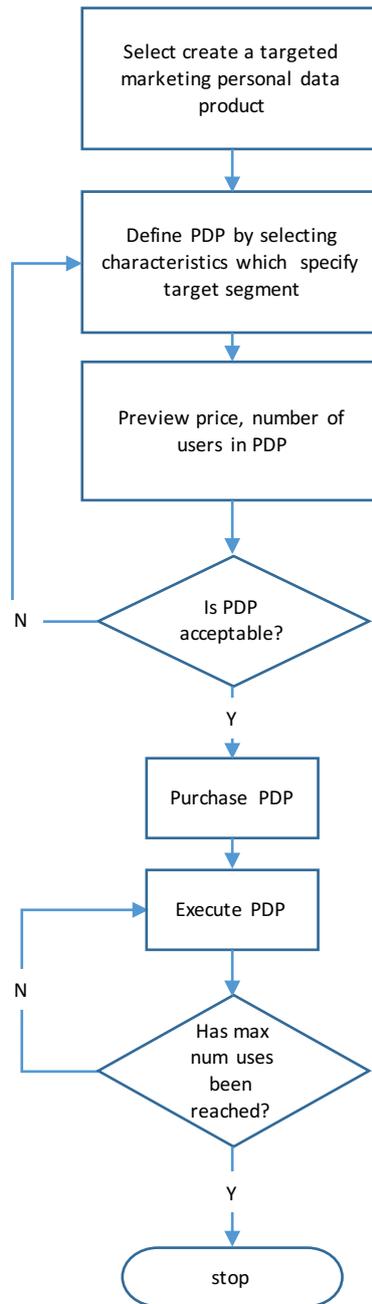


Figure 49: Workflow to construct and execute a targeted marketing personal data product

10.3 Workflow to create a personalised service response personal data product

Steps 1 – 3 are the same as in section 10.2.

Step 4: An end user wishing to know more about both existing customers and prospects at the point of interaction between end user and customer logs onto the PPIM system and constructs a

personal service response data product by specifying certain characteristics which they wish to know about individuals when they make contact with the end user (Figure 30 and Figure 31). These characteristics will vary according to the type of service or product being sold by the company. For example, a magazine subscription company may wish to know the interests and hobbies of the caller and his or her family in order to make cross sell suggestions, while this information would not be relevant to a company selling insurance.

Step 5: The end user constructing the PDPs then previews and refines the product before purchasing (Figure 26). He or she can iterate through the construction process by previewing the product and adjusting the characteristics that define it. Only individuals whose permissions coincide with the characteristics that define the product are included in the product preview and, once the product is constructed, in the product itself.

Step 6: Once the end user is satisfied with the definition of the product, the product is purchased and is added to the list of products owned by the end user's company (Figure 23).

Step 7: When a customer or prospect contacts the company, the customer service representative asks the caller for one identifying piece of personal information, such as an email address, and enters it into the system. If that customer is included in the PDP purchased by the company, the characteristics used to define the PDP are displayed to the customer service representative (Figure 34). Screen shot and copy and paste are disabled in the page to reduce the risk of unauthorised copying. This workflow is represented in its entirety in Figure 50.

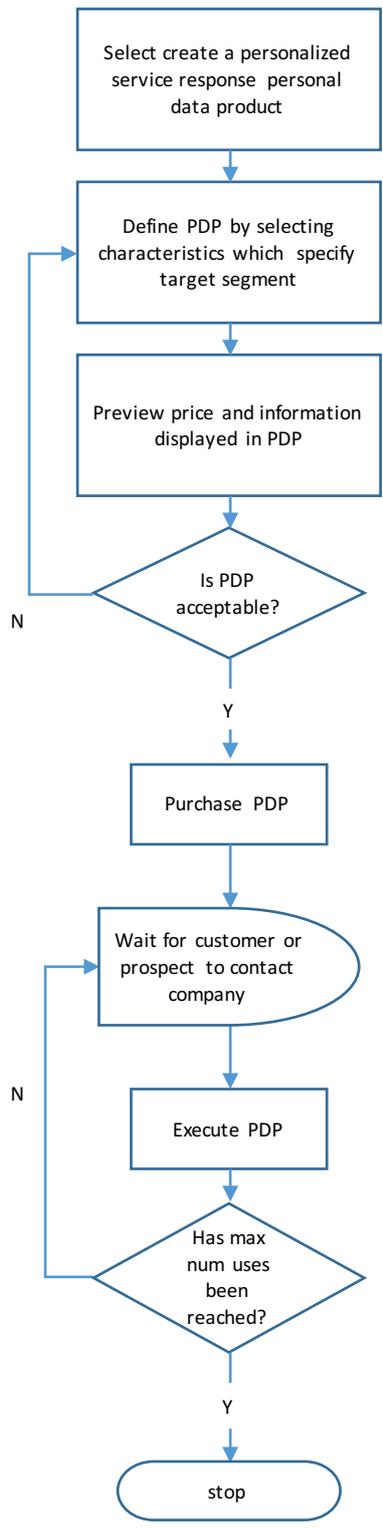


Figure 50: Workflow to construct and execute a personalised service response personal data product

10.4 Governance

Tiwana (2013) and Tiwana, Konsynski & Bush (2010) argue convincingly that good governance is essential to platform evolution. Tiwana (2013) argues that the aim of platform governance is to shape and influence the ecosystem and this is commonly achieved through three mechanisms: decision rights, voting rights and pricing policies. In this section, it is described how these three aspects of governance would apply to a PPIM.

10.4.1 Decision rights

Decision rights determine who decides what on the PPIM (Vázquez 2004). The role of the platform owner is to orchestrate platform evolution (Gawer & Cusumano 2002) as measured by market thickness (Roth 2008). Market thickness is predicated on the value proposition of all four stakeholders (or personae) being met by the platform (Muzellec, Ronteau & Lambkin 2015). Therefore, we propose that the platform owner take a democratic and transparent approach to decision rights and implement voting and feedback mechanisms to facilitate the participation of the four personae in decision making.

10.4.2 Voting rights

However, unlike a democracy, not every vote on a PPIM should carry the same weight. In accordance with the principle of aligning decision rights with specialised knowledge (Tiwana 2013, p. 113), voting should be weighted according to the contribution made by a participant and the investment they have made on the platform. This assertion is based on the observation that successful participation in the market demonstrates knowledge of the actions needed to ensure the successful operation of a PPIM. Therefore, those who have made the greatest contribution should have a greater say on decisions that will affect them. Contribution will be measured in different ways depending on the persona. A data contributing individual's contribution will be measured on the quantity, quality and longevity of the data they have contributed, and on the number of PDPs in which their data is included. End users of PDPs will have their contribution measured according to the number and value of the PDPs they have co-created, purchased and executed. Model developers and data feed developers will have their contribution measured by the number and popularity of the apps they have developed. Voting rights will be granted in proportion to the contribution of the participant.

10.4.3 Control mechanisms

10.4.3.1 Resource control

Clearly the resources of the PPIM need to be distributed fairly between competing end users, app developers and data feed developers either through differential pricing or other control

mechanisms (Tiwana 2013, p. 148) These resources include storage and compute power (as measured by CPU and IO). Resource allocation will be controlled by the use of virtual machine allocation. App developers will be allocated an initial portion of the available resources when they first sign up to the platform. The limitation on the virtual machine instance will automatically ensure that they do not exceed their allocated resources.

10.4.4 Quality control

The quality of personal information fed into and stored in the PDS is of paramount importance. Quality of traded goods is a factor related to platform reputation (Wiegand et al. 2015). If an end user creates a PDP in the belief that it comprises the data of a certain number of distinct, real individuals, it needs to be sure this is in fact that case. Trust in the quality of the transaction object being traded on an online platform can be undermined by replicated data or machine generated fake data (Norcie, De Cristofaro & Bellotti 2013). Poor quality metadata can add considerably to search cost and transaction cost.

To mitigate the risk of simulated and duplicated data, the PPIM owner will implement algorithms which can distinguish between real and simulated personal information data feeds and algorithms which can identify replicated data feeds which would be similar but not identical to those discussed by Kontaxis et al. (2011).

As a means to crowd source the curation of the large volume of user-generated content which will be collected on the PPIM, end users will have the ability to flag suspected fake or replicated data (Crawford & Gillespie 2016). These flags or alerts will assist the platform owner to maintain the quality of the data ingesting into the platform. Contributors who are found to be contributing fake or replicated data into the platform will be sanctioned or excluded from the platform.

There will also be a feedback mechanism by which the quality of the metadata will be rated. Since accurate metadata describe the content and utility of the personal information models, accurate metadata correspond to high rankings in response to searches. Users will be invited to rate models and data feeds which appear in search results in response to their search criteria. Data models and data feeds that garner high ratings will be judged to have more accurate metadata, while those with low or missing ratings will be judged to have sub-optimal metadata and will be encouraged to revise their metadata accordingly. Similar ranking and rewards systems have been shown to be highly effective on other platforms (Vasilescu et al. 2014). The platform owner would also implement a metadata collection workflow which developers would need to follow when submitting data feeds and models for inclusion on the platform. Platform owner reviewers would have the final say on whether or not the metadata is sufficiently accurate.

10.4.5 Governance examples

In this section, we describe a number of rules which may be used to govern the platform and how these rules may be implemented, policed and enforced with reference to the governance structures outlined above.

10.4.5.1 No duplicate apps

Platform participants are free to propose rules which they believe would improve the platform. Suppose a participant proposes a rule that there should be no duplicate apps on the platform. She argues that if an app that extracts a given insight already exists in the platform, then a duplicate or near duplicate should not be allowed on the platform, thereby reducing both search cost for consumers and workload on the platform. She posts a proposed rule change on the platform voting portal and participants discuss the proposed change. After a set period, a vote is taken according to the voting rights outlined in section 10.4.2. If the rule is adopted, the platform owner rewrites the app acceptance policy and informs developers of the change. The platform owner then rejects proposals for apps which are duplicate or near duplicate of existing apps, and implements mechanics whereby end users and developers can flag duplicate apps. In the interest of transparency and efficiency, the platform owner publishes rejected proposals, thereby allowing developers to understand how the policy works in practice so as to not waste time developing proposal for apps which have a high probability of being rejected.

10.4.5.2 No redundant apps

In a similar way, a rule that there should not be any redundant apps on the platform could also be voted on and implemented into the governance of the platform. If such a rule was implemented the control mechanism would be a report which lists the number of times each of the apps is included in PDPs each month. Apps which fall below a predetermined threshold for a certain number of consecutive months would be removed from the platform.

10.4.5.3 Predictive apps ranking

A third governance rule could be that predictive apps appear in search results in order of their predictive power. That is to say, apps with higher predictive power appear higher in the search results. The control mechanism for this would be a champion-challenger model comparison (Nath 2007) in which apps that aim to predict the same outcome must compete with each other on the same training and data sets. This competition would be managed by the platform owner and would form a natural hierarchy of apps. Apps would then appear in search results according to their rankings as determined by the champion-challenger comparisons.

10.4.6 Pricing

10.4.6.1 Pricing options

One key aspect of governance of the PPIM is pricing (Tiwana 2013, p. 127). There are multiple decisions that need to be made about how various participants on the PPIM will be compensated. Options for pricing include symmetric or asymmetric pricing for the two sides of the platform. Symmetric pricing occurs when the platform owner seeks to gain revenue from all sides of the market, while asymmetric pricing occurs when one or more sides of the market are allowed to participate on the platform free of charge while revenue is sought from the remaining sides. If pricing is to be asymmetric, a decision need to be made about how long the asymmetry is maintained.

The question of pricing is intimately related to the adoption and expansion strategies employed by the platform owner to grow the PPIM (Hagiu 2009). For example if a sequential entry strategy (Evans 2009, p. 15) is employed, the pricing model for the PPIM will be very different to that used if a simultaneous entry of sides strategy is to be employed (Evans 2009, p. 16).

Other decisions that need to be made are whether participants will be charged for access to the PPIM or if they will be charged for usage of the PPIM. Elsewhere (Farrelly & Chew 2017), a usage pricing model is proposed.

Decisions also need to be made about whether revenue splitting will be conducted using a fixed scale or a sliding scale pricing model. In a fixed scale revenue splitting model, the platform owner retains a fixed percentage of all revenue generated on the platform. For example, Apple, as the platform owner, retains 30% of all revenue generated on their app platform (Tiwana 2013, p. 129). With a sliding scale, the percentage retained by the platform owner changes according to the popularity of the app.

10.4.6.2 Optimising pricing decisions

The role of the platform owner in setting pricing policies is to maximise market thickness. That is to say, when setting pricing policies, the platform owner should select policies which will maximise the participation of all sides of the market, as measured by number of participants, numbers of apps submitted, and number of apps purchased and executed on the platform. This activity can be monitored in terms of the revenue generated on the platform.

Setting pricing policies to optimise revenue on a complex system which is yet to be built is a non-trivial problem and can have a major impact on the performance of the platform. Furthermore (Tiwana 2013, p. 199) argues persuasively that the ability for a platform to evolve in response to changing market conditions is crucial to its survival. Therefore, we propose that rather than set and forget the pricing structure of the platform, the platform owner must establish

a mechanism which continually conducts trials on alternative permutations of pricing alternatives and measure the effects they have on market thickness (Goh & Bockstedt 2013).

However, maintaining a stable and consistent interface, not just in terms of an API but also in terms of governance polices, is also a key determinant of platform success. Therefore, we propose that the platform owner publish and publicise their policy of continuous evolution of pricing models, so participants known in advance and understand the rationale behind changes in pricing.

10.5 Congruence between *i*SIM and end-to-end validation of PPIM

This section, explains how the service architecture, described in 9.1, is aligned with and enables the *i*SIM conceptualisation of a PPIM outlined in chapter 8. It represents the service business strategy, the customer type and value proposition and the service concept design from chapter 8 and inserts the relevant API calls from the ecosystem architecture to demonstrate the alignment.

10.5.1 Service business strategy

As detailed in section 8.4.1, the service business strategy is an instantaneous ‘build-to-order’ logic associated with a virtual marketplace. Building on the outline in section 8.4.1 the service business strategy is presented together with API calls to demonstrate that the ecosystem architecture is aligned with the service business strategy.

The raw materials (personal data) are ingested via data feed connections which have been developed by the data feed developers and added to the platform.

post	submit	<a href="https://platform_url.com/report/data_feed/<reuse_variation>/account/feeds/<feed_id>/submit">https://platform_url.com/report/data_feed/<reuse_variation>/account/feeds/<feed_id>/submit	submits a data feed to the platform owner for approval
post	create	<a href="/account/feeds/<feed_id>/create">/account/feeds/<feed_id>/create	creates a data feed
post	approve	<a href="/account/feeds/<feed_id>/approve">/account/feeds/<feed_id>/approve	approves a data feed
post	add	<a href="/data_feed/add/<data_feed_id>">/data_feed/add/<data_feed_id>	adds a data feed to a user’s account

The output from models that have extracted inferred information from the raw data are also present in the centralised data store.

get	model_search	/platform/models/<string>	given a string, returns models which contain that string
-----	--------------	---------------------------	--

The end user co-creates his or her customised PDP with a PPIM by selecting an appropriate PPIM-provisioned template and populating it with the requisite components to customise the PDP.

post	create_preview	/pdp/create_preview/<pdp_id>	creates a preview of a PDP
post	delete_preview	/pdp/delete_preview/<pdp_id>	deletes a preview of a PDP
post	add_feature_to_preview	/pdp/add_feature_to_preview/<pdp_id>	adds a feature to a preview of a PDP
post	delete_feature_from_preview	/pdp/delete_feature_from_preview/<pdp_id>	deletes a feature from a preview of a PDP
get	preview	/pdp/preview/<pdp_id>	returns a preview of a PDP
post	send	/pdp/send/<pdp_id>	sends request to data contributing individual, asking permission to include their data in an PDP

After previewing the PDP, and sending and receiving permission to include user's data in the PDP, the end user creates the PDP which is instantaneously built-to-order at the completion of this process.

post	create	/pdp/create/<pdp_id>	creates a PDP
------	--------	----------------------	---------------

When the PDP is executed, the platform facilitates the revenue distribution from the purchaser to the contributors to the PDP.

post	credit_funds	/platform/credit_funds/<pdp_id>	credit funds to data contributing individuals, model developers and data feed developers when a PDP is executed
------	--------------	---------------------------------	---

10.5.2 Customer type and customer value proposition

Data contributors have the role of granting permission for their data to be collected on their behalf, stored in the centralised data repository and for granting permission for their data to be included in the personal data products (PDPs) to which they give their consent.

post	accept	/request/accept/<request_id>	accepts a request to use personal information
------	--------	------------------------------	---

The value proposition to data contributors is the full control of their personal data and the financial compensation they receive in exchange for limited and transparent use of their personal data on which they have granted permission.

get	requests	/account/requests/<username>	returns the requests that have been accepted by a user
get	usage	/account/usage/<username>	Returns the uses to which personal information associated with an account is being put

Model developers participate on a PPIM by co-producing models with the data contributing individuals.

post	submit	/account/models/<model_id>/submit	submits a model to the platform owner for approval
post	create	/account/models/<model_id>/submit	creates a model

These models extract insights of value from the raw personal data. A model may be as simple as extracting an individual's age from his or her Facebook data or as advanced as predicting an

individual's preferred holiday destination using clustering techniques. PDPs are customised by selecting the output from one or more models together with an associated permitted usage and then are instantaneously built-to-order at the completion of the customisation process.

post	add_feature_to _preview	/pdp/add_feature_to_preview/ <pdp_id>	adds a feature to a preview of a PDP
post	delete_feature_from _preview	/pdp/delete_feature_from_preview/ <pdp_id>	deletes a feature from a preview of a PDP
get	preview	/pdp/preview/<pdp_id>	returns a preview of a PDP
post	send	/pdp/send/<pdp_id>	sends request to data contributing individual, asking permission to include their data in an PDP
post	create	/pdp/create/<pdp_id>	creates a PDP

The value proposition to model developers is the access right to the permitted personal data repository and the financial compensation they receive whenever a personal data product that comprises the results of one or more of their models is executed.

post	credit_funds	/platform/credit_funds/<pdp_id>	credit funds to data contributing individuals, model developers and data feed developers when a PDP is executed
------	--------------	---------------------------------	---

Data feed developers participate on a PPIM by developing feeds from new devices or services that generate or capture personal data into the centralised personalised data repository.

post	submit	/account/feeds/<feed_id>/submit	submits a data feed to the platform owner for approval
------	--------	---------------------------------	--

post	create	/account/feeds/<feed_id>/create	creates a data feed
------	--------	---------------------------------	---------------------

The value proposition to data feed developers is the financial compensation they receive whenever a personal data product that comprises the raw data collected via one or more of their feeds is executed.

10.5.3 Service concept design

Here the design of the service logic of a PPIM is outlined to fulfil the business logic and customer value propositions described above – from the perspective of four actor types: data contributing individuals, data feed developers, model developers and end users.

Data contributing individuals receive access to a centralised personal data store which captures (via a data feed / device interface) and securely stores their data. They receive requests for access to their personal data which, if accepted, results in financial compensation.

post	credit_funds	/platform/credit_funds/<pdp_id>	credit funds to data contributing individuals, model developers and data feed developers when a PDP is executed
post	accept	/request/accept/<request_id>	accepts a request to use personal information
post	reject	/request/reject/<request_id>	rejects a request to use personal information

Individuals receive access to a personal information dashboard which enables them to manage and control access to their personal data.

get	usage	/account/usage/<username>	Returns the uses to which personal information associated with an account is being put
-----	-------	---------------------------	--

Data feed and model developers benefit from participating on a PPIM by gaining access to a market to promote and sell the fruit of their labours. The data feeds and models they develop

are made available on the PPIM and are available to be incorporated into PDPs at build time on demand.

post	add_feature_to_preview	/pdp/add_feature_to_preview/<pdp_id>	adds a feature to a preview of a PDP
post	delete_feature_from_preview	/pdp/delete_feature_from_preview/<pdp_id>	deletes a feature from a preview of a PDP
post	create	/pdp/create/<pdp_id>	creates a PDP

When a PDP comprising a developer’s data feed or model is executed the developer receives financial compensation.

post	credit_funds	/platform/credit_funds/<pdp_id>	credit funds to data contributing individuals, model developers and data feed developers when a PDP is executed
------	--------------	---------------------------------	---

Developers also receive the ability to transparently track the use to which their components are being put and the financial compensation they are receiving.

get	models	/account/models/<username>	Returns the models associated with an account
get	usage	/account/models/<model_id>/usage	returns the usage statistics for a model

End users benefit by participating on a PPIM by gaining the ability to co-design, co-build, then purchase and execute PDPs with the PPIM. End users receive the ability to browse personal data assets such as templates and information which has been inferred and extracted from the raw materials of personal data by model developers.

get	model_search	/platform/models/<string>	given a string, returns models which contain that string
get	pdp_search	/platform/pdp/<string>	given a string, returns PDPs which contain that string

Via the end user facing PDP dashboard, they receive a means to transparently co-create, preview, purchase, execute, delete, renew and monitor the PDPs.

post	create_preview	/pdp/create_preview/<pdp_id>	creates a preview of a PDP
post	delete_preview	/pdp/delete_preview/<pdp_id>	deletes a preview of a PDP
post	add_feature_to_preview	/pdp/add_feature_to_preview/<pdp_id>	adds a feature to a preview of a PDP
post	delete_feature_from_preview	/pdp/delete_feature_from_preview/<pdp_id>	deletes a feature from a preview of a PDP
get	preview	/pdp/preview/<pdp_id>	returns a preview of a PDP
post	send	/pdp/send/<pdp_id>	sends request to data contributing individual, asking permission to include their data in an PDP
post	create	/pdp/create/<pdp_id>	creates a PDP
post	execute	/pdp/execute/<pdp_id>	executes a PDP
post	renew	/pdp/renew/<pdp_id>	renews a PDP

End users also receive the ability to validate the quality of the data to which they are buying access and the validity of the models they are incorporating into their products via a fair and transparent rating systems whereby the market can give feedback on the validity and utility of the models.

get	get_rating	/account/models/<model_id>/assign_rating	returns the rating of a model
-----	------------	--	-------------------------------

get	goodness of fit	/account/models/<model_id>/ goodness_of_fit	returns the goodness of fit metric used to assess a model's accuracy
-----	-----------------	--	--

On a PPIM, PDPs can be co-created from four types of components: templates, data feeds, extracted information and permissions. These templates are previews of a PDP created by the platform owner.

post	create_preview	/pdp/create_preview/<pdp_id>	creates a preview of a PDP
post	delete_preview	/pdp/delete_preview/<pdp_id>	deletes a preview of a PDP
post	add_feature_to _preview	/pdp/add_feature_to_preview/ <pdp_id>	adds a feature to a preview of a PDP
post	delete_feature_from _preview	/pdp/delete_feature_from_preview/ <pdp_id>	deletes a feature from a preview of a PDP

The end user can utilise a template for a given type of PDP (targeted marketing, personal response, customer identification, etc.). The template is then populated with the data feeds (made possible by the data feed developers) and the information extracted from the data feeds (by the model developers).

post	add_feature_to _preview	/pdp/add_feature_to _preview/<pdp_id>	adds a feature to a preview of a PDP
post	delete_feature_from _preview	/pdp/delete_feature_from _preview/<pdp_id>	deletes a feature from a preview of a PDP

Finally, permission to execute a PDP, in terms of time duration, frequency of use and mode of contact is incorporated into the PDP.

post	create	/pdp/create/<pdp_id>	creates a PDP
------	--------	----------------------	---------------

10.6 End-to-end fitness

In this section, the end-to-end fitness of the PPIM is assessed. To do so the *iSIM* interrelationship map is utilised which has been developed as a tool to assess it (Figure 51) (Chew 2015).

	strategy	Service architecture	Monetization	Customer type & CVP	Service concept	Service system	Experience
strategy		Industry position; businesslogic; platform choice	Platform choice; m-sided revenues per business logic	Business logic & CVPper customer type per platform	Business logic & CVPper customer type per platform	Business logic & CVPper customer type per platform	Business logic & CVPper customer type
Service architecture	E2E fitnessof modular business logic per industry & platform layer		M-sided revenue model per modular arch. layers & external fit	Modular cust. Interface; componentize d CVP & external fit	Modular concept design with CVP at cust. interface& E2E fit	Modular resource & systemmodel delivering CVP & E2E fit	Experience model delivers CVP;customer learning syst. & external fit
Monetization	Monetization intensity dynamics	Monetization intensity dynamics		Monetization intensity dynamics	Monetization intensity dynamics	Monetization intensity dynamics	Monetization intensity dynamics
Customer type & CVP	Alignmentto CVP; external fitness criteria	Alignmentto CVP; external fitness criteria	Alignmentto CVP; external fitness criteria		Alignmentto CVP; external fitness criteria	Alignmentto CVP; external fitness criteria	Alignmentto CVP; external fitness criteria
Service concept	Strategic adjustment requirements	Architectural adjustment requirements	Valuecreation dynamism requirements	Alignment of internal to external fitness		Service system E2Efitness design requirements	Alignment of internal to external fitness
Service system	Strategic adjustment requirements	Architectural adjustment requirements	Systemic value creation dynamism requirements	Alignment of internal to external fitness	Alignment of internal to external fitness		Alignment of internal to external fitness
Experience	Customer learning re external fitness	Customer learning re external fitness	Customer learning re external fitness	Customer learning re external fitness	Customer learning re external fitness	Customer learning re external fitness	

Figure 51: *iSIM* interrelationship map (Chew 2015)

10.7 Evaluation of the end-to-end conception of the PPIM as a four-sided market

An evaluation workshop was held to evaluate the end-to-end conception of the PPIM.

10.7.1 Workshop design and methodology

In accordance with design science, the PPIM digital platform was evaluated qualitatively (Yin 2013) using a workshop and a survey. The workshop was attended by reflective practitioners (Gummesson 2002) all of whom worked in data related roles. They were selected on the basis

of their subject matter expertise, both data related and business related. This is consistent with other studies; see for example (Karunakaran & Puro 2012; Knol, Sol & Van Wamelen 2012; Nenonen & Storbacka 2010).

In addition to the prototype of the end user facing console presented in section 8.5, wireframes were also developed and presented showing the interface which would be presented to data contributing individuals, model developers and data feed developers. These interfaces are presented in Figure 52 through Figure 82.

For the sake of the workshop we have given the name ‘Datamilk’ to the company which fulfils the role of platform owner.

10.7.1.1 Model Developer Portal

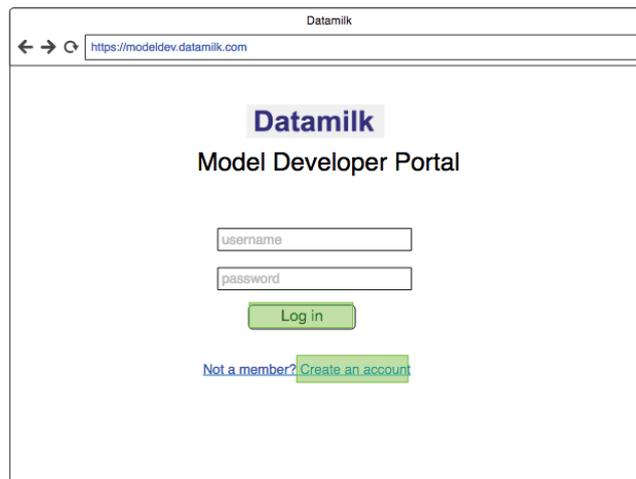


Figure 52: Model Developer Portal – login screen

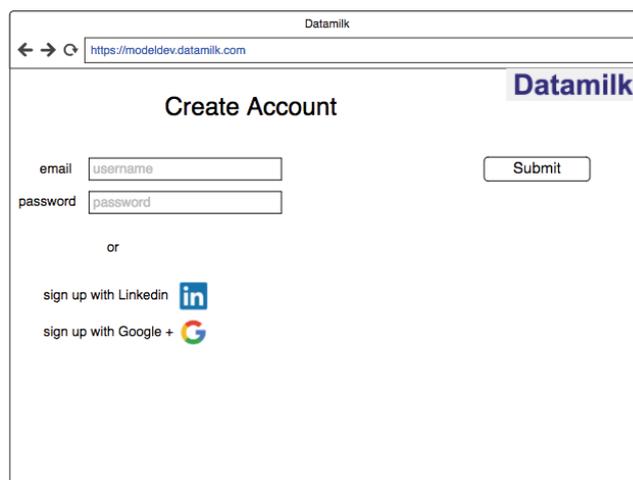


Figure 53: Model Developer Portal – create account screen

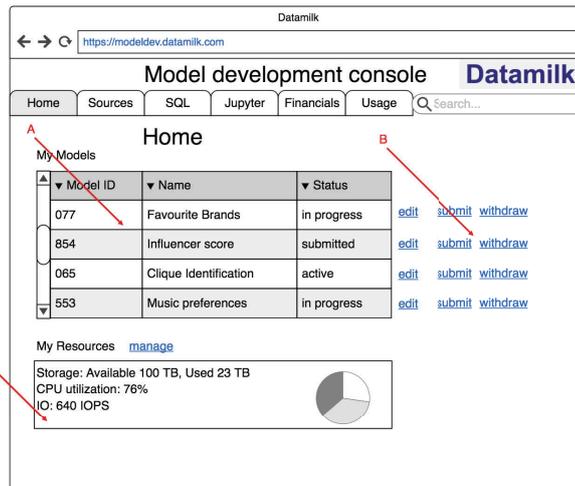


Figure 54: Model Developer Portal – home screen

On the home screen (Figure 54) the model developer can manage the models he or she has developed. The model developer can see the status of each model (in progress, active submitted, pending, etc.) (section A) and he or she can edit, submit or withdraw each of the models (section B). The model developer can also monitor the compute, storage, IO and network resources available (section C).

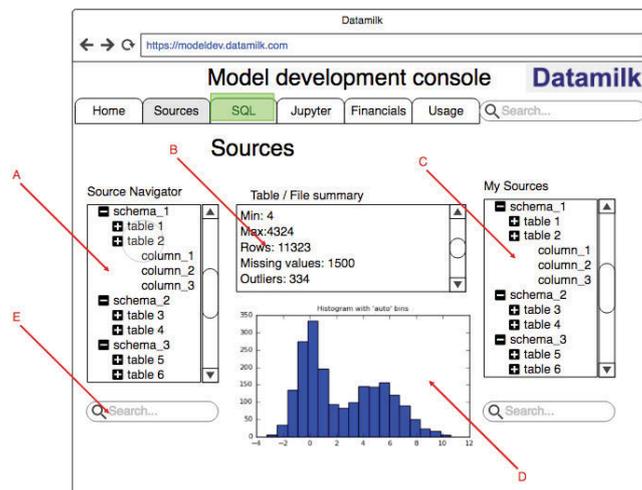
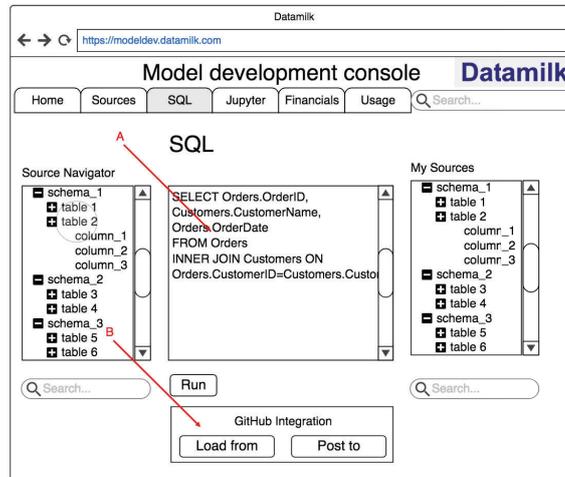


Figure 55: Model Developer Portal – sources screen

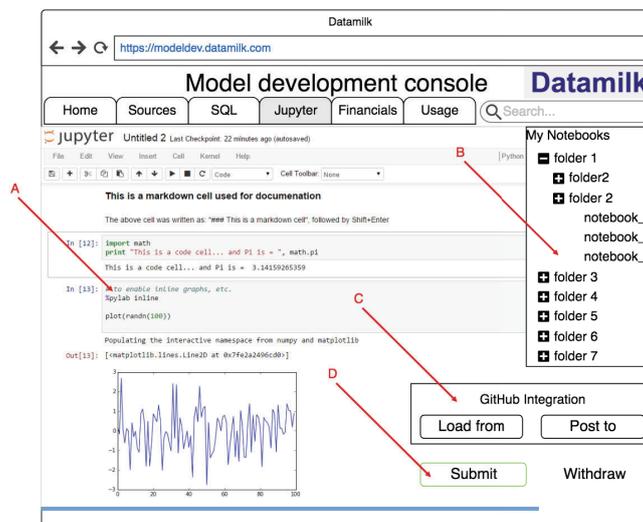
On the sources screen (Figure 55) the model developer can navigate the data sources in the centralised data store in section A. When a schema, table, view or column is selected a summary is displayed in sections B and D. The developer can also navigate schemas, tables and view

which reside in his or her own private partition of the centralised data store using section C. The developer can search for schemas, tables, views, columns and values using section E.



56: Model Developer Portal – SQL screen

On the SQL screen (56) the model developer can write and execute SQL statements to create new derived tables and views (section A). He or she can also integrate this code with a version control repository such as GitHub (section B).



57: Model Developer Portal – Jupyter screen

On the Jupyter screen (57) the model developer can develop predictive model using Jupyter notebooks (section A). The models can be written in the developer’s language of choice, executing again the tables in the centralised data repository. The code base can be integrated with a version control repository such as GitHub (section C). The models can be submitted to

the platform owner for inclusion in the PPIM or withdrawn from consideration if needed (section D).

▼ date	▼ debit	▼ credit	▼ balance
2017-05-07	\$ 0.80		\$ 15.90
2017-05-08	\$ 1.70		\$ 17.60
2017-05-09	\$ 1.10		\$ 18.70
2017-05-10	\$ 3.40		\$ 22.10
2017-05-11	\$ 1.70		\$ 23.80
2017-05-12		\$ 13.00	\$ 10.80
2017-05-13	\$ 1.60		\$ 12.40

[Withdraw Funds](#)

58: Model Developer Portal – financials screen

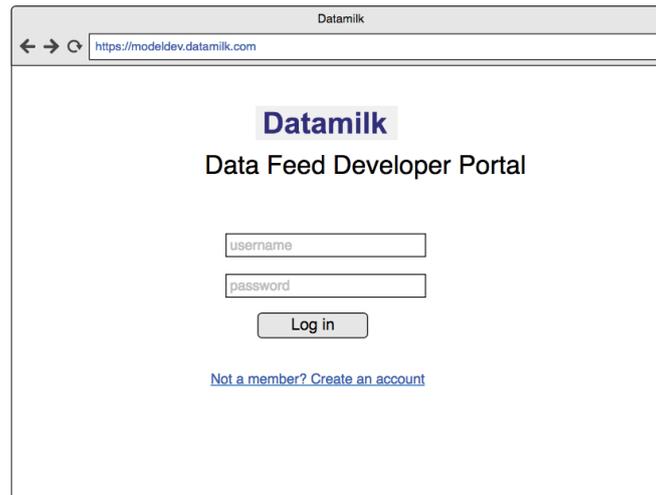
On the financials screen (58) the model developers can check their balance and monitor payments to and from their accounts. They can also transfer funds from their account to their external financial institutions.

▼ Model ID	▼ Name	▼ Status	▼ last used	▼ usage	▼ revenue	▼ details...
077	Favourite Brands	in progress	N/A	N/A	\$0.00	details...
854	Influencer score	pending	N/A	N/A	\$0.00	details...
065	Clique Identification	active	23/05/2017	3324	\$435.00	details...
553	Music preferences	in progress	N/A	N/A	\$0.00	details...

59: Model Developer Portal – usage screen

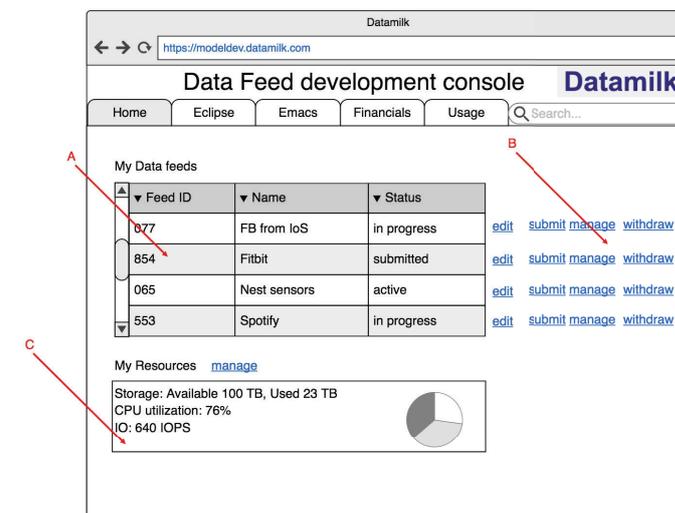
On the usage screen (59) the model developers can check the usage of the models they have developed and the revenue which has been derived from each model.

10.7.1.2 Data Feed Developer Portal

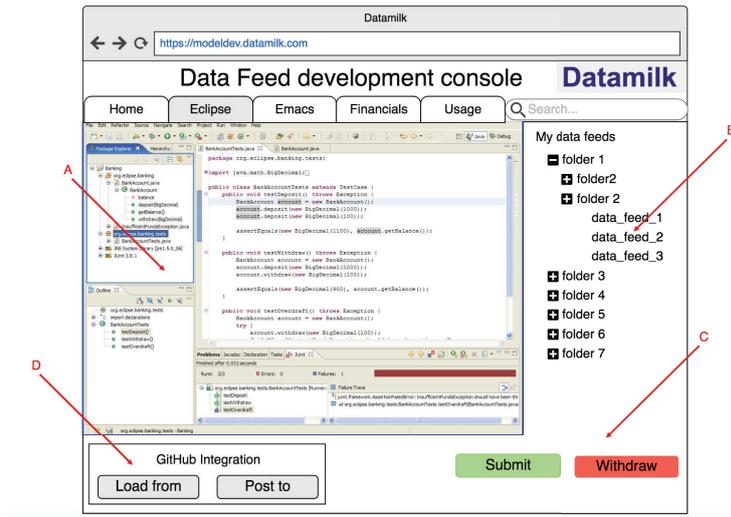


60: Data Feed Developer Portal – login screen

Data feed developers can log into the platform (60), and manage their feeds from the home screen (61). On this screen, they can view the feeds they are working on (section A), edit, submit, manage and withdraw their feeds (section B) and monitor the resources which have been allotted to them for the development and testing of data feeds (section C).



61: Data Feed Developer Portal – home screen



62: Data Feed Developer Portal – eclipse screen

On the eclipse screen (62), data feed developers can use Eclipse to develop feeds in their language of choice (section A). They can browse and edit previously developed data feeds (section B) and submit data feeds to the platform owner for inclusion on the PPIM (section C). The code they develop can be integrated with the code version repository GitHub (section D).

▼ date	▼ debit	▼ credit	▼ balance
2017-05-07	\$ 0.80		\$ 15.90
2017-05-08	\$ 1.70		\$ 17.60
2017-05-09	\$ 1.10		\$ 18.70
2017-05-10	\$ 3.40		\$ 22.10
2017-05-11	\$ 1.70		\$ 23.80
2017-05-12		\$ 13.00	\$ 10.80
2017-05-13	\$ 1.60		\$ 12.40

63: Data Feed Developer Portal – financials screen

On the financials screen (63), data feed developers can have visibility of the revenue they are accruing from the use of their feeds and withdraw funds from the platform into their external financial institutions.

▼ date submitted	▼ name	▼ usage	▼ description
2017-05-07	iPhone location	1322	... A
2016-07-08	Android accelerometer	334	...
2016-05-09	Nest energy usage	221	...
2017-04-10	Fit bit - all data	3324	...
2015-12-11	Gmail	6657	...
2016-03-12	Firefox browsing history	21	...
2015-11-13	Safari search history	78	...

64: Data Feed Developer Portal – usage screen

On the usage screen (64), data feed developers can have visibility of the number of times their data feeds have been used and, by clicking through on the details link (Section A), they can see the PDPs in which their feeds have been included.

10.7.1.3 Data Contributing Individual mobile interface

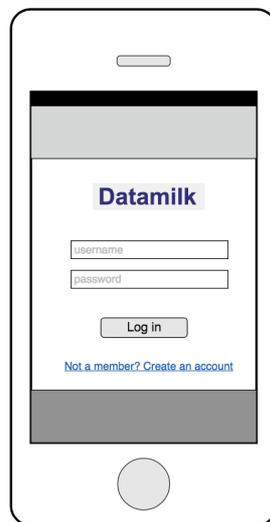


Figure 65: Data Contributing Individual – login screen

Data contributing individuals log into the PPIM via a smartphone (Figure 65).

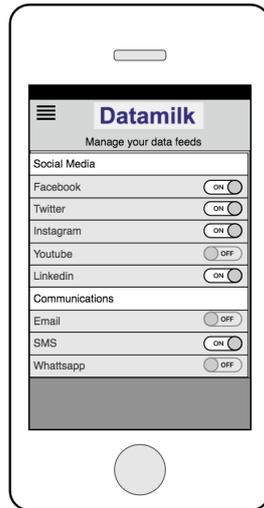


Figure 66: Data Contributing Individual – manage data feeds screen

Data contributing individuals can manage the information which is collected on their behalf (Figure 66) by selecting or deselecting the appropriate data feeds.

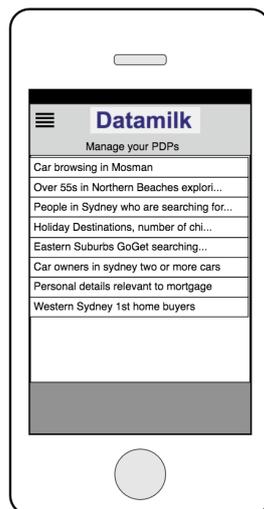


Figure 67: Data Contributing Individual – manage PDPs screen

Data contributing individuals can gain visibility of the PDPs to which they have agreed to contribute their data (Figure 67), and by selecting one particular PDP they can get visibility of the details of that PDP (Figure 68). At the appropriate time, they can withdraw from or renew their participation in the PDP.

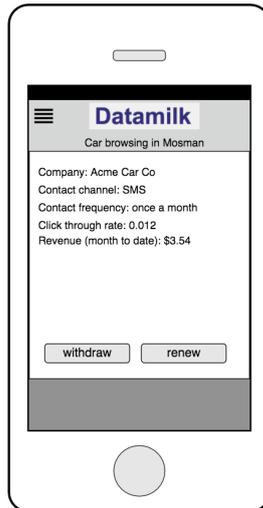


Figure 68: Data Contributing Individual – example PDP screen

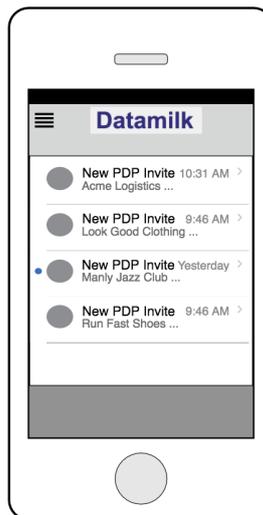


Figure 69: Data Contributing Individual – new PDP invitations screen

Data contributing individuals receive invitations to grant permission for their personal information to be included in new PDPs (Figure 69).

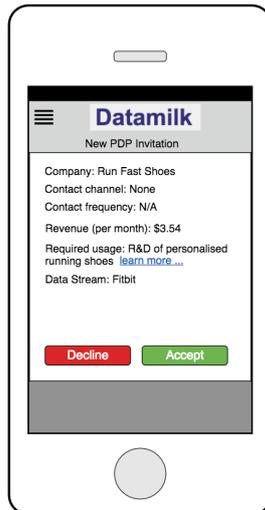


Figure 70: Data Contributing Individual – example PDP invitation screen

By selecting one particular invitation they can gain visibility of the details of that invitation and then either accept or decline the invitation (Figure 70).

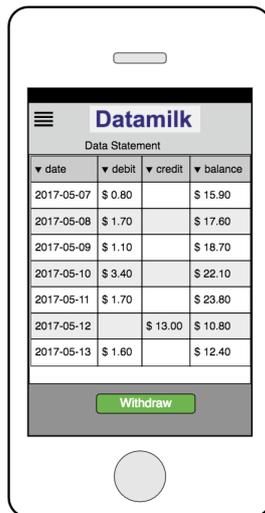


Figure 71: Data Contributing Individual – financials screen

Data contributing individuals can monitor the funds they are receiving from the PPIM and transfer funds to external financial institutions (Figure 71).

10.7.1.4 Platform Owner Market Management Portal

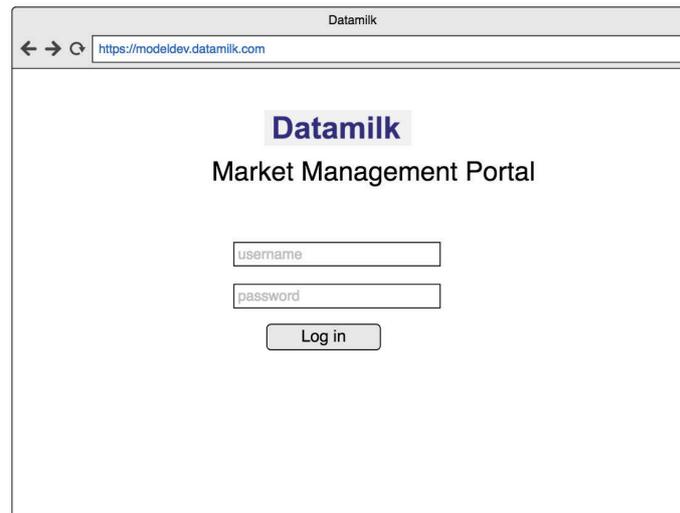


Figure 72: Platform Owner Market Management Portal – login screen

The platform owner can log into the Market Management Portal (Figure 72).

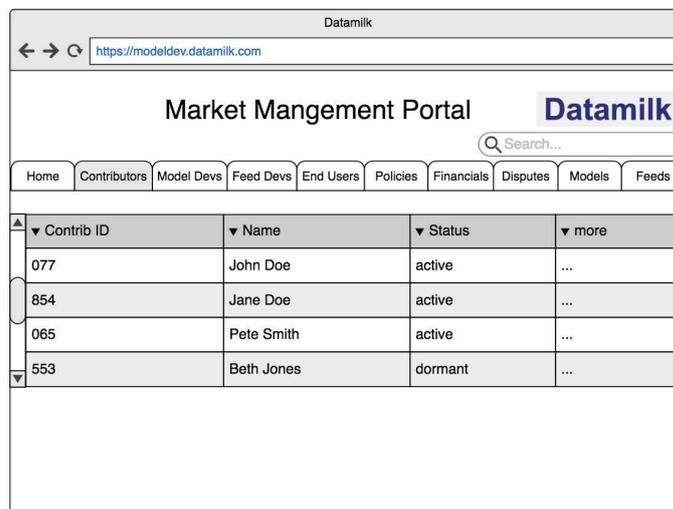


Figure 73: Platform Owner Market Management Portal – data contributing individuals screen

The platform owner can search for and browse the data contributing individuals (Figure 73).

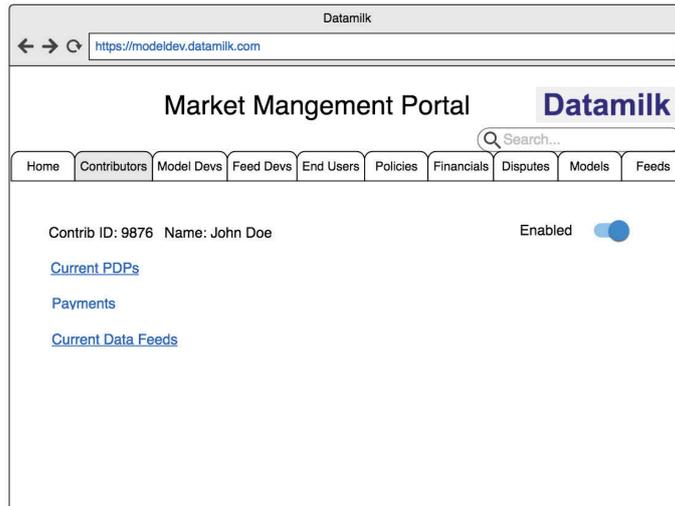


Figure 74: Platform Owner Market Management Portal – example data contributing individual screen

By selecting one particular data contributing individual the platform owner can view their current PDPs, payments and data feeds and enable or disable the data contributing individual (Figure 74).

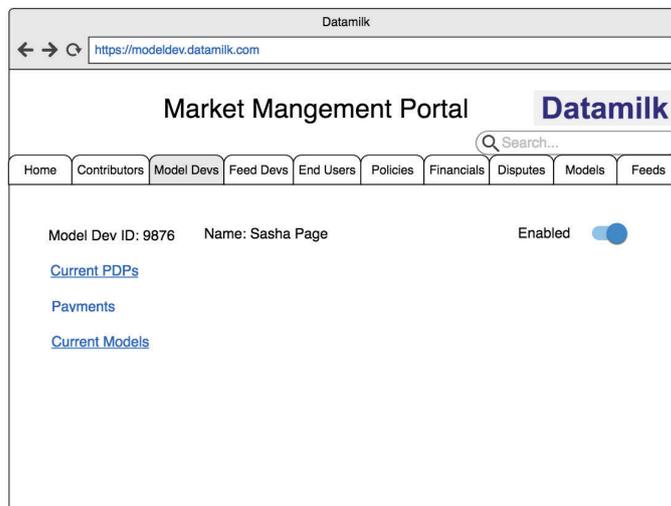


Figure 75: Platform Owner Market Management Portal – example model developer screen

The platform owner can search for and browse model developers. By selecting one particular model developer the platform owner can view their current PDPs, payments and models and enable or disable the model developer (Figure 75).

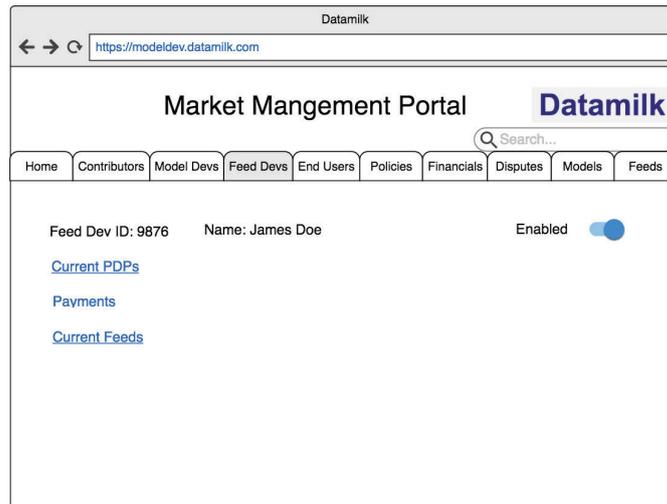


Figure 76: Platform Owner Market Management Portal – data feed developers screen

The platform owner can search for and browse data feed developers. By selecting one particular data feed developer the platform owner can view their current PDPs, payments and data feeds and enable or disable the data feed developer (Figure 76).

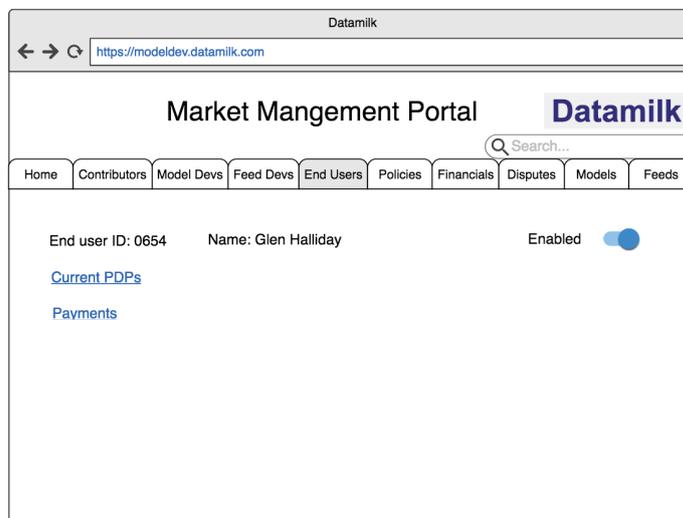


Figure 77: Platform Owner Market Management Portal – end user screen

The platform owner can search for and browse end user. By selecting one particular end user the platform owner can view their current PDPs and payments and enable or disable the end user (Figure 76).

▼ date created	▼ name	▼ owner	▼ description
2017-05-07	sql min records	JBL	...
2016-07-08	k-anonymity	RWF	...
2016-05-09	l-diversity	GJF	...
2017-04-10	fees	DVK	...
2015-12-11	resource allocation	PMW	...
2016-03-12	ranking	OLH	...
2015-11-13	Goodness of fit	JBL	...

[Create new](#)

Figure 78: Platform Owner Market Management Portal – policies screen

The platform owner can browse and search for policies which have been implemented to manage the PPIM. The platform owner can create new policies and edit existing policies (Figure 78).

▼ date	▼ user_id	▼ debit	▼ credit	▼ details
2017-05-07	343	\$ 0.80		...
2017-05-08	223	\$ 1.70		...
2017-05-09	765	\$ 1.10		...
2017-05-10	324	\$ 3.40		...
2017-05-11	221	\$ 1.70		...
2017-05-12	112		\$ 13.00	...
2017-05-13	112	\$ 1.60		...

Figure 79: Platform Owner Market Management Portal – financials screen

The platform owner can search for and browse the financial transactions associated with each participant on the PPIM in order manage the financial matter of those participants (Figure 79).

▼ date	▼ user_id	▼ description	▼ status	▼ details
2017-05-07	343	unpaid bills	resolved	...
2017-05-08	223	unauthorised usage	pending	...
2017-05-09	765	fake data	pending	...
2017-05-10	324	suspected robot	open	...
2017-05-11	221	faulty feed	open	...
2017-05-12	112	incorrect payment	pending	...
2017-05-13	112	missing payment	resolved	...

Figure 80: Platform Owner Market Management Portal – Disputes screen

The platform owner can browse, search for and manage disputes, which have been submitted to the platform owner for resolution (Figure 80).

▼ date submitted	▼ name	▼ type	▼ status	▼ usage	▼ description
2017-05-07	propensity to churn	logistic regression	active	2234	...
2016-07-08	propensity to churn	random forest	pending	N/A	...
2016-05-09	propensity to churn	ensemble	rejected	N/A	...
2017-04-10	propensity to churn	support vector machine	active	3324	...
2015-12-11	propensity to churn	ensemble	active	2245	...
2016-03-12	propensity to churn	logistic regression	active	21	...
2015-11-13	propensity to churn	logistic regression	active	78	...

Figure 81: Platform Owner Market Management Portal – models screen

The platform owner can browse, search for and manage models which have been submitted to the platform owner (Figure 81).

▼ date submitted	▼ name	▼ status	▼ description
2017-05-07	iPhone location	active	...
2016-07-08	Android accelerometer	rejected	...
2016-05-09	Nest energy usage	pending	...
2017-04-10	Fit bit - all data	active	...
2015-12-11	Gmail	active	...
2016-03-12	Firefox browsing history	active	...
2015-11-13	Safari search history	active	...

Figure 82: Platform Owner Market Management Portal – data feeds screen

The platform owner can browse, search for and manage models which have been submitted to the platform owner (Figure 82).

10.7.2 Attendees

The selection of interviewees was guided by theoretical sampling (Eisenhardt 1989) in which we sampled experts who could evaluate the constructs and structure of the proposed designs based on their domain expertise and years of experience. Workshop attendees had an average of 11.9 years’ experience in data related roles.

Workshop attendees were drawn from eight businesses, one from each of Accommodation, Cafes and Restaurants, Communication Services, Government Administration and Defence, Health and Community Services, Personal and Other Services and Transport and Storage. Three from Property and Business Services, four from Finance and Insurance and five from Education (ABS 1993). Workshop attendees were drawn from one micro-sized business, one sole trader, one small, five medium and ten large businesses.

The aim of the workshop was to present the end-to-end design of the PPIM, covering all aspects of the design including the interfaces for end users, model developers, data feed developers and data contributing individuals. The API design was presented. The design of the database used to store and analyse the data was presented.

The aim was also to ascertain whether or not the end-to-end design fulfilled the requirements of the seven design elements of *iSIM*: Service Business Strategy, Customer type and customer value proposition, Service Concept, Service System Design, Service Experience Design, Service architecture and Monetisation. The aim was also to gather input from workshop

attendees to highlight any shortcomings in the design and any area of improvement which should be considered for further research. The workshop was held in a conference room at the University of Technology, Sydney.

To glean the opinions of the workshop attendees a series of 63 questions were developed related to the seven design elements of *iSIM*. In addition to these, one other category of questions relating to the end-to-end design of the PPIM in its entirety was also included in the survey. The questions in the survey are listed in full in Table 26.

Table 26: Survey questions presented in end-to-end evaluation workshop

Category		Question	
Service Business Strategy	Q1	The designed platform delivers an instantaneous 'build-to-order' mechanism for personal data products.	
	Q2	The designed platform delivers a mechanism for the end user to co-create PDP.	
	Q3	The designed platform delivers a revenue distribution mechanism to distribute revenue to the participants on the market.	
	Q4	The designed platform could be extended to include additional participants such as template developers, plug in developers and personal data brokers.	
Customer type and customer value proposition	Q5	The PPIM presented allows data contributing individuals to grant permission for their data to be collected on their behalf.	
	Q6	The PPIM presented allows data contributing individuals to grant permission for their data to be stored in the centralised data repository on their behalf.	
	Q7	The PPIM presented allows data contributing individuals to grant permission for their data to be included in the personal data products (PDPs) to which they give their consent.	
	Q8	The PPIM presented fulfils the value proposition of improved control of their personal data for data contributing individuals.	
	Q9	The PPIM presented fulfils the value proposition of financial compensation in exchange for limited use of their personal data for data contributing individuals.	
	Q10	The PPIM presented fulfils the value proposition of improved transparency on the ways in which their personal data is being used for data contributing individuals.	
	Q11	The PPIM presented allows model developers to participate on a PPIM by co-producing models with the data contributing individuals.	
	Q12	The PPIM presented fulfils the value proposition of financial compensation for model developers whenever a personal data product, which comprises the results of one or more of their models, is executed.	
	Q13	The PPIM presented fulfils the value proposition financial compensation whenever a personal data product, which comprises the raw data collected via one or more of their feeds, is executed to data feed developers.	
	Q14	The PPIM presented allows PDP end users to realise the benefits of utilising PDPs.	
	Service Concept	Q15	The PPIM presented allows data contributing individuals to capture and securely store their data.
		Q16	The PPIM presented allows data contributing individuals to receive requests for access to their personal data which, if accepted, result in financial compensation.
		Q17	The PPIM presented allows data contributing individuals to manage and control access to their personal data.
		Q18	The PPIM presented allows data contributing individuals to fulfil expectations of transparency.
Q19		The PPIM presented allows data contributing individuals to fulfil expectations of opt out ability.	
Q20		The PPIM presented allows data contributing individuals to fulfil expectations of fair compensation.	
Q21		The PPIM presented allows data contributing individuals to fulfil expectations of low transaction costs.	
Q22		The PPIM presented allows data contributing individuals to fulfil expectations of security of data.	
Q23		The PPIM presented allows data feed and model developers to gain access to market to promote and sell the fruit of their labours.	
Q24		The PPIM presented allows data feed and model developers to transparently track the use to which their components are being put and the financial compensation they are receiving.	
Q25		The PPIM presented makes data feeds and models available to be incorporated into PDPs at build time on demand.	
Q26		The PPIM presented fulfils the data feed and model developers' expectations of discoverability of their models and data feeds.	

	Q27	The PPIM presented fulfils the data feed and model developers' expectations of fair comparison between models and data feeds.
	Q28	The PPIM presented fulfils the data feed and model developers' expectations of fair compensation for the use of models and data feeds.
	Q29	The PPIM presented allows end users of PDPs to co-design, co-construct, then purchase and execute PDPs.
	Q30	The PPIM presented allows end users of PDPs to browse personal data assets such as templates and information which has been inferred and extracted from the raw materials of personal data.
	Q31	The PPIM presented allows end users of PDPs to transparently co-create, preview, purchase, execute, delete, renew and monitor PDPs.
	Q32	The PPIM presented allows end users of PDPs to validate the quality of the models they are incorporating into their PDPs.
Service System Design	Q33	PDP templates and the data feed and model outputs are designed as modular components.
	Q34	The API interfaces share common standardised elements.
	Q35	The underlying Personal Data Store (PDS) can be adapted to support PDPs for various uses.
	Q36	The underlying service architecture is extensible to incorporate additional stakeholders, thereby facilitating the expansion of the four-sided market instantiated in the platform prototype to an m-sided market ($m>4$).
	Q37	A centralised cloud based PDS comprising a distributed file system and a massively parallel processing database stores the individual's personal data.
	Q38	Apps which can be installed on physical devices capture at source an individual's personal data.
	Q39	A data model in the PDS tracks which data feeds, models and individual's raw data have been used by the consumers to co-create various PDPs.
	Q40	The data captured at source on individuals' devices are transferred to the PDS.
	Q41	End users co-create PDPs using raw data captured via data feeds and the outputs from models which extract and infer additional information from the raw data.
	Q42	The data model is utilised to track which data feeds, models and individuals contributed to the PDP. When the PDP is executed the contributors' accounts are credited appropriately.
Service Experience Design	Q43	The PPIM facilitates the data contributing individuals' behavioural requirements by allowing them to install the custom built apps to collect their personal data on their behalf and to ingest and store those data in the PDS.
	Q44	Data contributing individuals' cognitive requirements are supported via a data contributing individual facing dashboard.
	Q45	The transparency and immediacy of the service which grants data contributing individuals visibility of and control over the uses to which their personal data are being put would engender trust in the PPIM.
	Q46	Data contributing individuals would have a sense of fairness when they are financially compensated for the use of the valuable personal data they have generated.
	Q47	The modellers' cognitive requirements for visibility and control of the uses to which their models and data feeds are being put are catered for via the modeller dashboards.
	Q48	The modellers' emotive requirements for fairness and equity are addressed as they are compensated for the use of models and data feeds they have programmed and a means to expose their products to a wider audience.
	Q49	The end users' cognitive requirements are catered for via a consumer facing dashboard which connects to the PDS via an API.
	Q50	The end users utilise the dashboard to browse, co-create, preview, execute and renew PDPs – thereby gaining knowledge of the products he or she has purchased and used.
	Q51	The end users' emotive requirement for trust in the quality of the PDP is enhanced through the rating and customer feedback mechanisms.
Service Architecture	Q52	Data feed developers co-create value by facilitating the ingestion of new data feeds.

	Q53	Model developers co-create value by inferring, predicting and extracting additional knowledge from the raw data.
	Q54	Data contributing individuals co-create value by granting permission for their personal data to be used.
	Q55	PDP end users co-create value by constructing the PDP, paying for the right to use it and co-creating value-in-use by executing the PDP.
	Q56	The PDP templates and the data feed and model outputs are modular components which enable reuse and adaptability to co-create customised PDPs.
	Q57	The four interfaces (APIs) share common standardised elements as they interface to a common PDS.
	Q58	The dashboards share common elements such as PDPs, usage, payments and status of PDPs.
	Q59	The underlying PDS can be adapted to support PDPs for various additional uses.
	Q60	The commonality between the API interfaces is extensible and could incorporate additional stakeholders.
Monetisation	Q61	The PPIM is monetised on a percentage pay per use model.
	Q62	Whenever a PDP is executed funds are transferred from the end user's account to the PDP contributors' accounts (model developer, data feed developer and data contributing individuals).
	Q63	A brokerage fee is calculated as a percentage of the transaction and transferred to the PPIM maker's account.
End to End	Q64	The PPIM presented comprises a coherent, end-to-end solution for a PPIM.
	Q65	The APIs presented form a coherent end-to-end API design for a PPIM.
	Q66	The prototype presented integrates the APIs and service architecture into a coherent end-to-end solution for a PPIM.
	Q67	The workflows presented from ingestion of raw data through to competed personal data products comprise a coherent end-to-end solution for a PPIM.

10.7.3 Results

Table 27 shows the percentage of responses for the entire survey which fall into each category. Of the 1206 responses (67 questions x 18 = 1206 responses) 43.0% were 'strongly agree', 35.7% were 'agree' and so on. The table has been colour coded for ease of understanding, with a higher percentage being green and lower percentage being red.

Table 27: Summary of all survey responses

	strongly agree	agree	neutral	disagree	strongly disagree
All responses	43.0%	35.7%	18.1%	2.3%	0.9%

Table 28 shows the percentage of responses for each category of question. For example, there were four questions that asked about aspects of the service business strategy (questions 1 – 4). Of the 72 responses (18 respondents x 4 questions = 72 responses) 52.8% were 'strongly agree', 37.5% were 'agree' and so on. The table has been colour coded for ease of understanding, with higher percentage being green and lower percentage being red.

Table 28: Summary of survey responses by question category

	strongly agree	agree	neutral	disagree	strongly disagree
Service Business Strategy	52.8%	37.5%	6.9%	2.8%	0.0%
Customer type and customer value proposition	43.9%	40.0%	12.8%	3.3%	0.0%
Service Concept	45.1%	30.9%	19.1%	3.7%	1.2%
Service System Design	33.3%	35.0%	28.9%	1.7%	1.1%
Service Experience Design	34.0%	40.1%	21.6%	1.9%	2.5%
Service Architecture	52.5%	32.1%	14.8%	0.0%	0.6%
Service Experience Design	37.0%	44.4%	18.5%	0.0%	0.0%
End to End	48.6%	38.9%	9.7%	2.8%	0.0%

Table 29 contains the full results from the surveys.

Table 29: Survey responses

		strongly agree	agree	neutral	disagree	strongly disagree
Service Business Strategy	Q1	38.9%	55.6%	0.0%	5.6%	0.0%
	Q2	50.0%	38.9%	5.6%	5.6%	0.0%
	Q3	61.1%	22.2%	16.7%	0.0%	0.0%
	Q4	61.1%	33.3%	5.6%	0.0%	0.0%
Customer type and customer value proposition	Q5	38.9%	50.0%	5.6%	5.6%	0.0%
	Q6	50.0%	44.4%	0.0%	5.6%	0.0%
	Q7	61.1%	33.3%	0.0%	5.6%	0.0%
	Q8	27.8%	38.9%	27.8%	5.6%	0.0%
	Q9	50.0%	33.3%	16.7%	0.0%	0.0%
	Q10	33.3%	50.0%	16.7%	0.0%	0.0%
	Q11	38.9%	38.9%	11.1%	11.1%	0.0%
	Q12	55.6%	27.8%	16.7%	0.0%	0.0%
	Q13	44.4%	38.9%	16.7%	0.0%	0.0%
	Q14	38.9%	44.4%	16.7%	0.0%	0.0%
Service Concept	Q15	27.8%	33.3%	27.8%	11.1%	0.0%
	Q16	66.7%	27.8%	5.6%	0.0%	0.0%
	Q17	33.3%	44.4%	11.1%	11.1%	0.0%
	Q18	44.4%	22.2%	27.8%	5.6%	0.0%
	Q19	50.0%	38.9%	11.1%	0.0%	0.0%
	Q20	38.9%	33.3%	22.2%	0.0%	5.6%
	Q21	33.3%	16.7%	38.9%	5.6%	5.6%
	Q22	33.3%	16.7%	38.9%	11.1%	0.0%
	Q23	66.7%	27.8%	5.6%	0.0%	0.0%
	Q24	61.1%	33.3%	5.6%	0.0%	0.0%
	Q25	38.9%	38.9%	22.2%	0.0%	0.0%
	Q26	50.0%	27.8%	11.1%	5.6%	5.6%
	Q27	44.4%	16.7%	27.8%	5.6%	5.6%
	Q28	44.4%	33.3%	22.2%	0.0%	0.0%
	Q29	55.6%	27.8%	16.7%	0.0%	0.0%
	Q30	38.9%	38.9%	16.7%	5.6%	0.0%
	Q31	44.4%	33.3%	22.2%	0.0%	0.0%
	Q32	38.9%	44.4%	11.1%	5.6%	0.0%
Service System Design	Q33	33.3%	27.8%	38.9%	0.0%	0.0%
	Q34	16.7%	27.8%	50.0%	5.6%	0.0%
	Q35	38.9%	33.3%	27.8%	0.0%	0.0%
	Q36	22.2%	44.4%	27.8%	0.0%	5.6%
	Q37	27.8%	11.1%	61.1%	0.0%	0.0%
	Q38	38.9%	27.8%	22.2%	5.6%	5.6%
	Q39	33.3%	44.4%	22.2%	0.0%	0.0%

	Q40	33.3%	50.0%	16.7%	0.0%	0.0%
	Q41	27.8%	50.0%	16.7%	5.6%	0.0%
	Q42	61.1%	33.3%	5.6%	0.0%	0.0%
Service Experience Design	Q43	55.6%	27.8%	16.7%	0.0%	0.0%
	Q44	27.8%	44.4%	22.2%	0.0%	5.6%
	Q45	22.2%	50.0%	22.2%	0.0%	5.6%
	Q46	38.9%	33.3%	16.7%	5.6%	5.6%
	Q47	27.8%	44.4%	27.8%	0.0%	0.0%
	Q48	33.3%	33.3%	27.8%	0.0%	5.6%
	Q49	33.3%	44.4%	22.2%	0.0%	0.0%
	Q50	22.2%	61.1%	16.7%	0.0%	0.0%
	Q51	44.4%	22.2%	22.2%	11.1%	0.0%
Service Architecture	Q52	66.7%	27.8%	5.6%	0.0%	0.0%
	Q53	77.8%	16.7%	5.6%	0.0%	0.0%
	Q54	66.7%	33.3%	0.0%	0.0%	0.0%
	Q55	61.1%	33.3%	5.6%	0.0%	0.0%
	Q56	44.4%	22.2%	33.3%	0.0%	0.0%
	Q57	33.3%	33.3%	33.3%	0.0%	0.0%
	Q58	38.9%	55.6%	5.6%	0.0%	0.0%
	Q59	50.0%	33.3%	11.1%	0.0%	5.6%
	Q60	33.3%	33.3%	33.3%	0.0%	0.0%
Monetisation	Q61	44.4%	33.3%	22.2%	0.0%	0.0%
	Q62	33.3%	55.6%	11.1%	0.0%	0.0%
	Q63	33.3%	44.4%	22.2%	0.0%	0.0%
End to End	Q64	55.6%	33.3%	5.6%	5.6%	0.0%
	Q65	55.6%	27.8%	16.7%	0.0%	0.0%
	Q66	33.3%	50.0%	11.1%	5.6%	0.0%
	Q67	50.0%	44.4%	5.6%	0.0%	0.0%

Based on the responses to the survey, we conclude that overall the PPIM presented to the workshop respondents was well received. Considering the PPIM as a whole, Table 27 shows that $43.0 + 35.7 = 78.7\%$ of all response to all aspects of the PPIM were either agree or strongly agree.

Considering the various aspects of the PPIM, the aspect that was thought to be most robust by the respondents was the Service Business Strategy – with 90.3% of responses being agree or strongly agree. The second most robust aspect of the PPIM was the End-to-End viability of the PPIM with 87.5% of responses being agree or strongly agree.

The aspect that was thought to be least robust by the respondents was Service Concept with 68.3% of responses being agree or strongly agree and 4.9% of responses being with disagree or strongly disagree.

Following Yin (2013), we identify, compare and contrast the main themes in the participants' responses in the free text response section of the survey and from the transcript recorded during the workshop. Of the eighteen experts who attended the workshop, eleven elected to provide responses in the free text section of the survey. Of these eleven experts, nine were of the opinion that the permissions-based PPIM was, broadly speaking, viable, useful and worth pursuing in the form it had been presented.

Participants who expressed positive reaction to the PPIM cited the following aspects specifically: utility of the PPM for marketing, utility of geospatial data for the automotive industry and government services, importance of transparency on the PPIM from the data contributing individuals' perspective, the novelty of the PPIM as a means to the monetisation of personal data, the detailed nature of the PPIM and the utility of the prototype to end users.

There were a number of issues raised by the participants in their free text responses. These concerns were:

- Doubts that data contributing individuals would receive enough revenue to compensate them sufficiently for the effort required to participate on the PPIM.
- The difficulty of verifying the validity of the personal information as it is fed into the PPIM.
- Uncertainty about the number of active users which would comprise a sufficient critical mass to ensure a PPIM is viable.
- The concern that the data contributing user base would be biased towards contributors who are drawn from a low socio-economic background, thereby reducing the utility of the PPIM.

Finally, we calculate the sentiment of each of the verbatim responses using the R package 'RSentiment' (Bose & Goswami 2017). This analysis reveals that of the eleven responses, seven were classified as very positive and one as positive.

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
[1,]	"Sarcasm"	"Negative"	"Very Negative"	"Neutral"	"Positive"	"Very Positive"
[2,]	"2"	"0"	"1"	"0"	"1"	"7"

Figure 83: Output from sentiment analysis of workshop participants verbatim responses

10.7.4 Conclusions to be drawn from the workshop

From the workshop response we conclude that while, overall, the PPIM is well conceived and fundamentally sound, further research is required to resolve a number of important issues raised by the workshop attendees.

11 Successful digital platform adoption strategies as applied to a PPIM

11.1 Introduction

In previous chapters, conceptual models for a PPIM have been designed and tested. The question to be addressed in this chapter is how exactly should a PPIM be initiated in order for it to become successful and widely adopted?

For any m-sided industry platform to be successful it must have sufficient thickness, that is to say there must be a sufficiently large number of participants to make the market operate and for all participants to benefit from their participation.

In order to establish a viable PPIM, there is a need to attract and retain a sufficiency large number of each of the four types of participants: data contributing individuals, end users, model developers and data source developers. However, until there is a sufficient number of all four types of participants, there is little incentive for any of the participants to participate in the market.

The aim of this section is to research and summarise platform adoption strategies and design decisions which have been applied in the past to other successful industry platforms.

These ‘lessons learned’ are then applied to the case of a PPIM and then how these strategies would be implemented to ensure a PPIM is widely adopted and has sufficient thickness to be viable is outlined.

There is a need to analyse the adoption factors in three ways: adoption by end users, adoption by data contributing individuals and adoption by developers (both data feed and data model developers)

This topic is worth researching because designing and expanding an industry platform is ‘a complex and daunting process’ (Hagiu 2009, p. 29) and there is a need to consolidate and synthesise research on this topic. Salazar (2015, p. 2) says that:

The microeconomic, and strategic and innovation management literature, which focuses on issues ranging from network effects, through standardisation and complimentary products, and to multi-side markets, remains fragmented.

According to Thomas, Autio & Gann (2014) there are four distinct types of platform in the literature: organisational platforms, product family platforms, market intermediary platforms, and platform ecosystems. According to this classification a PPIM would be a market intermediary platform since it acts as an intermediary between the four groups of market participants. Gawer & Cusumano (2014) would describe a PPIM as an external or industry-wide platform since it is not designed to operate within a single company but rather provides a platform upon which outside participants can define, develop, market and trade complementary products.

11.2 Determinants of platform adoption

11.2.1 Introduction

This section explores how best to develop ‘business strategies that encourage partners and customers to adopt a particular technology’ (Gawer & Cusumano 2008, p. 30).

In chapter 8 it was established that a PPIM is a good candidate for a multi-sided platform, so now the factors that determine whether or not the platform will be adopted by a sufficiently large number of participants can be addressed.

There is no doubt that the solutions to our research question are missing from the literature for the obvious reason that no working model of a PPIM currently exists. Research into strategies for emerging platforms is sparse and disparate. According to Shuradze, Wagner & Wagner (2015, p. 8) ‘studies on factors that stimulate the platform adoption are very scarce and the topic is under researched’. Furthermore, most research into industry platforms concentrates on pricing and competition between existing platforms rather than on factors that bring platforms into existence in the first place.

The emerging business and economics literature on two-sided markets up to now has not been of much help in this direction either, as it has mostly focused on pricing and competition between platforms, taking their existence as given and without tackling broader strategic questions regarding design, vertical and horizontal scope of MSPs.
(Hagi 2009, p. 2)

The results is that there is a lack of understanding among researchers and practitioners as to how and why platforms emerge (Gawer & Cusumano 2014, p. 429). Indeed there is a dearth of research into the reasons influencing platform adoption (Shuradze, Wagner & Wagner 2015).

11.2.2 Importance of first mover advantage

One of the factors that influence the success or otherwise of an industry platform is the timing of the entry of the platform into the market. There is considerable evidence that pioneers enjoy significant first-mover advantage over later entrants (Zhu & Iansiti 2012).

From a strategic point of view, the platform market maker must decide when to enter the market. There needs to be a balance between the maturity of the infrastructure and business processes and the need to enter the market quickly to gain first mover advantage.

11.2.3 Expansion strategies

Hagiu (2009) argues that to be successful, multi-sided platforms (MSPs) must perform one or both of the following fundamental functions. They must reduce search costs and/or reduce shared transaction costs among its multiple sides. He then argues that to expand, an MSP must make certain strategic decision in order to expand. They must choose the relevant platform 'sides', decide on which fundamental activities to perform and decide on the appropriate levels for the depth and scope MSP functions.

Hagiu (2009) discusses drivers and limitation on depth and breadth for platform providers and proposes that both are import determinants of platform success. The depth of a platform's offerings is important because the deeper the offerings, the more additional value the platform provides for all participants on MSP. The breadth of a platform's offerings is important because, by broadening the platform's reach, additional markets come into scope for the platform's operations. However, adding breadth to platforms involves risks of creating unnecessary competition with firms who may have been otherwise motivated to act as complementors on the platform.

Zhu & Iansiti (2012, p. 88) claim that the success of entrants into platform based markets depends on 'the strength of indirect network effects and on the consumers' discount factor for future applications'. By 'consumers' discount factor for future applications' they mean those factors which measure how much consumers care about applications to be released in the future.

11.2.4 Network effects

By network effects we mean the fact that 'the more users who adopt the platform, the more valuable the platform becomes to the owner and to the users because of growing access to the network of users and often to a growing set of complementary innovations' (Gawer & Cusumano 2014, p. 417).

The literature refers to direct and indirect network effects. Direct network effects can be driven by a technical compatibility or interface standard (Cusumano 2010b), whereas indirect network effects (Zhu & Iansiti 2012) occur when a growing number of participants use a platform making it increasingly valuable to both prospective and existing participants (Cusumano 2010a).

It is well known that network effects play a significant role in platform adoption and increase market thickness. For example, Ayers et al. (2009) found a significant relationship between increase in market-level electronic medical record adoption and overall physician adoption intention in that market.

Liu, Lee & Iyer (2006) studied the factors which influenced how quickly companies would form alliances Microsoft as a platform provider. They found that in the case of Microsoft, as network effects increase, the likelihood of alliance formation increases and thus time to formation decreases.

Employing that real options reasoning, Fichman (2004, p. 139) demonstrate that susceptibility to network externalities, defined as ‘the extent to which a technology increases in value to individual adopters with the size of the adoption network’ have a positive effect on platform adoption.

11.2.5 Ease of adoption

There is much that platform owners can do to make it easy for participants to use their platform. Claussen, Kretschmer & Mayrhofer (2013) describe steps Facebook took to encourage developers to create Facebook apps. These steps include: the provision of developer tools that decrease development costs (Piezunka 2011), strategic subsidies, open and well-documented app programming interfaces, free test facilities, and support forums and conferences.

11.2.6 Variety of product offers

Claussen, Kretschmer & Mayrhofer (2013) also found that a large variety of product offerings encouraged participation on platforms. Zhu & Iansiti (2012) also noted that in addition to the quality and variety of products on offer at the time, participants also take into account products likely to be offered in the future – the so called consumers’ discount factor for future applications.

11.2.7 User innovations

Van der Boor, Oliveira & Veloso (2014) claim that user innovation is key to platform adoption. They argue that innovations contributed by users are more agile and more closely related to needs of consumers, thereby stimulating participation on platforms. Well designed digital platforms also facilitate generative innovations (Yoo 2013) which occur when users use the

platform to create new innovations independently of the platform owner – innovations often not envisioned when the platform was initially designed.

11.2.8 Solving business problems

In order for users to participate on an industry platform the platform must solve a real business problem for participants from all sides of the market (Gawer & Cusumano 2014).

11.2.9 Availability of complementary products

By studying the video console market, Cenamor, Usero & Fernández (2013) found that platform adoption is significantly higher when complementary products are available for the platform. They also found that the impact of complementary products was enhanced when the platform provider managed the complementary products portfolio.

While not studying industry platforms per se as defined in this dissertation, Fichman (2004) examined investment decisions regarding IT platforms broadly defined as technology which enables multiple applications and facilitates related business opportunities. He found that a range of factors had a positive effect on platform adoption. Some of these factors pertained to the circumstances prevailing when the firm was looking to adapt the platform while other factors pertain to the platform under consideration. It is this second set of factors which are relevant to the design and implementation of a new industry platform.

11.2.10 Number of complementors

Piezunka (2011) argues that that a ‘different number of complementors and different types of relationships might be optimal at different stages of the platform’s development’ (Piezunka (2011, p. 204). He suggests that early in a platform’s evolution it may be beneficial to develop a small number of trust-based relationships with chosen complementors. This would allow the platform to maintain the quality of offerings and technology and to avoid the pitfalls associated with a ‘the-more-the-merrier’ approach, including ending up on a ‘too big too fast’ trajectory (Sterman et al. 2007).

11.2.11 Pricing and network effects

Research into two-sided markets suggests that pricing on platforms is a critical issue. Piezunka (2011) recounts how Yahoo, a significant player in the online auction industry, increased its seller fees. This led to a significant reduction of participation by both buyers and sellers.

The issue of whether or not to charge access fees and/or usage fees is also crucial to the success of a platform. An access fee is a fee charged to participant to access the platform. A usage fee is charged every time a participant uses the platform.

Clearly, fees provide a disincentive for participants, so it is important when attempting to encourage platform participation on a nascent market to reduce fees as much as possible to the side of the market which has the greatest network effect. This strategy of discounting the side of the market that has the stronger positive indirect network effect is commonly adopted by nascent platforms (Parker & Van Alstyne 2005; Rochet & Tirole 2004; Weyl 2010).

It may also be possible not to charge fees at all or it may be strategically advantageous to charge negative fees, whereby participants are incentivised financially to participate on the market in order to build a critical mass of users. This is related to the sequential entry strategy discussed in section 11.3.1.1.

11.2.12 Addressing the chicken and egg problem

The chicken and egg problem in platform theory is a problem faced by platform start-ups, namely, that they only deliver value to side A of the platform if there are existing participants on side B. To solve this problem, the platform start-up needs to develop a strategy to get participants on both or all sides of the platform. Evans (2009) gives detailed descriptions of strategies to overcome this problem which will be reviewed here. In section 11.3.1 we discuss how these strategies could be applied to a PPIM and which would have the highest likelihood of success.

The strategy of *sequential entry* involves participants entering on one side of the platform on board and subsequently getting participants on the board at a later time. The *entry with significant pre-commitment investment* strategy involves inducing participants on one side of the platform to invest significantly in the platform before participants from the other side have arrived, in the hope of economic benefit when they eventually do. The strategy of *simultaneous entry of sides* involves inducing participants on both sides of the market to adopt the platform simultaneously through advertising and other marketing activities. When looking to attract participants simultaneously there are five main sub-strategies which can be adopted. The *basic zig-zag* strategy involves alternating between attracting participants from the two sides of the platform, incrementally attracting users from both sides. The *pre-commitment to both sides* strategy involves persuading a number of early adopters on both sides of the platform to precommit to use the platform once it opens. This provides sufficient credibility to initiate the zig-zag strategy. Other strategies to initiate the zig-zag process are the *single and double-marquee strategies*. In the *single-sided marquee* strategy the platform convinces a prestige or marquee user on one side of the platform to pre-commit to participating on the platform. In the *double-marquee* strategy, users of this nature are signed up to both sides of the platform. The *two-step* strategy involves designing a platform in such a way that it initially provides value to

one side of the market, despite the absence of participants from the other side. Once sufficient users are participating on one side of the market, the platform is opened up to participants on the other. The *zig-zag with self-supply* strategy involves the platform provider contributing sufficient first party content (Hagi & Spulber 2011) to attract users on one side of the platform. Once these users are participating in sufficient numbers, the platform provider pulls back from generating first party content and induces participants from the other side of the platform to contribute this content.

11.3 Applying platform evolution best practice to PPIM

This section summarises best practice to platform adoption and apply these best practices to the initiation of a PPIM.

11.3.1 Addressing the chicken and egg problem

This section addresses the question of which strategies from (Evans 2009) will be adopted to initiate a PPIM, how will they be adopted and why.

11.3.1.1 Sequential entry strategy applied to a PPIM

If the PPIM platform provider was to employ a sequential entry strategy (Evans 2009, p. 15), they could look to raise funds from investors, use the funds to compensate data contributing individuals for subscribing to the platform and, once there were sufficient data contributing individuals participating on the market, look to attract end users to the platform.

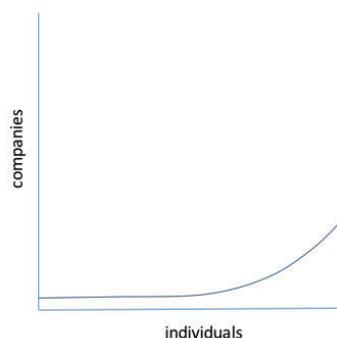


Figure 84: Effect of applying a sequential entry strategy to a PPIM

An alternative means to employ would be for the PPIM provider to provide a free personalised service which entices people to contribute their data to the platform, not for financial compensation but for some other non-monetary reward. This could be the convenience of having a centralised personal data repository or it could be the ability to track one's personal data over time or to compare oneself to one's peers. Information such as how one is tracking compared to

one's friends and acquaintances in terms of wealth, career advancement and educational attainment could be enough to induce individuals to store their data on a nascent PPIM. Once there are a significant number of individuals on the market, the offer to monetise their data could be made.

11.3.1.2 The simultaneous entry of sides strategy applied to a PPIM

To apply a simultaneous entry of sides strategy to starting a PPIM, the platform provider could approach firms, explain the value proposition and try to convince the firm to join the market and also invite their customer base to join the market. Some interviewees suggested in the interviews in Appendix 2 that many companies would benefit from purchasing access to the personal information of their existing customer base. If this proves to be the case it could initiate a 'basic zig-zag' adoption pattern (Evans 2009, p. 16). Once multiple firms (along with their corresponding customer bases) have signed up, PDPs that draw on the personal information across their customer base could then be developed. In this case the adoption pattern would follow a pattern illustrated in Figure 85.

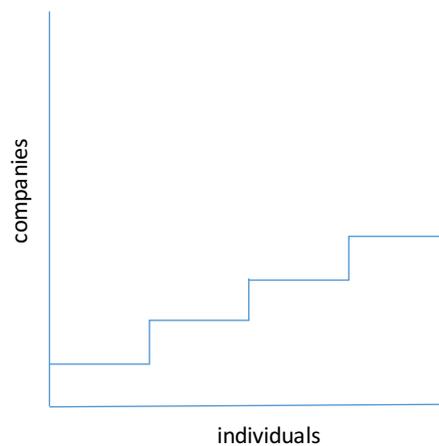


Figure 85: Effect of applying a simultaneous entry of sides strategy to a PPIM

The benefit of this strategy is that the investor's funds would not be spent on attracting users on one side of the market as would be the case if the sequential entry strategy were to be followed. Instead investors' funds would be used for building the infrastructure and paying staff to sign up companies and individuals, thus populating both sides of the market. A second advantage is that the PPIM could be prototyped and tested with real customers as soon as the first company and customer base was enrolled in the market, thereby organically testing and improving the infrastructure, the business model and the PDPs being traded on the market. If the entire customer base, or a large proportion of a customer base agreed to be enrolled in the market

simultaneously it would ensure movement to the ‘north east’ (Evans 2009, p. 8) of the diagrams above which is a preferable strategy.

11.3.1.3 The entry with significant pre-commitment investment strategy applied to a PPIM An ‘entry with significant pre-commitment investment’ strategy (Evans 2009, p. 15) could be applied to a PPIM by getting pre-commitment from companies that they will participate and buy access to PDPs when there are enough individuals who meet their demographic requirements. These companies could also be invited to invest in the platform to build a war chest to attract individuals to the platform. Having a number of large companies who have bought into the concept and who have pre-committed to participating on the market would give the platform and the concept of a PPIM credibility and legitimacy and would act as an incentive for individuals to participate and also to encourage their contacts to participate, as these companies would only start trading on the PPIM once sufficient individuals had signed up. Under an entry with significant pre-commitment investment strategy the participation of companies and individuals would follow a trend as shown in Figure 86.

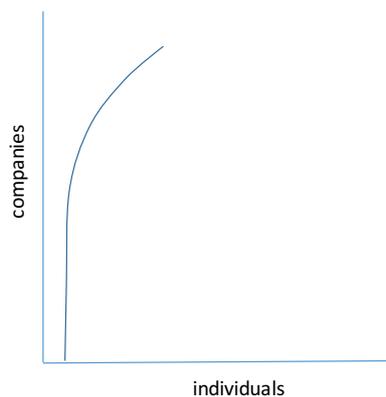


Figure 86: Effect of applying an entry with significant pre-commitment investment strategy to a PPIM

The benefits of this strategy are that, once the first few companies have signed, others may also sign up out of fear of missing out. Furthermore, it would give the platform provider an avenue to test the idea of the PPIM with companies.

11.3.1.4 Preferred strategy

Of the strategies described in sections 11.3.1.1 to 11.3.1.3 the strategy of simultaneous entry of sides appears to have the most benefit. It would enable the platform to grow organically and most importantly it would enable a testing and refinement of the PPIM with real customer data in the very early of stages of development.

Steps have already been taken to implement this strategy. During the course of this research, discussion with four companies were held and the prototype was demonstrated to these companies. Each of the four companies indicated their willingness to trial the solution once a minimal viable product (MVP) had been built. These companies indicated that they would trial the solution and invite their customer base to enter the market and agree to have their personal data streams ingesting in the platform. In this way, as more companies enter the market, additional individuals are simultaneously entering the market. This would lead to creating the network effect necessary for scaling up the platform into a viable market solution. Furthermore, preliminary discussions have been held with potential partners with similar technology strategies to share IP to solve the 'chicken-and-egg' problem and enhance the commercial viability of PPIM. An initial government funding has also been secured to initiate the pilot implementation of PPIM.

11.3.2 Exploiting network effects on a PPIM

There are a number of avenues to explore to exploit network effects on a PPIM. One person's data on its own is of little value. No end user will bid for a PDP comprising one person. But as every additional individual joins, each person's data becomes more valuable. Therefore, there is a natural incentive of participants to encourage their contacts to join the PPIM. The question is how the platform provider could make this incentive obvious and encourage the viral spread of the platform use through the social network.

Network effects work best when the platform in question will not function unless the participants' contacts are also using it. For example, instant messaging services such as WhatsApp only benefit a participant if his or her contacts are also using the product.

One way to exploit this avenue would be to proceed as follows. Suppose an individual signs up to the PPIM and ingests his or her data into the platform. When the individual wants to easily share something about themselves with an end user, for example his or her location or holiday plans, the individuals send the end user a link inviting them to log in to the PPIM to access the information. The platform provider then pursues a 'land and expand' strategy once the end user has begun participating on the PPIM. This is an example of the individuals granting free access to nano PDPs – or factoids about themselves that they want to share without having to enter the data repeatedly.

Additional incentives can be offered, such as paying a premium for access to personal data of individuals who have successfully recruited a certain number of users to the platform.

11.3.3 Developing a network of complementors on a PPIM

The question we address in this section is ‘How can platform-leader wannabes successfully encourage other companies to join their ecosystems and develop essential complementary applications?’ (Gawer & Cusumano 2008, p. 30). As suggested by Hagiú & Spulber (2011) it is proposed that the PPIM platform provider seed the market by developing first party content, i.e. in-house PDPs that can be traded on the PPIM. Following this, there is a need to build and establish trust-based relationships with a small set of high quality complementors who can be relied on to supply high quality PDPs to complement the stock of first party content. Before opening the platform to a wider audience of model contributors, an integrity and quality checking methodology for PDPs must be developed to ensure the quality of PDPs allowed onto the PPIM. There is also a need for a rating and feedback system for PDPs (Hagiú 2009, p. 15) so that users may be enabled to judge the quality of the products they are paying for access to. However, these rating systems can be problematic if not used well (Frean 2016). Quality certification (Hagiú 2009, p. 15) – how much quality assurance will PPIM providers make available to ensure people are getting access to high quality PDPs. Multiple low quality PDPs will undermine trust in the platform. Once the number of PDPs and associated models have increased on the PPIM there will be a need to develop algorithms which will detect fake data, by distinguishing data sets generated by genuine users from those which are generated by scammers.

Finally, when opening the platform to qualified PDP developers, the PPIM platform provider will pull back from creating first party content in order to develop an ‘arms-length relationships so as to achieve a critical mass of complementors’ (Piezunka 2011, p. 204). Furthermore, contributors will only participate on a PPIM if it is easy for them to adopt the platform (see section 11.2.5). The PPIM owner will need to encourage ‘development of complementary applications by providing developers with sophisticated toolboxes that help them to create new, useful applications’ (Piezunka 2011, p. 197). This tool box will comprise a set of open, comprehensive, well documented and supported APIs to allow model developers to search the corpus of existing models, access data on which to develop and test new models and submit models for including on the PPIM.

An equally important consideration is the degree to which API integration into third party tools, such as ERP systems and enterprises data warehouses, will be provided by the PPIM platform provider. Just as it is essential for complementors to be able to integrate with the PPIM, so it is also crucial for the PDPs produced on the market to be integrated into existing business systems.

11.3.4 Timing of introduction of a PPIM

The timing of the introduction of the PPIM will largely depend on the strategy employed to solve the chicken and egg problem (see section 11.3.1). There are a number of possible competitors or similar platforms already appearing on the market (*Personal Black Box 2016*;). Therefore, to secure first mover advantage the PPIM should be launched as soon as a minimum viable product has been developed. When deciding on a launch date, the PPIM platform provider will need to balance the benefits of first mover advantage against the risk of launching prematurely with an inferior offering, failing to deliver and possibly jeopardising the platform's reputation.

11.3.5 Openness

This section specifies which components are to be bundled into the platform and which will be provided as complementary products (Piezunka 2011, p. 198). In general, 'increase in openness results in higher efficiency and innovation output'. However Almirall & Casadesus-Masanell (2010) examined whether the optimal degree of openness is contingent on a multi-product system's complexity and found that for a high complexity platform, a closed organisation is superior.

Trust that the PDPs comprise accurate, real life data generated by real people is critical to the successful operation of a PPIM. Furthermore, trust that individuals' personal privacy will be maintained in line with the permission they have granted when agreeing to be included in PDPs is also crucial to the successful operation of a PPIM.

Since there is no doubt a PPIM is a highly complex platform, we propose that the PPIM platform operates in a closely regulated manner in order to maintain the quality and integrity of the PDPs. The PPIM platform owner will check and certify PDPs before they may be traded on the PPIM.

The PPIM will also come bundled with a small number of pre-built high value PDPs. This will seed the market and allow end users to participate on the market from day one.

In a related issue we discuss here whether and how the PPIM platform owner should itself engage in the production of complementary products (Piezunka 2011, p. 199), that is to say whether or not first party content (Hagiu & Spulber 2011) will be traded on the PPIM.

11.3.6 Pricing and fee distribution on a PPIM

When pricing access to a PPIM the initial strategy will be to charge neither access nor usage fees for data contributing individuals in order to build a critical mass of contributors as quickly as possible. Companies will also be able to browse available PDPs without being charged a fee.

Since the platform provider will contribute first party content in the form of data models and PDPs in the early stages of the platform the question of charging fee for model developers and data feed developers will not arise.

Once a critical mass of data contributing individuals is participating on the market fees will be charged for companies to purchase access to PDPs. Those fees will be distributed equally to all data contributing individuals whose data were incorporated into the PDP.

Once the market develops, and third-party developers start contributing models and data feeds to the market, fees will be distributed to data contributing individuals, and model developers.

12 Conclusion

To conclude this dissertation, it is worth noting that the questions surrounding ownership and value of personal information have not decreased in importance during the course of this research. Media interest (Harrison 2018; Jamrisko & Miller 2018; Shadbolt & Hampson 2018), start-up activity (Smith 2017; Zimmerman 2018) and political concerns (General Data Protection Regulation 2017; Harris & Cilento 2016) about the issue of how best to deal with the problem of the valuable personal information generated daily by e-citizens across the globe has continued to rise.

The issue of our attitude as a society towards personal data continues to evolve, not only at a national level but at an international level. It touches on questions of questions of ethics, equity and, when we consider the digital persona, the very question of what it means to be human in a digital age.

12.1 Contribution

12.1.1 Theoretical contribution

To review the theoretical contributions made in this dissertation we represent Table 30 here for ease of reference.

Table 30 Summary of theoretical contributions by chapter		
	Theory of design and action	Generalisations in the form of a design artefacts
Chapter 5	A partial theory of design and action which says how to construct two-sided PPIM	
Chapter 6	A partial theory of design and action which says how to construct a	Conceptual model of a personal information transaction object (PITO)

	transaction object which could be traded on a PPIM	
Chapter 7		1) A simulation to test the appetite for individuals to participate on a PPIM 2) A data model to store information on a two-sided PPIM
Chapter 8	A partial theory of design and action which says how to construct a four-sided PPIM as an industry platform (developed using <i>iSIM</i> and industry platform theory)	1) A prototype PPIM 2) Service experience blueprint for participation on a PPIM
Chapter 9	A partial theory of design and action which says how to construct a four-sided PPIM as a digital platform (developed using <i>iSIM</i> and digital platform theory)	A service architecture for a four-sided PPIM
Chapter 10	An extension to the partial theory of design and action developed in chapter 9	An extension to the PPIM prototype developed in chapter 7

As claimed in section 1.4, this dissertation makes a theoretical contribution of two kinds: partial theories of design and action that says how to design and construct certain novel artefacts, and generalisations in the form of design artefacts (Gregor & Hevner 2013, p. 339).

The first of these contributions is a partial theory of design and action which says how to construct two-sided PPIMs which can be found in chapter 5. It is a partial theory because it is the product of the first iteration of the design science process and was the initial conception of a PPIM which was improved upon in subsequent iterations.

Chapter 6 makes two theoretical contributions. It contributes a partial theory of design and action which says how to construct a transaction object which could be traded on a PPIM and also contributes a conceptual model of a personal information transaction object (Figure 7).

Chapter 7 contributes two generalisable design artefacts: a simulation which can be used to test the appetite for individuals to participate on a PPIM and a data model to store information on a two-sided PPIM (Figure 9).

Three theoretical contributions can be found in chapter 8. The first is a partial theory of design and action which says how to construct a four-sided PPIM as an industry platform. This theory was developed by drawing upon industry platform theory and *iSIM*. This chapter comprises a novel design approach of utilising *iSIM* to design an industry platform for a specific use case (in this case – the design of a PPIM). The second contribution in chapter 8 is a generalisations

in the form of a design artefacts, namely a prototype PPIM (Figure 21 - Figure 34) and the third is also a design artefact - a service experience blueprint for participation on a PPIM (Figure 19).

In the next iteration of the design science process to address shortcomings identified in the feedback on the theory in chapter 8, we extend our partial theory of design and action to construct a four-sided PPIM as a digital platform. This contribution can be found in chapter 9. This chapter draws upon both *iSIM* and digital platform theory. Chapter 9 also included a design artefact, namely a detailed service architecture for a four-sided PPIM (section 9.8).

Finally, in chapter 10 we extend the partial theory of design and action developed in chapter 9 and also contribute one additional design artefact: an extension to the PPIM prototype developed in chapter 8 (Figure 52 to Figure 82).

12.1.2 Managerial contribution

Chapter 9 also makes a practical (managerial) contribution for potential PPIM platform owners and all four participant types. It demonstrates a novel approach to monetising personal information and for model developers and data feed developers to generate a passive income stream.

This research also presents a specific design example of a number of key features of digital platforms that must be incorporated in the PPIM implementation. It demonstrates a novel example of how to incorporate the contribution of two classes of developers onto a digital platform. It also presents an example of a modular platform design, in the form of a modular service architecture and outlines a stable API interface for a PPIM. Businesses owners and senior executive in industries which are undergoing ‘platformization’ (Constantinides, Henfridsson & Parker 2018, p. 1) will benefit from considering how to monetise the information they collect in the light of this research.

12.2 Implications

The implications of this research are twofold. There are implications for the field of platform design and implications for society as a whole.

Regarding the implications for the field of platform design, the novel use of *iSIM* for the design of both an industry platform and a digital platform shows that the application of a service innovation approach to platform design can be a fruitful one. As an increasing number of industries undergo ‘platformization’ (Constantinides, Henfridsson & Parker 2018, p. 1), the need for effective and successful design methodologies for platforms will be important. Because of

our work, future researchers will not have an example upon which to build when addressing future platform design problems.

Our application of the design science methodology to the problem of platform design also has implications for the field. The innovative use of prototypes, simulations, evaluations workshops, multiple iterations and an atomisation step represents an exemplar of platform design by means of design science.

The implications of this research for society as a whole are wide ranging and far reaching. If implemented, a PPIM has the potential to act as a platform for the exchange of access to personal data of any digitally connected e-citizen. This research lays the foundation for further research into a workable PPIM and ultimately, if implemented, the establishment and maintenance of such a market. This could radically transform the lives of many digitally connected e-citizens, especially in developing countries as they increasingly gain access to the internet. Currently e-citizens receive no direct financial compensation for the personal data they generate and the balance of power in the digital economy is heavily weighted towards multi-national corporations; a PPIM could help to redress this imbalance.

This research shows that individuals are not in principle opposed to participating on PPIMs. This implies that the benefits they perceive in terms of financial compensation, control of their personal data and transparency about the uses to which their personal data are being put outweigh their concerns about loss of privacy and potential security breaches. This research also found that there is significant appetite on the part of companies to create and consume personal data products which could be constructed from accessing personal information.

Together, these two findings indicate that there is sufficient evidence to pursue the PPIM project further and explore ways in which it might be realised in practice.

To contextualise our findings, we need to recognise that the concept of the monetisation of personal data is a controversial one and there are arguments both for and against. These arguments take place in the context of ‘the trade-offs between open access to information which enables economic efficiency—and an individual’s right to privacy’ (Milberg, Smith & Burke 2000, p. 35). The main arguments in favour of monetising personal data are that the primary producers of the personal information, the data contributing individuals, as we have referred to them in this dissertation, deserve fair compensation for the value of the data they have produced. Also in favour of a PPIM is the potential it has for unlocking the value of personal information through the additional transparency it provides. Furthermore, joined streams of personal information identified with the individuals who generated them allow fine-grain segmentation of a customer base. Queries with multiple conjunctions can be run on these data which allows organisations to identify and develop niche markets (Dalgic & Leeuw 1994, p. 40) and to

personalise products and services (Searby 2003). Companies use detailed aggregated data to create targeted offers, innovate coupon offers and improve loyalty (Acquisti 2010) which benefits both individuals and companies.

In contrast to these arguments stands the concerns many individuals have about the very concept of monetising their personal data which they view as part of their digital personae (Clarke 2014), a digital extension of their physical, mental and emotional being. They object to selling access to this aspect of themselves and aim to redress the ‘asymmetry in power that broadly favours institutions (both public and private)’. In addition to this there are the very real concerns about the maintenance of privacy and the security of personal data which have been found to be of primary concern to many digital citizens (Milberg, Smith & Burke 2000).

Governments are also taking a greater role in regulating the way in which personal information is accessed and utilised. Evidence for this change in attitude can be found in changes to Australian federal law which give individuals increased control and visibility of the fate of their personal data. For example, in February 2017 amendments to the Privacy Act were passed meaning that ‘if an organisation subject to the Privacy Act incurs an “eligible data breach”, it will have to alert ... the people whose data has been compromised.’ (Redrup 2017). Furthermore, the Australian Productivity Commission (Harris & Cilento 2016) in its 2016 draft report of the value of both private and government data assets recommended:

greater control for individuals over data that is collected on them by defining a new Comprehensive Right for consumers. This right would mean consumers: retain the power to view information held on them, request edits or corrections, and be advised of disclosure to third parties. (Harris & Cilento 2016, p. 2)

Internationally, Europe has strengthened its privacy and data protection legislation, with the introduction of the General Data Protection Regulation (GDPR) in May 2018 (General Data Protection Regulation 2017).

However, although the exact effect a workable PPIM would have on the personal information market cannot be known in its entirety, we can make some informed comments on the likely effects.

First, it has been shown that opaqueness in organisations, such as financial institutions, telecommunications companies and auto manufacturers, obstructs economic freedom and, as a consequence, economic development (Miller & Holmes 2010). Therefore, it is reasonable to infer that the increased transparency afforded by a PPIM would promote economic progress for both companies and individuals. Openness and transparency increases both value and participation on digital platforms (Parker, Van Alstyne & Jiang 2016). Muschalle et al. (2013, p. 11) argue that ‘price transparency can lead to a positive development of the entire market

place'. The heightened transparency afforded by a PPIM would also go some way towards ameliorating the 'market failure' (Holdren & Lander 2014) brought about by the binary take-it-or-leave-it set of terms offered by many online service providers. This situation is compounded by the fact that many individuals do not read the terms of service in the first place (Stephenson 2017).

Secondly, it has been shown that in other markets, such as those for books (Waldfoegel & Reimers 2015), travel (Tse 2003) and apparel (Oxborrow & Brindley 2014), disintermediation of traditional participants causes significant transformation of the market. Therefore, it is reasonable to infer that transformation of a similar magnitude in the market for personal information could be brought about by the creation of a PPIM.

Thirdly, anecdotal evidence collected by the authors in the course of previous research (Farrelly & Chew 2017) confirms that many companies would dramatically change the way they interact with customers and prospects should a workable PPIM be enabled. Respondents in previous interviews stated that a PPIM could transform multiple functions within their businesses including: recruiting, store location planning, direct marketing, call centre interactions and research and development for new products.

Classical economics states that complete information and transparency creates economic efficiency (Miller & Holmes 2010). The creation of a workable PPIM would increase transparency and therefore economic efficiency. We have argued that the raw materials of a valuable information asset are being generated but individuals currently lack the means to capture, aggregate and trade permission to access their digital aggregated assets. A PPIM would address this need.

12.3 Fulfilment of aims

Section 1.3 outlined the research questions (RQs) we aimed to address in this dissertation. It was stated that the aim was to identify theoretical frameworks which would be useful in designing a PPIM. This RQ was addressed in chapter 2. The aim was also to examine how market engineering theory, value co-creation theory and industry platform theory could inform our understanding of PPIM design. These questions were addressed in sections 4, 3.3, and 8 respectively.

The second research question concerned past attempts to establish a PPIM and the research as to why they failed to succeed. This RQ was addressed in chapter 4.

Chapter 5 addressed the third RQ concerning the design of a two-sided PPIM. This chapter outlines an initial conception of the IT infrastructure and for a two-sided PPIM and presents a number of scenarios which would be enabled by this design.

Chapter 6 comprises an in-depth examination of our RQ concerning the nature of the transaction object on a PPIM including an evaluation by a panel of industry experts. The fifth research question, concerning the design of a PPIM as four-sided industry platform, is addressed in chapter 8.

The sixth RQ concerning the thickness of a PPIM was addressed in two parts. Chapter 7 examined individuals' appetites to participate on a PPIM – this addresses one side of the thickness question – while in section 10.7 the end-to-end effectiveness from an end user perspective of the PPIM with a panel of industry practitioners was evaluated. This section addresses the thickness question from the side of the end users.

Finally, the seventh RQ concerning the practical implementation of the PPIM is addressed in section 11.3.

12.4 Limitations

There are a number of limitations to this research which could be investigated in further research.

First, the evaluation was conducted on a very small number of experts and due to time constraints, the interviews were somewhat brief. While the opinions of a small panel of academic and industry experts are valuable, the limitations of this type of evaluation must be recognised when designing a platform which is intended for a mass market.

The second main limitation of this research is that the prototypes designed in chapters 8 and 10 were somewhat limited in functionality and the user interface was not as well developed as it could be. Some parts of the prototypes were still at the wireframe stage of development. While these were sufficient for presentation and evaluation purposes, a more fully functional prototype would be needed for an in-depth evaluation by a wider audience.

The third major limitation in this research is the absence of a proof-of-concept to test the theory of a PPIM in action. For example, the research in chapter 7 is based on simulated requests from companies to access individuals' social media and other personal information. To obtain a more informative and accurate understanding of individuals' willingness to grant access to their personal information it would be necessary to undertake a pilot or proof-of-concept in which actual payments were made to individuals. The research in this dissertation presents schematic designs and prototypes for evaluation by expert panels but does not collect behavioural data on how individuals would participate on a PPIM in practice.

Other limitations are that the transaction object designed in chapter 6 is somewhat rudimentary. There is scope to explore alternative PITO design such as a schema-driven design (XML Schema or RDF Schema). This development could enable more advanced processing of the personal information and consequently enable some relevant functionality based on this.

13 Further Research

This dissertation has investigated the feasibility of a permission-based PPIM from two sides, that of the data contributing individuals and that of the data product consuming companies. Further research could be conducted to investigate whether model developers and data feed developers would be willing to participate on such a market. We have concentrated on the data contributing individuals and data product consuming companies as they are the most important participants on a PPIM. Without these two groups of participants the market most certainly would not be viable.

One of the key factors in building a successful industry platform is the reduction of transaction costs and the exclusion of the ‘middle man’ wherever possible. Areas for future research include the possibility that while the aim of the PPIM could be to create a tradable asset class, it may not be suited to controlling the trading of that asset. Research into the use of the Blockchain distributed ledger to manage and transact the exchange of tradable permissions for access to the new asset class created by the PPIM may well be fruitful. For example, Tapscott (2016) claims that by using the Abra Blockchain (Abra 2016) technology it was possible to reduce the transaction cost for transferring money from Toronto to Manilla from ten percent to two percent. If these and similar claims can be verified the application of Blockchain technology to a PPIM could be worth further investigation.

Other aspects related to the development of a PPIM which could also benefit from further research include the variations in human perception of online interfaces (French & Hancock 2017), novel architectures for clouds, programs, devices, and users (POMI 2019), the privacy economy (Lam 2019) and the Actor Model - one of first conceptualization of a flexible, adaptable model for multi-sided end user applications (Russell 2011).

However, the proof of the pudding is in the eating, and the real test of this research and the most pressing and, to my mind, most exciting next step would be to test this research in practice and attempt to initiate and run a working PPIM in the real world. As indicated above, steps to implement a PPIM in practice have been taken. Initial discussions with various sources of

funding have commenced and are proceeding in a positive manner. Partnerships with startups who are aiming to create a PPIM similar to that described in this thesis are also underway. Clearly further details about the implementation of a PPIM in practice are commercial in confidence and cannot be published in this thesis and the academic literature.

Basing the prototype for a PPIM on simulated data is a clear limitation for this research. To better understand end user's appetite to participate on a PPIM there is a need to conduct a proof-of-concept project in the real world. An additional limitation is the limited sample size of the evaluation panel. A larger survey of potential end users would give a more informative insight into the viability of a PPIM.

As with every start-up, the potential risk factors associated with a company which aims to establish a working PPIM are numerous. The major risk factor is that one side of the market (either data generating individuals or end users) would not be sufficiently engaged in the market to make the market viable. Another major risk factor is that one or more of the major online social networks would alter their end user license agreements to exclude the inclusion of their data in a PPIM, thereby undermining the value of the market. Adverse changes in government regulatory frameworks such as GDPR which govern commercial activities based on insights derived from personal information represent a third category of risk.

Furthermore, this chapter concerns itself with the design and evaluation of a workable PPIM and its utility from a business perspective. Further research is needed to refine the proposed design taking into account the feedback gathered in the semi-structured interviews, such as ability to track and measure return on investment and integrate the insights derived on a PPIM into existing business ecosystems.

There also needs to be additional research into how different types of personal data will be accommodated on a PPIM. Some personal data such as health and financial data is highly privacy sensitive and would need to be treated in a very different manner to less sensitive data such as social media posts. The willingness of data contributing individuals to share insights derived from data of varying sensitivity is explored in chapter 7 but further research into the kinds of guiding principles which would be needed to be considered for factoring in such sensitivity in the PPIM design is needed in this area. This research could center around developing guiding principles to deal with personal data of varying sensitivities.

14 Appendices

14.1 Appendix 1

Selected transcripts of interviews regarding the nature of the transaction object

14.1.1 Transcript 1: Interview with Interviewee 1

INTERVIEWER: Just before we go any further, do you mind if I record our discussion just so I can make some notes on it later?

INTERVIEWEE 1: No, I don't mind that's OK.

INTERVIEWER: Thanks very much. So have you been busy as always?

INTERVIEWEE 1: Yes. It's pretty good.

INTERVIEWER: Did you have a chance to look at the paper that I sent through?

INTERVIEWEE 1: Yes, I quickly skimmed through the paper but could you maybe quickly repeat the key points from your view?

INTERVIEWER: I don't know if you've got your paper in front of you but there is one diagram, particularly that probably sums it all up, which if you have it makes it easiest to go through. Let me just find which page it would be on. Do you have the paper on your computer there? It's on page 10.

INTERVIEWEE 1: Page 10, OK.

INTERVIEWER: Figure 3 page 10. Basically the idea is this, right, that people as they go about their daily lives with wearable technology and with mobile technology, they're generating a vast amount of very valuable personal information and at the moment there's a secondary market for that information, whereby people give it away for free basically, but in this in return they get services such as Facebook and Gmail and so on. Or Skype for example. That is no primary personal information market whereby you can exchange your data or permission to use your data for actual cash and secondarily and I think this is something you pointed out in your work is that it's very, there's no visibility or there's very little visibility around what the companies are doing once you hand over your personal data to them, what they're doing with it. My idea is there's an opportunity here to create a primary personal information market and the way it would work is the individuals, you would build some apps and APIs, that at source, as the individuals create personal data whether it's email, SMS, location data, whatever it is, as they're created it's captured on their behalf and put in a personal data store. So that is part 1, 2 and 3 on

the diagram and then what happens is that data is put onto a logical data model, it's joined with other users who have similar data, and then companies, if they have a need to use perhaps a novel personal data product, what they do is they put out a request to the market. They say this is the type of list that we're looking for or this is the type of personal information that we want and they can put out a bid for how much they're willing to pay for it. And in the market maker, passes that bid on to individuals whose data fits their profile, they either accept or reject the bid. If they accept it, then their data is included in the data product which is created. If they reject it, then it's not. And then the company gets access to the personal data and can use it. The actual data product itself will be based around permissions. So they don't get to own the data, they get permission to use the data. And those permissions can be a whole variety of different parameters which can be tweaked to vary the permission. It could be exclusive right to the dataset for a certain period of time. It could be the ability to run a list or run a query against that data and use the results. It could be permission to send direct marketing to the people who created the data through various media. So the permissions and the way those permissions are scoped out, they will be the parameters that define the personal data product. So depending how those permissions are framed up, there are the parameters that would define what the personal data product is. To keep visibility and transparency there will be two dashboards. One on the business side, that's number six on the diagram. And that will be the dashboard that summarises what permissions the company has purchased and when they expire and so on. And at number five on the diagram there's a personal data dashboard which summarises for the individual the permissions that they have sold, and the revenues they're getting. So that in a nutshell is how it would work. And I gave a couple of use cases about how it might apply to different industries and it's got a lot of applications in marketing and politics, where you want to contact a certain micro sector of the population to sell your message. So really the purpose of the interview was to ask you first of all in a very general way, what your views of the proposed solution are: do you think it's feasible, do you see any major flaws in it. Do you see any interesting, any sort of novel approaches in there?

INTERVIEWEE 1: The first question I think that's a pretty crucial one is which comes to my mind is how much transaction costs in the market and which prices do you expect or personal data being traded so what is a single profile of data worth, what are companies willing to pay for it which definitely determines whether the model is feasible or whether it's not feasible.

INTERVIEWER: That's a very good point. There's a bit of research out there already around quantify how much for example, there's a paper that looks into how much each user is worth to Facebook so there's actually a little bit of research out there which will help the market maker get an idea of the sort of starting price to propose. The good thing about a market mechanism is that it helps establish the price and that price will obviously vary based on demand and supply.

But one of the things I see in working in the commercial world is that personal data is on its own, individual streams of personal data is pretty valuable. But what's really valuable is when it's joined and that's the thing about a personal data market is because the individual is at the centre of it. They have the ability to make that join and identify the diverse streams with the single individual which multiplies the value of the data exponentially.

INTERVIEWEE 1: As I get this the idea is to pull the data off a great many individuals and sell this pool of data to the companies and the market mechanisms then manages to distribute the profits to individuals again.

INTERVIEWER: Just one key point to really clarify here, we are not selling the data per se. So the individual owns the data and cannot alienate that, there's some very interesting research by Schwartz that demonstrates that this point about the data being inalienable is rather important because the data is almost especially your personal data is almost like a digital persona and it's important that that can't be sold. What the market is doing is it's commodifying and selling permission to use that data within certain bounds. That's what's going to be called the transaction object. The transaction object on the market is permission, not the actual data itself. So there'll be no physical transfer of that dataset.

INTERVIEWEE 1: Another issue that I see is how do you get the individuals to actually contribute the data to the marketplace, so there's the issue of convincing and motivating individuals. There's the issue of privacy concerns that may inhibit people and there may be the issue of having a quite high effort of managing personal data through a personal data store, that individuals may not be interested in or lose interest in after some time. So how do you deal with these issues in the model?

INTERVIEWER: These are absolute key points so there's a few different points you've raised there. One is getting the critical mass of users. That's going to be a separate part of my research because basically what I'm doing here is I'm canvassing a very high level view and my dissertation is going to be looking into these aspects and drilling down hard into them so I'm writing a paper at the moment focusing just on the transaction object and what it will be. Getting critical mass is going to be a key point so they may actually have to be an intermediary step between going from nothing to a full market where you offer a data aggregation service which basically makes it convenient for people to collect and manage their personal data. And when you get to critical mass you can start marketing on their behalf. As far as privacy goes I guess the point here is that for most of us, we already the point is that everybody gives up their privacy already if what they're getting in return is worth it. So you know you go to the doctor, you tell them very personal and private information because what you're getting in return is something more valuable. And you go to your financial planner, you give away very personal information

about your plans and your future and your income because what you get in return is worth it. The point here in the market is my argument is, people are giving away these personal data. But what they're getting in return is not commensurate for the value of it. They're getting these few services like Gmail and Facebook and Skype but their data is actually worth a lot more than that. So the market will hopefully quantify that and facilitate it. There is a lot of research that shows especially younger generation, they're a lot less concerned about privacy than older people are. They're almost the online generation and it will be a lot more palatable to younger generation who are so used to sharing data online.

INTERVIEWEE 1: What I see in here is if you mention younger generation is also the problem that the population you get to engage in managing and trading its own data via a market or personal data store is biased. So it will be those who are already interested in these topics and motivating the other parts of the population which might have less knowledge and willingness to participate in such a market may be difficult. As a company that reduces the attractiveness of the permissions and results you can obtain from the datasets, because you basically get only a small subset of the population. That's one limit, I also have no solution for it.

INTERVIEWER: I guess the point there which I think is a very important point and I guess the way I look at it is, if you want to make an inference about a population for example then yes you do need a representative sample, right and hopefully a random sample. But if you're actually looking for to target a micro segment, you're happy to take as many members as you can who meet that profile. And whether or not they're representative or whether the sample doesn't really bother you so one of the use cases I demonstrate the market using for an example a cafe. Cafe in a suburb, called Manly, and their booking rate is down on Tuesdays. And they sell Mexican food and they play jazz. So what they're looking for is adults who use Facebook, who like Mexican food, who like jazz and who eat out more than 4 times a month. That's the profile they're looking for. That's a very, very specific micro segment. But if they're in a market place where they could put that up as a data product and what they want to do, they want permission, they don't want to know the people's names. What they want permission for is to advertise their restaurant to them on Facebook. So it's pretty non-invasive, people are already seeing ads on Facebook, they'll just see slightly different ads. It's totally anonymous so the user's not giving up much privacy, but they're going to get paid for that, that small amount of privacy, they do get paid up. That's a scenario where the fact that it's not representative of the population is not really a concern to the company that's buying the list. Or buying the use of the list. Just to go back to your point about the effort of the transaction costs. I think that was a very good point and so one way that typically happens in markets is you get brokers. So brokers will act on behalf of a large group of individuals and what they do is they say well we'll build a risk profile or a privacy profile for you and if you sign up as a member and you're in a low to medium

privacy, and you want to opt out of all geo-location data but you're OK with social media, then they will auto accept on your behalf any data products that get proposed that fit your privacy profile. And that way you still get a revenue. Obviously they take a cut but you still get a revenue without the effort of having to vet every data product which is proposed on the market.

INTERVIEWEE 1: I like the idea of having brokers but I'm a little bit sceptical of whether you can nail down privacy by having categories of data or having different sets of data such as location data, house data, consumer data and so on because much of the privacy issues arise if you for instance you use consumer data in an unexpected way so for instance if someone sees that I go shopping, I change products every week, they may be able to infer certain circumstances of my life seeing that I'm either sick or pregnant or I have so and so many kids so having just a data category seems to me like a naive approach to opt out because the whole picture is more than a simple sum of its products.

INTERVIEWER: That's a very valid point. So how would you ... just exploring this a little bit further. How does this model that I'm proposing, how does it relate to the work that you're proposing? What similarities and differences do you see there?

INTERVIEWEE 1: The biggest similarity I see is the big stress on giving individuals control by having these technologies such as personal data stores trying to engage them into managing their own personal data, storing it into a repository and deciding for themselves whether they want to sell it, give it away or to license it.

One difference that I see is that in your model as far as I understood it's possible to commission data to some parties that you actually don't know, so you're commissioning data to someone anonymous in the market who you actually have no contextual relationship to. So this wouldn't be possible in our model because it opens up a lot of privacy issues that individuals do not have control over what happens to the users of the data once they don't know who is in actual possession of it.

INTERVIEWER: That's a good point. The way I'm trying to address it is with this dashboard, if you have an individual say at Number 5 on the diagram, the individual personal information dashboard which allows the individual to see which companies have purchased which access and when they expire and what permissions are being granted. But I also agree with you that exactly how that dashboard will be designed, it could be a vast amount of information and as we know the way information like that is presented can influence how easily it is to digest and how easily it is to understand and highlight the key risks and so on. I think the actual design of the user interface would be incredibly important and it could be a good area for some experiments and some prototypes to see what the best type of interface is there for people to get full transparency as possible.

INTERVIEWEE 1: There's so many contextual factors played into the use of the data that it's difficult to represent in a kind of structure. But it's a general issue we have to deal with also in other areas.

INTERVIEWER: Excellent. I think that's it. Do you have any other comments or anything else you wanted to raise on that?

INTERVIEWEE 1: I find your model very interesting and I see a lot of similarities to our work and it's good to see that there is something happening and there are some workable platforms coming up that are in operation so if you have something up, a demo or prototype, I would be very happy to engage in that prototype.

INTERVIEWER: Thank you, that would be excellent. I would like to take you up on that. Are you familiar with the work of Sandy Pentland? He's doing some interesting things with the open answers or the open personal data store. It's pretty interesting. He built some demos, he and his research team.

INTERVIEWEE 1: I like his work. I think his whole idea is more centred around the idea of having open data, about social behaviour, social circumstances and having data on an aggregate level, so his data is less personal than what I think your model is trying to target.

INTERVIEWER: That's correct. It's a bit of a different focus but I think some of the technology they're building could easily be adapted or could be used in a different context so something to watch about. Thanks very much for taking the time and your comments were very, very good, I like it when somebody who, I'm speaking with someone who obviously knows the area and pushes back. That makes me think more deeply about some of the issues because that's always a good way to improve your work so I really appreciate your time and I'll definitely be keeping an eye out for forthcoming publication and perhaps we might meet up in a conference one day.

INTERVIEWEE 1: I look forward to that. Thank you for the call.

14.1.2 Transcript 2: Interview with Interviewee 2

INTERVIEWEE 2: My thought about this is that there is information asymmetry here in the sense that it's the buyer who knows more about the value as well as the specifics of ... basically the value at the end of the day. And the seller really is at an absolute disadvantage as they have no idea as to what ends or basically to what benefits their personal information can generate. So here's an idea that I was thinking, is that, you can give the buyer the option to pay for the information that they're already collecting from the user or how are you going to call, just call

them the individual, right. Yes, so we could give the buyer an option for the individual to not disclose to anyone else, to any other buyer information they're already giving to that buyer.

It's a reverse auction. From Facebook's point of view, if you have an app that's basically harvesting information that typical user would already be providing on Facebook, but now you're creating a secondary market where you offer the same information provided in Facebook to other Facebook wannabes like eBay and Amazon. Now can you imagine how Facebook would bid up the value of that information because the information asymmetry is now in the seller's advantage. Because all the buyers would have the information to know what that data from the individual means to bottom line, because at the end of the day, data is useless without some way of auctioning it, right?

This also takes the onus of classifying of data from the app from the individual to the buyer. So if the buyer has a new way of tapping individual data. Not a problem, because app just begun harvesting it and makes it available to a secondary market and it's the onus of the original buyer who may setup in novel data capture of the individual to then ensure that their channel is secure from competitors.

It comes back to other ideas I have about how to secure intellectual property. There's more in that space. I thought I might share with you that this would give you a way to simplify your approach, there is still a need for a personal information market but I would say it's a secondary rather than a primary. So rather than acting as a gatekeeper, you provide an open auction and the original buyer basically has the option to lock it from other competitors, at a price.

INTERVIEWER: That's one thing I've been toying with, writing a paper at the moment around the transaction object and what exactly is going to be bought and sold and one of them is buying and selling exclusive access to the data, as opposed to open access. Go on then...

INTERVIEWEE 2: This is basically these suggestions I'm giving you assumes by default that there is no exclusivity to personal data which is actually the standard today.

INTERVIEWER: The way I was thinking of it is each of these companies like Facebook and so on, they all have access to the data stream generated on their platform. But in fact, none of them have access to the joined data stream which only the individual is in the position to join right and a joined data stream is actually much more useful and much more valuable than an individual data stream. So that's one point. And the other point is, companies have access to that individual data stream of existing users or customers. But not of prospective users or customers. And that's another need I've seen my idea would fill as companies would, obviously people would, companies interest in acquiring new customers. So if there was a market on which they could bid for access to data of individuals that they don't currently do business with then that would be valuable to them.

INTERVIEWEE 2: I do agree that the Holy Grail for everyone is to have that single stream, right? To be able to profile an individual and that has non-commercial implications as well for all things identity. I guess my suggestion would help you do that because it's going to be a lot of effort. There's no one company ... going by Game Theory, there's no one company incentivised to create this open stream. We're not going to be able to capture the entire benefits of that open stream.

INTERVIEWER: That's what my idea is. I'm saying that only the individual is one, incentivised and two, has the permission to actually create that stream. But do go on ...

INTERVIEWEE 2: But then the individual as the average grandmother who buy a Wii and so on may possibly be incentivised but they do not have the capacity, the drive, it's the default state is the do-nothing state. It's in addition to your primary, you setup a secondary market, it's almost the way how modern standards have evolved today where at some point it gets too difficult for everyone to have their own proprietary, ports of proprietary protocols and just for, just out of self-interest. Because it's cheaper to collaborate in the end than to up bid to each other to arrive at that same open stream, single open stream standard without any additional investment from app developer.

INTERVIEWER: Like open APIs and so on.

INTERVIEWEE 2: That's correct. It would lead to a conversion from a secondary to the primary.

INTERVIEWER: Thanks that's a very good perspective actually and that's partly why I'm doing this interview to get different perspectives because people from different backgrounds look at it from different ways. So one thing I wanted to explore with you is because obviously with a market there's obviously two sides, there has to be individuals who are willing to share and there's companies who's willing to buy. So I'm not sure with your background, if have you some examples with companies who are looking or you think would be interested in accessing this type of...

INTERVIEWEE 2: Absolutely. I was a consultant and I've also been in banking for quite some time in government. It's an interesting one with government because they're the ones who most needed data but least understand that need. But I know the ones who are most incentivised financially are the telcos. You're looking for enterprises which deal basically commodity products, ones which they want to tailor their product but basically have no means of doing so because the product itself does not lend itself to specific taste without tailoring it in the first place, a bit like a catch 22 whereas we look at premium brands, their products are basically shaped toward a particular customer they really want to capture and the investment is in the past. The investment has been around a perception whereas you talk about telcos or even banking, telcos, banking, I guess to a lesser extent, retail? Like just generic retail like groceries, I guess.

Those sort of companies, I know right now, we're getting data from Vida, we're getting a lot of data from Australia Post, just to get an idea of what individuals are doing and we even have like, Head of Customer Experience now whose sole purpose is to know everything about the customer and lift up the customer value of business.

INTERVIEWER: That's certainly been my experience and there'll be a lot of appetite up for that. I guess the other question is obviously some members of the community are going to be more comfortable with sharing their data. I'm thinking Gen Y and so on you have almost been brought up living their lives online, I think they're going to be more comfortable with it and used with that.

INTERVIEWEE 2: There's 2 ways I guess to tackle this proposition for a personal information market. On one hand is self-interest, you're basically monetising what you're giving away for free. On the 2nd level, it also provides transparency into what you're giving away and portray it as an altruistic exercise which to some extent it is actually, to say that the cost of adding transparency to basically the data you're generating for others to take advantage of requires someone to pay for it and the best way to pay for it is to monetise it and someone else to clip the ticket.

INTERVIEWER: I think there's another sort of angle another lens that you could look at it from altruistic lines which is it's a bit like a trade union for individuals who generate data because individually they don't have any power to fight back against the corporations and individually their data is not that valuable. But when the collective power of a large number of individual data, it becomes, they get a lot more leverage and they also get a lot more value in their data.

INTERVIEWEE 2: Could I suggest that rather than see it as revenue for unions we talk about it as an eBay. A marketplace eBay where everyone is the buyer. You choose what bid to accept, the star rating on everyone. Everyone knows where they stand where currently today it's in the corporation's interest to not disclose that.

INTERVIEWER: Exactly right.

INTERVIEWEE 2: Maybe use the eBay example because everyone knows about them. Democratisation of auctions.

INTERVIEWER: That makes sense. In your experience, which departments, which sort of roles within banking, for example, do you think this sort of idea would be most appealing?

INTERVIEWEE 2: Absolutely, marketing. Marketing, absolutely. Marketing followed by strategy.

INTERVIEWER: This is probably more forward thinking too because I'm thinking along product development. So that companies actually develop a new product, if they have a better

sense of the micro segmentation in the market they can be a lot more agile, a lot more targeted with their product development.

INTERVIEWEE 2: I don't have a healthy view of product teams right now. For the last 5 years, I've been sincerely disappointed. It would be preferable for a top notch product team to appreciate micro segmentation and all that but a lot of them are supply driven. They're not demand driven. They leave it to marketing to tailor the proposition. It's a fraught exercise between marketing and product and I.T. is the meat in that sandwich. Product really have no idea what it's about, what sort of customer base they're aiming for, what they look like, they have ballpark figures but they don't even begin to understand the customer until it goes out to market. Unfortunate as it sounds it's how it's done right now. I don't think they're even close to being ready analytically. But marketing definitely is because that's the nature of marketing work, it's every advertising dollar has to be justified.

INTERVIEWER: I'm also talking to some of the spin doctor firms that work with political parties about this and I think they would love to be able to micro segment and tailor their message to voters of particular interest groups.

INTERVIEWEE 2: They have a heck of a lot more information today now than what people understand they do just because of the exemptions of the Stamp Act and so on. I'm involved with the Liberal party by the way. From my insider's point of view, there's also a possibility and this is one thing you might want to consider in your paper, it's the on-selling of personal data. And this comes back to identity theft, identity cloning or basically identity trading, the middle man. I worry about the handing of information to political parties because they have a rich body of information that not very well structured, it actually can be really sensitive because they carry between electoral cycles and there's nothing to stop an intern lifting that information or selling it to a secondary market pool. A lot of money through the system. Master data management will be central to your proposal because someone has to make sure that there isn't someone else duplicating that identity.

INTERVIEWER: That's the thing about my proposal is you don't sell the data to the company. They don't get the data. They get permission. What we're selling, the trade of the object here is permission to run a query against it and then take an action off that result. They don't get the data loaded onto their systems. In a way, it's fighting back against a shadow secondary market that you mentioned which I know is a major issue and I think there's a lot of legislation comes in around soon that's exactly what the problem that this proposal is trying to address in some way.

INTERVIEWEE 2: The thing is, I just don't know how the query, what the particular type of query language, but what it will look like. Basically it sounds like a whole new communication paradigm.

INTERVIEWER: I'll give you an example. So this one is very easy to understand. I think it illustrates it so actually it's in the paper, it's the one about the cafe. So the cafe is in Manly, they serve Mexican food and they play jazz and the sales are down on Tuesdays and Wednesdays. So basically what they want is a micro segment of people who live in Manly who like jazz, who use Facebook and who eat out more than 3 times a week. All of that appears in personal information. They put out a bit for that product, certain number of people accept for a fee, they can identify that micro segment they can then push ads for their cafe on Facebook to those people. They never know their names because it's all done through an intermediary that runs a query, identifies the people and then passes on the message. So they never get the data but they get access ... that would be the personal information product that identifies the micro segment and then permission to push market to them.

INTERVIEWEE 2: That requires the individual to exclusively store their data behind that file.

INTERVIEWER: So that would be their data stored if you look in that diagram.

INTERVIEWEE 2: The centralised data store.

INTERVIEWER: That's right. 3 and 4. And so the company, all the company sees is the dashboard that says yes you have purchased access to this segment, there's this many people in this segment and this many Facebook ads are being pushed to them and we can even show you the return like how many of them went in into the restaurant. So that's what the company sees. It doesn't say, it doesn't give a list of email addresses or names or anything. That's the sort of idea behind it. So it's permission you're purchasing not the actual file. Too abstract?

INTERVIEWEE 2: It's not that it's too abstract. It's very easy to convince an individual to bypass that. Very easy. because again information asymmetry they have no idea of the value of their personal data. So things like lucky draw prize, fill in your name, phone number, address, how many kids you have, income. Too easy to do. That's being done today by telcos and banks. Too easy to bypass this.

INTERVIEWER: So you're saying the main concern is that the company will bid too low and the individual has no idea of the real value of their data and they'll give it away for too cheaply?

INTERVIEWEE 2: Exactly.

INTERVIEWER: That's a very good point.

INTERVIEWEE 2: Which is why secondary markets will help a lot if only just to inform the individual by locking out his competitors, oh we're paying a lot, why I'm sharing on Facebook with CBA. Well, OK. Over time as individuals get educated it comes back to you have segment level permissions and presuming like segment level portal or something for individuals to understand what the peers are valuing it at. Then you can have a more informed community, a more informed market. This is what to aim for in an efficient market, an informed marketplace.

INTERVIEWER: The question I have is how enforceable would that be and would companies be willing to pay for it if they couldn't enforce it? So they say ok we want exclusive use of this data. Person says yes, they buy it and then how do they enforce the other guy doesn't just turn around and say the same to the next company.

INTERVIEWEE 2: That's the beauty of it, right? The person does not have the technical capacity to share it. The app is basically the means of which they can share it. It's coming back to information asymmetry. It means that the data that individual has is priced by the individuals as well as misused, have no idea what to do with it, how to sell it or how to describe it. The app makes it possible so the buyer can guess the security they need by understanding what the app reports, so it comes back to almost like a trust building relationship between the app builder and the buyer. Which is basically core to this idea anyway. So once the buyer has assurance that there are real safeguards in place like there must be some level of public access by the buyer, just to ensure you know, when it comes to auditing that there's definitely been no breaches, whether the company behind inadvertently disclosed data, all of that I guess will be covered in the contractual agreement between the buyer and the app developer. Then that's the point of control to ensure that they buy exclusivity they would guarantee exclusivity.

INTERVIEWER: Makes sense. Certainly something I'll take into account when developing this. Thank you. Alright. Any other comments or questions on that because I need to wrap up for this.

INTERVIEWEE 2: It's great, I also need to head off. Good luck. It's very ambitious and I would really like to meet you in person sometime.

INTERVIEWER: Yes, where are you based?

INTERVIEWEE 2: Sydney.

INTERVIEWER: You're in Sydney? Oh great. OK. We'll catch up sometime. Do you go to the big user groups or data meetups or anything?

INTERVIEWEE 2: I go to data science meetups.

INTERVIEWER: OK. I go for those too so we'll probably bump into each other there.

INTERVIEWEE 2: Will do. All the best, Interviewer.

INTERVIEWER: Thanks for your time.

Transcript 3 Interview with Interviewee 3

14.1.3 Transcript 3: Interview with Interviewee 3

INTERVIEWER: Did you get a chance to look at the paper that I...

INTERVIEWEE 3: Unfortunately not, I apologise.

INTERVIEWER: That's fine. So what I'll do is if you'll like I'll just quickly describe the concept that I've been working on and then we can have a discussion around that, that would be great. Basically what I've been researching is what I call a primary personal information marketplace and the idea is at the moment with wearable technology and mobile technology as we go about our daily lives we create enormous amounts of very valuable personal information. At the moment you can exchange that for free products online like Skype for example and Gmail and Facebook but there's no way or no really available way for you to exchange that for money essentially. My belief is that people are giving away their personal data far too cheaply. That it takes their worth a lot more than what they're getting in return for it at the moment. What I'm researching is the feasibility of creating a marketplace whereby we can facilitate people to exchange access to their personal data in exchange for financial compensation. The way it basically is we create apps on the phone and on the laptop that capture the personal data at source. As you create it, it's captured on your behalf and when you're on the internet, that's synchronised with your personal data store or a centralised repository. The whole idea is that the tradeable object on the market is not the data itself because there are very serious privacy concerns. I'm not proposing that individuals actually sell the data set and deliver the to the company. What I'm proposing is that they sell permission to query against that data. So that's actually the tradable object. It's the permission. The idea would be that the company has a need for the personal data product so it could be a list of people in a certain micro segment of the community. They post the product that they want upon the market and people who meet the criteria have the option of either opting in or opting out of their product. If they opt in, their data is being used to create their product and they're getting compensated for that. There'll be 2 dashboards; one from the individual's point of view, that can see what permissions they've sold to what company and what their data is being used for and when those permissions expire so there will be everything is opt in and everything has a grandfather clause on it too. So when you're buying permission, you're not buying permission ad infinitum.

INTERVIEWEE 3: How do you address the pricing problem?

INTERVIEWER: Yes, that's a very good question. What I'm doing at the moment is I'm now working my way through different aspects of this proposal. I'm writing a paper right now on the transaction object so exactly what do we mean by permissions, how are they defined, how they're defined in such a way that there's an almost infinite number of different ways you could conceive of permission to access personal data so how do we make that manageable. There's quite a bit of research already around pricing. First of all pricing access to databases. And also pricing personal data. So there's some research around how much each user is worth to Facebook for example. Or to Google. But also the way I'm seeing it is it would be an auction. An open auction so people would be able to see what has been bid in the past, what's being bid currently. To be honest, I haven't dug deeply into that. That's probably a topic for a separate paper.

INTERVIEWEE 3: Because you've mentioned that actually what's being traded here is the permission to query rather than the data itself. I presume that the pricing would in some way be dependent on the nature of that permission, right? You could presumably hypothesise the variety of different queries on your personal data. You could presumably price them differently.

INTERVIEWER: There's two ways that the pricing could be set. One would be the company says this is the personal data product that we seek and we are willing to pay this. I will give you an example it makes it more concrete. This is an example I wrote up as a hypothetical case study in the paper. I'm a restaurant owner in Manly. And business is slow on Wednesdays and Tuesdays. I sell Mexican food and I play jazz at my restaurant. The micro segment of people I'm looking for are adults who are from a certain socio economic group, who live in Manly, and the surrounding suburbs, who eat out more than once or twice a week and who are on Facebook. I'll explain why in a moment. That's quite a lot of conditions. I post that up, part of the data product is what I want to do with that data. Because permission is not just permission to run the query but also permission to take an action off it. The action I want to do off this is to push market, ads for my restaurant onto the person's Facebook page. I can do all this anonymously so if I have an intermediary which is the market that I am proposing, they can verify it's found the list of people who meet that criteria. I then pay for access to that list, I then create the ad and those ads are pushed to the person's Facebook page. It's quite non-invasive, those people are already using Facebook, they're already seeing ads, they just see slightly different ads. But from the restaurant's point of view they know they're targeting exactly the micro segment they want and if they also wanted to say have x query purchase access to credit card transactions or mobile phone data, again anonymised, they could track whether or not they're getting extra traffic because of their advertising campaign. There would be a price for that. We're willing to pay this many dollars a month and people will say it's worth it to me, basically I'm generating all that data anyway.

INTERVIEWEE 3: In your research have you come across any reasonable explanations as to why such a market doesn't yet exist?

INTERVIEWER: There are a number of reasons why. One, perhaps nobody's thought about it? Perhaps one's thought about it but done nothing about it.

INTERVIEWEE 3: That one I can guarantee's not the case. I can guarantee because I've thought about it, about the idea. Of course not the details you're talking about. Just the idea of people with whom I've been talking, the question is in the air, but I am not aware of anything. Your research seems to confirm this. I'm really curious as to why that's the case, I really don't know.

INTERVIEWER: One thing I've looked into was back in the 90s, just before the tech bubble crash, there was more than 20 different companies that tried to do something like this. They were called intermediaries, it was a whole sort of movement, kicked off by a book called *Net Worth* by Hagle and Singer. They all tried and they all failed. I read a paper that studied the different types of intermediaries and what they were offering and I analysed why they failed and it's my belief that technology has changed and people's attitude to privacy has changed in such a way that this is entirely feasible now. One of the reasons I think why it hasn't taken off is because of the, they're going to be major challenges as far as getting a critical mass, because this solution will only work if you have enough people on the market. People are only going to be on the market if there's an incentive for them to be there, right? That's the subject of another paper. How to initiate and to scale it. I have some ideas around it. Perhaps not going straight to introducing a personal information market but providing a platform where people can store and coordinate their personal data. Once you have enough critical mass, you then introduce this as an optional use for the data. That's another story. As to why it's never been done, it is a good question. People have dabbled with it, I've been researching and I keep a close eye on many start-ups and a lot of them are doing things around it but I guess my way of research is to find out how to make it feasible, and how to get it going.

INTERVIEWEE 3: I think your aim is fascinating. You're working on such an interesting problem, when you do something like that. I'm afraid I don't have much to contribute because I don't really have any experience in the literature and you've done a lot of reading papers and so on. I think it's more like, I was really curious to talk to you about this, because I just wanted to learn more about it. You were allocating half an hour of your time to teach me, really. That's a great opportunity for me to learn. I was curious to ask this question like I just asked you because to me, it's a really good question. Why such a thing doesn't exist. Of course there's a critical mass problem that you allude to. You could ask different questions which is why the big ones haven't done it themselves. You can hypothesise that maybe well they're already getting it for free, why should it be different?

INTERVIEWER: There's a good answer to that. In my opinion, and from what I've read, the answer is that Google gets their data for free, Facebook, they have to provide a product. Facebook gets their data and Twitter gets their data, nobody gets all the data. There's only one person who can join all that data and that's the individual. My idea is quite unique because it's putting the individual back in the centre where they belong because they created the data in the first place. So they can create that join. Joined personal data is exponentially more valuable than individual streams of personal data. In my work, one of the most common problems we get asked to solve is how can I identify individuals across these anonymous streams of data that our company collects. I'm sure you've come across examples as well. Only the individual can do that. The thing is Google and Facebook are unlikely to cooperate and start sharing their data because of interest in each other and they've got too much to lose and they want to beat each other. They don't want to cooperate whereas the individual has every incentive to make that join because it increases the value.

INTERVIEWEE 3: The individual probably needs a tool or mechanism to do that. And to develop that tool is probably not trivial, it may require a third party so how do you address that?

INTERVIEWER: That's what I'm proposing. I'm proposing that the tools that would allow them to join and store and combine and sell their multiple streams of personal data.

INTERVIEWEE 3: I see. Go ahead please.

INTERVIEWER: The sort of question I'm asking the people I'm interviewing, first of all is do you think it's completely crazy but also do you think this is something companies would actually find of value and would they spend money on it. For a market to work, you need both people to contribute their data and for companies to buy it. What's your thoughts on that?

INTERVIEWEE 3: The main concern regarding the particular question is lack of critical mass. Companies are typically, especially enterprise traditional enterprises, are very conservative. They tend not to jump into something's that's new by definition and that's why there isn't entrepreneurship, a market of start-ups and all that. I'm not sure, I'm sceptical in general about large corporations making large investments in, not this particular project, but any project that would rely on the existence of critical mass for it to actually work. It's similar to social network problem. How do you build a social network? That's how Google has thrived.

INTERVIEWER: I'll just clarify. My question isn't around how to help build it but suppose it was running would they be participating in the marketing, buying into access of people's personal data do you think?

INTERVIEWEE 3: I think the clever ones will. By all means, if I was the chief of a corporation, I would be very happy to do so. It depends who you're talking to. Also as you very astutely

pointed out previously, if you want to get this right, you need permission that's not only associated with the nature of the query that's being run. Although in certain applications make it pretty intuitive, which kinds of data will be useful for certain types of task. History has shown over and over that usually data is useful for a variety of tasks that you don't think it's useful for. My view would be therefore that it would be optimistic in that sense. There's a lot of data out there that would be potentially useful for a variety of tasks that you've never thought about. I'm not too well read of the literature of that sort of thing, but I did read quite a few papers two years on specifics of how you can for example identify personal traits of very intimate traits of people by simply looking at traits that are not actually sensitive and looking at which correlations between the non-sensitive traits would actually predict very well the existence of sensitive traits. I can certainly point to some papers for you. For example, your Facebook likes can predict very accurately whether your parents were divorced by age 21 or not. It can predict very accurately your sexual orientation, your political inclinations. Just from the Facebook likes without using any other contextual information or any textual information or anything else. Results like that have emerged over the past few years so in general my answer to your question would be yes if I was in a position to be able to play into this market, I would be very interested. My concern would be the pricing. You can't predict really the value that a given data would have for a given problem until you try it. It could be very interesting. The dynamics of pricing and of course if you're just using auctions you can deal with this in a more principled way. Just building a market for pricing but I would be very interested indeed. That's my personal view.

INTERVIEWER: That is one of things I'm struggling with because it's hard to do empirical research on something that doesn't exist - semi structured interviews with business owners, individuals and academics is the best you can really do to begin with.

INTERVIEWEE 3: There's one observation I'd like to make though, you alluded to, which is the fact of collaboration between organisations like Google and Facebook, it's hard for them to collaborate and to share data for very obvious reasons. I would like to point something you may want to look at which is, there is some recent research that looks into how you could basically address that problem. How Google and Facebook could potentially actually create positive sum game for both by sharing their data under suitable circumstances. I think you need to probably take a look at that to make sure that the claims you make in your research are well informed by that research. I will send you an email with the details. Just to make sure that you have as much of information coming from the other end of the spectrum, you are prepared to refine if required. If it doesn't really in anyway jeopardise your assumptions or anyway, at least you know why and are well informed.

INTERVIEWER: I'm very interested in that too at least be able to defend my position and be aware of alternative views.

INTERVIEWEE 3: I myself will be part of this research not very closely but certainly part of it. Some of this research is being developed at an organisation that I used to be part of before. I would be very happy to give you the references and also to point to the people with whom you could talk who are here in Sydney as well. I think it could potentially be useful.

INTERVIEWER: I just have one last question. We've talked a little bit about companies, assuming this market is up and running and companies benefit from it. The other big question is do you think there's enough people out there especially Gen Y and people who're basically brought up digitally who would be willing to sell access to their personal data in the sorts of anonymised ways. Do you think there's an appetite there to say well actually we've given away our personal data for free for too long?

INTERVIEWEE 3: You can split the question into two parts I think. One question would be Gen Y, do they care about money or not? Second, do they care about privacy or not? I think they're two different questions, entirely different questions but it's the interplay between how these two questions are answered that actually would be able to answer the question you just asked. For those people who do care about money, the answer's pretty simple. If they care about privacy, they won't be interested in the first place because they're not thinking about this type of questions. Let's assume that the young generation does care about money. Maybe a reasonable assumption for a significant proportion of them. I really don't know what the research says in terms of the true perception of how much release of individual information is damaging for these individuals if they have this association. Clearly many of them don't realise the potential ramifications of what they do, that's certainly clear but care about those ramifications at all. You can create some examples. Some teenagers may be wary of releasing pictures that may reveal what they won't their parents to be aware of, like they're drinking or having much more fun than they should on Monday afternoon when they should be at school whatever it is. It's not entirely clear to me the 2nd part of this question. I'm not sure, I can't give you a very well informed answer. My gut feel is that the more people are educated about the potential consequences on releasing their personal information the more the awareness will grow and the better they will be able to make a choice. This will point in the direction of probably aligning with your proposal in a positive way because the more people become aware of what they do, the side effects, the positive and negative of releasing their information, the more they will realise the value of information. The more they realise the value of information, they care about money, going back to the first question. It's a possible conclusion, they may actually enter the game. My gut feel but I'm unaware of research which supports this.

INTERVIEWER: That's all the questions I have unless you have anything you want to comment on?

INTERVIEWEE 3: First, thanks for hosting this. I think it's a great subject. I think you've made a great choice, something interesting to work on. They are really interesting questions, interesting problems and I think there's probably lots of challenges, probably most of them you've already stumbled upon. I think what's key here is that you obtain as much information as possible from the other side, the sides. What I mean is playing the devil's advocate in the sense that hypothesising every reason why and try to really hunt them and try to identify as many challenges as possible. You may stumble upon a given reason, right that actually you've overlooked and that explains why the thing hasn't taken off yet but still you'll be able to sort of maybe circumvent that issue by making changes in your proposal, you see what I mean? The whole business of research is about that. You start the question and certainly over time you change your question a little bit because you realise that your original question is irrelevant given the evidence you've acquired during the process of trying to answer the 1st question in the first place. That's my experience. In that context, I will be happy to share with you some of the work I've alluded to that tries to actually promote developed systems that make those who currently hold significant amount of data to collaborate between themselves. The competition in a way with what you are proposing. The more you are informed about that of the potential competition, probably the more credible will be your claims and the conclusion that you come up with.

INTERVIEWER: I've already come across a few alternative views. There's one quite strong movement in a sort of start-up in the Valley at the moment which is that putting any limits on personal data is a lost cause. The whole movement is to open data, it's an unstoppable movement and you'll be fighting against history if you do that. There's more value in having everything open. It's different alternative views but I agree with you. I do need to explore ideas by playing devil's advocate and studying them.

INTERVIEWEE 3: I should really congratulate you on the courage to address this problem. Because it's a truly important question and these days when people pick their PhD topics, this is actually a very, very good question and you are not afraid to address it. I think this is absolutely fantastic. Please keep me updated of your progress. At some point I'd be happy to meet you in person as well whenever suitable.

INTERVIEWER: Thanks very much for your time.

14.2 Appendix 2

Selected transcripts of interviews regarding companies' appetites to participate on a PPIM conceived as an four-sided industry platform.

14.2.1 Transcript 4: Interview with Interviewee 4

INTERVIEWER: Have I already shown you this diagram?

INTERVIEWEE 4: No.

INTERVIEWER: So you're already familiar with the concept because you already did the other survey, didn't you? You did the one with where we surveyed which individuals would allow their data to be collected?

INTERVIEWEE 4: Absolutely.

INTERVIEWER: That's cool. So this is the project, right? That's known as a primary personal information marketplace (PPIM) and our concept is as you know, we're generating all these valuable data but there's no way you can get money for it at the moment.

INTERVIEWEE 4: Yes.

INTERVIEWER: Alright. So the way it works is we build an app that you put on your phone and on your laptop, and on your tablet, it collects all your data on your behalf when you're online. It synchronises it into a centralised database and you have a dashboard through which you receive offers to access your data. So what's happening is, what's being traded on the market is not the data itself, no one ever gets it but what's being traded is permission to use the data. That's what people are paying for and that's what being bought and sold.

INTERVIEWEE 4: So I own the data?

INTERVIEWER: You own the data. They use it for a specific purpose for a specific period of time. That's the thing that's actually being bought and sold. Now that's quite normal. From a company's point of view what happens is they have a dashboard –

INTERVIEWEE 4: Which company is this?

INTERVIEWER: Any company. So in this situation, that's you.

INTERVIEWEE 4: That's me.

INTERVIEWER: You have a dashboard, you can look around the personal data store without seeing the individual's name and addresses but you can browse and find what data's available, you can then create a data product that you want, or you can buy a premade one. And when you create it, then the individuals who accept your offer, their data goes into it, then you can use it. So that's the overall view. So I've already done this dashboard here and I've got 204 people to accept offers, I've simulated a whole lot of offers and 85% of the individuals who participated accepted one or more offers.

INTERVIEWEE 4: So when you say you simulated 204 people, not just making up data here, you've actually got 204 people?

INTERVIEWER: In the previous research paper, I've got real people but then I've made up offers from companies so I researched the literature and understood what would companies want to do with different people's data and then I say alright, here's your first offer, in fact I could show it to you, this is what one looks like. So this is the paper that summarises all of that research. Here's an example of an offer, okay.

This company is a charity that supports disadvantaged children. It requests permission to run queries against your WhatsApp activity, collected from your mobile phone so they can deliver more timely and relevant information to you. It would like to send you relevant messages via your phone, they'll pay you \$2 a month to accept the offer. And then you either accept or you reject.

INTERVIEWEE 4: So then, and this is probably not – I'm sort of granting them access to my phone? Or is my data still somewhere else or they have access to that?

INTERVIEWER: Yes. What happens is they definitely don't have access to your phone. They have data that's sitting in here but they can only use it for what they paid to use it for. I'll show you how that works. The bottom line is from the previous paper. We know that quite a lot of people are happy their data are put on the market.

INTERVIEWEE 4: The first step is you can have my data put on your database and I'm not going to get paid until someone actually buys that data.

INTERVIEWER: Exactly. Until someone uses it, you're not going to get any money. What we're talking about today though is the other side. Real companies, would they actually want to use data that's in there? So what I'm going to do is demonstrate this dashboard here and ...OK so this is what the dashboard looks like. What I can do is browse around it and look at prepackaged products and I can also create my own and at the moment, there's all types of products you could make. I've only built two for the prototype. One is the targeted marketing product and another one is a personalised service market product. But there's lots of other ones. Once you've bought your product, you can manage them and delete them and use them. So let's start by browsing. So we go into browse and we can see that there is four targeted marketing products I've made for you. This one over here is called Over 55s in the northern beaches. Anyway so I view it to find out what the story is. There's 281 people in this product. I can see approximately where they live. You get the idea, you can see how much it's going to cost you, \$320 a month. You can see the parameters that I've used. They're 55 to 72 years old. The keywords they've been searching, AMP, so they've been browsing AMP website and they've been searching for SMSF

and that's only the one we've specified to build it. If I wanted to then say – what's the mix between male female or also other parameters, I would click on this and it would fill them out.

INTERVIEWEE 4: \$320 a month. I could turn it on or off each month? So do I get for \$320 a database with 280 people who were looking for?

INTERVIEWER: No. What you do is, say you want to purchase them. What you get is you get to use this data product 20 times. That means you can send 20 marketing messages to that group of people. And so you say, right, I want to use it now I've bought it. Now I'm going to open it up to look at it and I'm going to use it. So I open it up. Select a file to send to the individuals. So click through here. This is my brochure, what's my new offer for this month. 10% off, whatever it is, whatever the communication is and then I execute. And it says, your data product is being executed. The message has been sent to the 281 individuals. But the system has sent it to that specific niche market that you have identified. Click OK and now if you go back and look at Manage, you'll see that the number of uses have gone down by one.

INTERVIEWEE 4: Obviously, a very effective product, it all comes down to database and the rest of it. I guess if you're taking it from a web browser you're going to get people who're accidentally there, people who're doing assignments all those sort of things. You never guarantee it. I guess you've got no track record which makes it harder.

INTERVIEWER: I've talked you through on how to create your own month. You might look at it and you go yeah...it's alright but it's not quite what I want.

INTERVIEWER: Right. So let's go and create a new one.

INTERVIEWEE 4: So do you want information about how I use this for my business specifically or generally for businesses?

INTERVIEWER: So what I want to know from you is would you use it? Is it relevant? Secondly, how it could be improved, make it more useful. Thirdly any issues you can see that would prevent not solely yourself but other companies using it and other ways it could be improved to make it even more useful for yourself and other companies as well. Essentially the topic of this research paper is to try and get a feel for the appetite for various companies to use such a market. Not necessarily in the same exact format that's prototyped here. What I'm doing is I've got a matrix of different types of companies from different sectors and different sizes. So you're small to medium, I assume, and then I've interviewed the guy from a large bank. His issue was, 'It's good but how am I going to integrate it into my CRM system?' That would be a big issue for him.

INTERVIEWEE 4: Whereas for us, it wouldn't matter as much.

INTERVIEWER: Let's create a new product. We'll call it Mortgages. So now we start building it up. Do we care about the gender of the person?

INTERVIEWEE 4: No.

INTERVIEWER: Do we care about the age?

INTERVIEWEE 4: Well, it's all going to be self defining, if you're looking for a mortgage, chances are you're between 25 and 35 years old, but you've also got people who're going to refinance so you really want the full range.

INTERVIEWER: Plus this would be the way to filter out the students who are researching what mortgages are. We can bump it up to 30 say whatever, it's done. Also just a note, when we're building it, the next step is to preview it and that will tell us how many people are in the segment and we can always go back and adjust it. And that's something I've got off feedback from other interviewees. They would like to see how many people there are so far as we go. And the price.

OK. Location. It's obviously pretty important.

INTERVIEWEE 4: Mortgage you can go statewide if you want.

INTERVIEWER: I'll put in Palm Beach. I quite like this. Yeah, that's right. If I go all the way down there, it will capture all of the Northern Peninsula. Relationship status, do I care about that? Not really. Type of customer, so this is key. I don't want to contact my existing customers.

INTERVIEWEE 4: Then we say the income to be over, let's just say \$150K.

INTERVIEWER: And then this is key, right? We don't want existing customers or we waste our money.

INTERVIEWEE 4: So when you say existing customers...

INTERVIEWER: If you only want to contact prospects, then you have to attach a file of existing customers with email addresses, which is going to be used to exclude people in the database so you don't waste money spending.

INTERVIEWEE 4: We don't have to do that because we send out our own information, we don't want to double up that. We can do that.

INTERVIEWER: And now we have Preferences. Companies or product liked by the individual's Facebook. Do you care about that? Maybe? Maybe not?

INTERVIEWEE 4: Not really.

INTERVIEWER: Browsing history? Yep. Care about that.

INTERVIEWEE 4: So a lot of database is going to come from people's browsing history, is that right? A lot of data is going to be collected from there?

INTERVIEWER: We're collecting everything, everything they want to collect. Yes, browsing, email, SMS, location data on your phone, where you go on holidays, all of that. Because all of that can be used to create data products. Different types of companies care about different things. Things they've been searching for, maybe they've been searching for loans, interest rate. You can refine it further, for example did they use a borrowing calculator.

INTERVIEWEE 4: Could you set up something like ... what we're trying to create here is almost like an algorithm to find the right people for mortgages. Could you set up something where instead of me trying to build it, a template or anything, have some common ones for financial planning and mortgage investments, you can get a list of products that comes up with, a list of suggestions then you go and select them. You try different things and you don't think of things that you should be thinking of. Something's that templated out.

INTERVIEWER: From what you're suggesting, it's sort of halfway between this and browsing. When you're browsing for the product, you just select a pre-built one and what you're suggesting is something like a wizard almost that helps you create...

INTERVIEWEE 4: Customise it, it gives you all the data and you customise it.

INTERVIEWER: Very good idea. Here's another really cool little feature, do you want to filter out people who don't know anyone in your customer base for example. Because we've got all the connections on Linked and Facebook. I know that most of my businesses are based on referrals. Let's see how many people, Let's say well, I want people who are no more than five degrees away from my existing customer.

INTERVIEWEE 4: Like a family tree.

INTERVIEWER: So you might say, yep, that's what I want ... and then we've got some really nice little predictive attributes, we built propensity models in the database and we've scored the customers about, and how likely they are to stay with you. You want propensity to churn well. I want people who're going to stick with me. I don't want people who are complaining all the time. Takes me too long to onboard them or maybe I don't.

INTERVIEWEE 4: Well, mainly you don't, especially mortgages and again this templated thing, a lot of information we're trying to get in here. You want people to return again and again because they come to you ... oh yeah great idea ... I wouldn't have thought of putting that in unless I was in the template.

INTERVIEWER: And then down here you select how you want to contact them. It might be you're never going to contact them. You just want to analyse their data, right? But it could be I

want to email them. And then once you've done all of that. You get the concept, right? And then I preview. I've found 42 people that fit my specific requirements. Yep, it's only going to cost me \$60 a month and I get to contact them 20 times by email and I know that's the area. Zoom in, double check that.

INTERVIEWEE 4: Whichever that finishes first, a month or you use up 20.

INTERVIEWER: That's another thing I need to clarify. What I'll do is grant a set number of contacts or you can have a time period.

So we purchase that one and now in here is ... let me see ... Mortgages is the one we've just made. Now we can view it and use it. So that's one type of data product. Another one is our personalised data response. The concept here is a new prospect rings up. Rather than you have to go through a long, time consuming process asking them about their circumstances, getting to know them, you can use their pre-existing personal information to know more about them straightaway. And that way you can give them more personalised service very quickly.

INTERVIEWEE 4: That would work. I'm just trying to think. When you have the personalised information, address or email, their phone numbers.

INTERVIEWER: You don't get to keep their data you get to use that on the screen, it's not downloadable, it's not capturable.

INTERVIEWEE 4: So you just have to write it down. The purpose of doing that would be to save time to prepopulate forms, so when you sit in front of a client you've got the address, name, telephone number, email address, key in all their details as well, it's all prepopulated.

Let's talk about all of your objectives and the rest of it, what you want to do, that's where the value would come rather than saying, we can have a look at where you're living and most of it. I'm just going to write that down on screen, I'm better off sitting down with them ... I'll write this down.

INTERVIEWER: This product is more for companies who have a wide range of products and they try to treat each customer more individually and also who've got high turnover customers. Like a call center for example.

INTERVIEWEE 4: So they just get the data with a note, they just ring them.

INTERVIEWER: Or the person's rung them. Let's get one up. So this is personalised service. So what we've done here. We've chosen the fields you want to see when the person contacts us and then the person rings up and we know that this person, whoever they are. We've got things like they're divorced, they've just moved into their first job, what are some of their issues, timelines, we've analysed the text. These are the things they've been searching for recently.

They've got one child, their postcode, they own their house, recent employment history, this is their income, this is their relationship status, their age. So rather than having to ask them that, extract that from your database, you've got it there. Especially if it's a prospect. You can see that it doesn't have the contact details apart from what you have entered.

INTERVIEWEE 4: Just so I understand how this works. You've got someone off the street, walking to the office, they've already got their data sitting in this database. They come in and they say, ok what do we want to talk about? We can go to this, we probably get the information there and then from them. Rather than going to this I guess the thing is they will have some of the information, not all of the information we need as well. This is very business centric to us, I don't think we use some of that, just by nature of the business.

INTERVIEWER: The scenario that other companies would typically use, somebody who rings up and says ok, I'm looking for a financial advisor. I've been talking to AMP, I've been talking to BNZ, and I'd like to find out a little bit about you, how you can help me. If you have the person's information in front of you, you'll be able to do much a better pitch to them, than you could, if you didn't have this in front of you. You could pay the fee because you possess valuable information about that person. The way you describe your services, what you emphasise. The fact that they're 28 and not 58 for example.

INTERVIEWEE 4: So they have to have registered for your database for you to get this information. They're not on your database, you can't get it.

INTERVIEWER: No. The way it works is, you want to create a type of personalised response data product so when anyone contacts us, we can find them. You can access this information. Let's go through the process. Everyone who agrees to be a part of it, would be in it.

INTERVIEWEE 4: If they're not part of that database, someone walks in who's not part of it, I wouldn't know all that.

INTERVIEWER: No. Just to help us, can you give us your email address and you put it in and not found. Nothing would come up.

INTERVIEWEE 4: The database, find it through their email address. Ok.

INTERVIEWER: Alright. So they're the two that are prototyped. And there's the other ones that we're going to work on personalised recommendations. So you sell a lot of different products, you don't know which ones to recommend to each customer. You can infer that by understanding more about the person to recommend the right one hopefully. Personalised content delivery. You've got a big customer base and you're doing marketing to your own customers, you want to personalise the message right. What sort of person they are, how they receive it. From other interviews, I've got suggestions around you're trying to grow your customer base. Step one is

to understand your existing customers and segment. There's a group of your customers here who really like us for these reasons. Can we go and find similar customers on the prospect database? Personalised education, that's a whole different ballgame. There's another that was suggested to me in an interview which is around obtaining verification so this is where businesses are conducting business almost anonymously. It's something like eBay, where you don't get to meet the customer but you have to verify in some way they're legitimate, using this as a way of doing that.

INTERVIEWEE 4: I guess the proof would be in using it two or three times and see exactly what results you get from it. In that case it's how good is your data. Like the first one, it would be brilliant to use. You often get where people will sell you databases and it's actually hard to get the right information out of it. The key people would be the key driver of this one. Because if I could find over 55s on the northern beaches and I'm able to send them an email, they're part of a database that they've registered for so I know exactly where its coming from. If I could sell a mortgage over \$700,000 or more ... It's worth more than \$320, without a doubt. It's incredibly powerful. That's the bit you have to work on, you know. I've got all these data, you've got to extract the right stuff out. You probably find it, you might use it, two or three times and if you don't get results from it ... no that's right. So that's the key on that one. Incredibly powerful tool to have, especially if you break it down to wage and age and what they're looking at and all that sort of thing, particular area. Incredible. The other one I guess depends on the size of your database.

INTERVIEWER: Absolutely, that's a key point. That's another issue but basically what I'm trying to do in this paper is assuming that it's large enough and it works as described on the packaging, what's the appetite like? But another really interesting ... that's great feedback. Another really interesting point that I found in my research was just then, right even before you purchase, just to be able to do that would be worth money to the company. So if you were looking at any new offers, you open that up and, to be able to do that, that's how many individuals in that area have been searching for that. That doesn't involve you contacting anyone. Simply involves you using their data to answer questions that you've been trying to answer about where to open an office for example.

So next question is assuming you've got this massive database of individual information, are there other products, other data products that would be worth creating for your business?

INTERVIEWEE 4: Good question. In a different sort of space, I guess. We sometimes have issue in looking for staff in the Northern beaches. It could be easier to pull off LinkedIn and all the rest of it, advisors who are looking for work, advisors who are looking for a change of role, advisors who are looking for a new role and all that sort of thing. You can actually get databases

of people like that. People pay very good money for those databases of recruitment agencies. So that would be handy for us.

INTERVIEWER: That's an awesome use case, really targeted recruitment basically talent acquisition.

INTERVIEWEE 4: That would be very handy if you could do that. If you can get the data and clean it, make it so that it works but again its the thing, you use two or three times doesn't work. It's not very time consuming, you make it where you go in there and it's all templated. Even looking at that personalised data you have there, there's like some sort of information coming through. That's way too hard. Summarise beautifully and that's all I need. You go to Google and you just view the top 5 choices. The algorithm works so well. It's getting the algorithm right. The first one you've got there. Just brilliant.

INTERVIEWER: The big picture of what I really want to do is create this marketplace where you've got APIs facing out three ways, you've got people who create a new data feed, coming into the database so not just download but creating, anybody who wants to create a data feed can feed into it can. Then you've got modellers who can create models to extract information out of it and you've got companies who want to create products out of this. So you're creating a marketplace which facilitates the raw materials which is the data. The extraction of value out of the raw material, the predictions and then the consumers who are looking to buying this.

INTERVIEWEE 4: So you might have a set algorithm?

INTERVIEWER: When we go and create one of these, right, each one of these little boxes. That's a little way to extract something out of the data. Right here we're going to extract the gender, here we're going to extract the age and so what we can do, we can allow other modellers to build new options further down.

INTERVIEWEE 4: Right. So what you've got there are fixed algorithms set for the rest of it and you're basically...

INTERVIEWER: It would be up to the individual modeller to say, actually look, I can extract from that data what the next job that this person really wants is this. You click on it and I add it to your product, I'll get a small amount and then you use it. The person who does fit into it gets some money, that's me and you get the –

INTERVIEWEE 4: that's important.

INTERVIEWER: That would then make it great because each company in industry have their own specific requirements and it would be impossible to build all the right ones for everybody. People with knowledge of the industry would be building those.

INTERVIEWEE 4: You put one of your programs here you can, you can view one from your database. You just pay for the time to use it, that's an income stream for you. You understand the building blocks underneath it, but to me, that's something so obvious. I just don't know where to start. Here's the thing. People have been searching for this sort of thing forever. Targeted marketing is what it is.

INTERVIEWER: That's the other thing about this concept. See there are the data sources right that are feeding into it. Google would be able to tick that. And Facebook can tick that one. But only you, the individual can tick more than one because you're the one who actually controls the data and that's where the richness comes from.

INTERVIEWER : Alright, that's really good feedback. It's really fascinating, though taking suggestions out of people from different backgrounds because each person sees it differently. Nobody has said it's rubbish. Very different. And then they also have issues with it which are very different, which is fantastic. That's what I'm really enjoying is just getting different perspectives on it.

INTERVIEWEE 4: Well it's going to become a generic tool, very specific which is not an easy thing to do. It can't be too complicated, it can't be too long or drawn out, that sort of stuff.

INTERVIEWER: Appreciate that. Any final comment or questions? Otherwise –

INTERVIEWEE 4: I'll have a think about it. How many people have you interviewed on the phone so far?

INTERVIEWER: You're the third one so far.

14.2.2 Transcript 5: Interview with Interviewee 5

INTERVIEWER: This is what my research is all about. Basically it's called a primary personal information marketplace. The whole concept is, because you wander around with your smartphone in your pocket, you're creating really valuable personal information. You don't have to try and do it, you can't help doing it basically. At the moment, companies all around the world are collecting their data and they're using it for their benefit. We give their data away in exchange for free products like Gmail or Skype but at the moment there's no way you can exchange it for money. My research is around can we create a marketplace where that data is captured on behalf of the individual and the individual can control who has access to it, have visibility on what's being done with that data. Currently there is a thriving secondary market in ...

INTERVIEWEE 5: What do you mean by secondary?

INTERVIEWER: So that's where companies collect your data and then sell it to other companies. There's no primary market where the primary user, that's you, can exchange –

INTERVIEWEE 5: I guess individually we're too little, aren't we?

INTERVIEWER: That's one of the analogies I've used is, it's a little bit like a trade union. Individually you don't have the power. So this is a diagram that sums up what I'm trying to create, so basically we follow the numbers. Basically your data is collected at source. It's collected on your behalf, and when you're online, it's synchronised to a centralised data repository. And then what happens is you've got a customer-facing dashboard which faces the individual and that allows you to see what data is being collected. Then come over here to companies, they can browse around that data, it's anonymised. They can create data products, if they want to purchase access to. The whole concept here is you're not buying or selling the actual data, you're buying selling permission to use the data. So the actual commodity that's being traded is the permission.

INTERVIEWEE 5: I guess the issue there would be anonymised data?

INTERVIEWER: Yes.

INTERVIEWEE 5: If I wanted to create a model, I would want to be able to match my existing customers to the customers in this virtual hub to then create the model. I don't expect to be able to see it, who I matched, against customers that were matched? So how would you see that sort of working?

INTERVIEWER: So what I would do is take you through a demo that I've built. I built a prototype that allows the company to look at the dashboard and create two types of product at the moment so what I want to do is demo it to you now. Get your feedback on it. Shortcomings of the system, objections and other types of data products that could be created.

INTERVIEWEE 5: I guess this data that's being collected, if you think about the smartphone which is what we've started with, Apple already knows where my phone is but then we're also going to have this and this. So then, who owns it? Does Apple own it or does that become my information?

INTERVIEWER: So that's part of my research. This is a really big concept – there are many aspects to the problem. One of it is data ownership. And so as part of my research, I've got the end user license agreement of all the major social networks. They all explicitly say you own the data and that you're giving us the right to use it. In fact, you do actually own the data and what's more if we collect it before it gets ingested into another system, so that's definitely the case, you're going to own the data. Also Google may have access to your Google data. What they don't have is your Facebook data, your Twitter data. There's only one entity –

INTERVIEWEE 5: Everyone's only owning a little bit of it

INTERVIEWER: You're the single point, you can join all of them together. As we know, joined data is more valuable than disparate sources. This is the concept here is that it will be an app on your iPhone. So what I've done in a previous paper is research are people willing for their data to be collected. Around 80% of people agree so I actually manufactured these fictitious data products that companies are interested in purchasing and they have to agree/disagree based on a certain price and 80% people agree to participate in some way. It gave me enough confidence to say, well it's not ludicrous. Would companies be willing to buy access to the data. So that's the subject of today.

This is the prototype. So you're a company, you come in here and you can browse created data products, you can create them and you can create targeted marketing or personalised service response. And then you can manage the ones you bought, it's basically how it works. So we'll start by creating one. We want to create a targeted marketing product. We'll give it a name, let's say we're interested in this group of people. Now I start selecting the characteristic of the need I want to identify. Should I care about the gender? No. Care about their age? Probably. Remember the more people in it, the more it's going to cost. Location, very important. That goes to Brisbane, I input zero. I quite like this one a bit, properly visualise it. Relationship status, maybe not. Things they like on Facebook. Recently they've been browsing around. Turn it on. And then I can go through some more sophisticated matrix, these are ideas about if we did have an extensive repository of personal data. I'm interested in people who are connected to existing customer, may not be relevant in this situation, maybe.

INTERVIEWEE 5: What do you mean by existing customer?

INTERVIEWER: Say you've got an existing customer...

INTERVIEWEE 5: How will your system know that these are existing customers?

INTERVIEWER: They're using online banking.

INTERVIEWEE 5: How would I get that information? Because they've got the app on their phone and they've used the app before. Fair enough.

INTERVIEWER: And then we can have some propensity scores as well and we can calculate things like propensity to churn etc.

And then once you've got that niche of customers, what you want to do, how you want to contact them. So let's say we want to SMS them, OK? Remember that people will only be included in this product if they only agree if they want to be part of that product. Preview it.

INTERVIEWEE 5: Which is going to have its own buyers to it, anyway.

INTERVIEWER: Ok, so we've got 56. So what we're doing is we have 56 people that meet our criteria, they're willing to be SMSed and we want to purchase it or not. Let's say we do.

INTERVIEWEE 5: Do we have enough information there to purchase at that point?

INTERVIEWER: Ok, so if we go back ...

INTERVIEWEE 5: Would I be able to see a profile of those customers that were selected? You need some confidence in terms of who these people really are, what is their information that we know about them. I don't know what you are actually buying sorry, if you have like a different profile of these other things then maybe you could add things like to say, yeah, they would also like to return all those information, as opposed to this.

INTERVIEWER: That's a really good point actually because the way it works at the moment, the way you review the data product you only see the information on the parameters you specify.

INTERVIEWEE 5: I'll say, just give me a profile of the others and then if occupation was something that was really valuable, then you could like, add that, even though you're not restricting on that, then you purchase that information, you can add this for \$10. 'Would you like fries with that?' type of thing.

INTERVIEWER: That's a really good point. I'll capture that. What I'm doing with these interviews is capturing people's suggestions about how to enhance it.

Alright, so we've bought our product and now we want to use it. What we've purchased is permission to send this group of people a message. So we connect the file that contains the message that we want to send them. A new online offer for first time buyers. The previous guy I interviewed really needed to preview how it's going to appear.

INTERVIEWEE 5: Well it's more than just on that phone itself because it's Android vs this, vs that, you know, because they all render differently as well.

INTERVIEWER: That's something missing in my prototype. In this case the information is plain text. Then we execute it. Your message has been sent to these individuals. You click OK and then transfer it to your management page and you can see the number of uses have been reduced by 1. You bought the permission to SMS them 20 times. You just used up one of those.

INTERVIEWEE 5: Why would I want to do that?

INTERVIEWER: Right, so you might only want to purchase 1 SMS. That's fine, you can do that as well.

INTERVIEWEE 5: Like next month, are you going to give me the ability to re-purchase the people who I've already contacted?

INTERVIEWER: Yes, they may have opted out, in which case there'll be less people in the product.

INTERVIEWEE 5: Actual different set of 20 people?

INTERVIEWER: Yeah, and that's actually something else I have to include in this is, the ability to union and intersect your existing products, because you don't want to keep spamming the same people. I've already captured that.

So that's personal targeted marketing. Any other feedback on that?

INTERVIEWEE 5: I've got a lot of ideas on what you could do with it. I'm from a banking point of view, so you've got the coordinates of these devices and if you're smart, you would know where they live because that device would be at a house for twelve hours. So effectively you can get a detailed ID for that x,y coordinate. You can do a bit of location, who these people are, where they live then you would potentially know where they work. From a banking context, I might want to know people who bank with us or bank with CBA. You can probably do that via your app thing or if I was really, really smart, I'd get all the x,y coordinates of all the CBA branches and I would say who is going to these particular things.. Now what if you're trying to find a home? What if you're trying to buy a house? If I knew all of the houses that were open for inspection, and I know their x,y coordinates, and you went there, between 10 o'clock and 10:30 when the open house is on, I know you're in the market for a new house. What if you're trying to buy new car? If you knew the x,y coordinates of car yards and you're always going to car yards, maybe I should be offering you a personal loan.

INTERVIEWER: Plus the searching of cars on the internet.

INTERVIEWEE 5: Because you already know the x,y coordinates where anyone is at any given time, it's limitless, this sort of thing that you could do. How would I upload the information that I was interested in to be able to do that match for them.

INTERVIEWER: The overall concept of this is – I've created two of these products. What I really want to do is create a marketplace where developers can develop new data feeds. We need an API for that. We also need an API for people who want to build models and then turn into information of value. They can create a model and place it in the marketplace.

INTERVIEWEE 5: That's what I'd be interested in as well. I actually have no real problem with my data being anonymised in your system, as long it wasn't returned to somebody else. You can match on my information, I don't really care. It's a little bit like the census data, right, unless you've got more than 3 houses in a particular area, to not being able to expose someone, otherwise we're just saying, that's ok.

INTERVIEWER: You can see from here it's pretty anonymous. You don't know them. You don't know people's names etc. You know what you need to know about them in order to market to them.

So the other one was the other type of product that I've prototyped was personalised service. The concept here is you're in a call centre. Either a prospect or an existing customer phones up, rather than have to go through a whole process of asking them questions or the whole sort of thing, you can use the information that they've given permission to use, to know all about them, on a single screen without waiting time.

So if we, let's say, we go browsing for a personalised service product.

So here's one, personalised details relevant to banking, a personalised service response. OK, so I view that, and what I've got here is, I want to know the person's main issue, when they get dealt with, what do they care about. Timeliness, politeness, whatever. That's being inferred from their correspondence. I want to know about their life events, things that are impossible, things that are forthcoming, things that are recent. I want to see some recent search terms, number of children, their accommodation type, their home postcode, their work postcode, their employment history, the sorts of stuff that are really relevant. So then somebody rings up and they give you their email address for example then rather than have to work all of that out, or ask them about that, I can refer to the screen and I can get that information.

INTERVIEWEE 5: For a big organisation, that would be an API call.

INTERVIEWER: That's right, exactly. Especially for prospects.

INTERVIEWEE 5: There's a lot of text there, you use things like word maps, you know. The things that they search for most are the biggest text in the middle and the smaller ones on the outside. Might make that a little nicer. These things are important to us.

INTERVIEWER: Of course you can tailor this. When you create it, let's create a personalised service response, you select what you want to show up. Because a lot of stuff would be irrelevant, for example ethnic background, might be totally relevant. Age, maybe gender ...

INTERVIEWEE 5: One of the things with prospects is you won't actually know what might be relevant.

INTERVIEWER: Correct. Or if you're selling a subscription to a magazine that would be relevant. So basically you build up an interface that allows you to create a dashboard that gives you the information you want, either prospect or customer comes up, put in there what you do know about them, and then it returns relevant information.

INTERVIEWEE 5: So let's say I signed up and download your app, are you going to pay me one off or every time someone selects me or once a month? You figured out the commercial model for the people to be compensated?

INTERVIEWER: When I did the research for the individuals, I did it by individual request. So let's say the bank creates a dashboard right, purchase it, and you get a request on your phone that says the bank would like to purchase access to your information. One way we might do it, we might say select your privacy setting; High privacy, medium or low. And then you'll be put in a group where all these requests that fit into that High privacy bucket. And then you get a monthly statement that shows you the company, the data they're accessing, and the usage and you get your payment. Also there'll be a – just as we are managing our data products here, these are the companies managing the data product, there'll be a similar dashboard for the individual. So they can logon, they can manage –

INTERVIEWEE 5: I don't want that company using my information, I want to exclude certain types of companies.

INTERVIEWER: Yes. For example you could exclude anything to do with gambling, you can also do that when you setup your profile. Pornography no, sex work no, gambling no, military no, banking yes, telecom yes, retail fashion yes,

INTERVIEWEE 5: It all depends on the numbers, there will be a point where you reach critical mass and you would need a few people to get on.

INTERVIEWER: That's another major hurdle is getting to critical mass. People won't join up, if they're not going to get paid and people won't pay if there's not enough people inside. That's another area of research to look into. Probably one of my ideas is to start off as a free service where it collects your data on your behalf, you can understand about your own behaviour, quantified self maybe but also lets you compare yourself to your peers anonymously. You might say well, I went through university, where are they now, employment-wise, wealth-wise. If it was compelling enough and enough people joined, I guess one of the big questions is out of these ones right, we've got targeted marketing, we've got the personalised service. The previous person I interviewed, he said another really good personal data product would be a verification component. When you're doing business online, and you need to be able to check that person to verify their identity using their personal information.

INTERVIEWEE 5: I think we're trying something like voice signatures at the moment. I can imagine it would be quite fancy. If you could know what device they were calling from, it would be a fairly high accuracy. You could be pretty sure they are who they say they are.

INTERVIEWER: And if you ask them some questions, something to compare their answers to. These are also some other ones in the pipeline; personalised recommendations. Based on your profile, based on profiles of similar people, much, much more targeted because we know much more about you. Personalised content delivery so we're sending out advertising or marketing messages, we can tailor it to you. Do you have other types of personalised data products you'd like to use?

INTERVIEWEE 5: We don't currently do a lot with prospects. We've got a large number of customers already, we don't actually do a very good job of understanding our existing, the way that they behave and things like that. So there is a lot of growth in their existing customer base rather than running out and try to find new ones. Saying that, if you had a mechanism to better identify customers that are likely to switch. You sort of want profitable customers. I want customers to have a mortgage with us, want customers who are in occupations with room to grow, long-term value. Anyone can get customers but are the customers the right sort of customers. I do like finding out more about the customer. I know we can do more of that. This is very much focused on individual people, massive opportunity for the small and medium enterprise people as well, traders, builders, professionals; I'm just trying to understand their behaviour as well. Completely different set of things you can do with B2B type of information. Where do the painters buy their paint from? Could be anything. From a bank point of view – So we're trying to bring a number of data sources together.

INTERVIEWER: That's right. I guess the differentiator here is there really is one entity that is uniquely placed to join the disparate sources of data, of course they actually own the data because they generated it. They are the ones who can facilitate what's really fine grain view of the customer from 360 degrees. If you're at the centre of the circle, you actually have that 360 degree view. There's definitely no working market at the moment where you can exchange your data for money. For example you can give away data for discounts but it's a lot less comprehensive than what I'm researching.

INTERVIEWEE 5: There are a lot of companies that try finding customers for surveys.

INTERVIEWER: Another thing here would be recruitment, even if you were a very small company start-up with three people, they know they need a PhD and he or she lives in Sydney, he earns that much, boom boom boom, you'd be cutting out the recruiting firm and go directly to the person. And that person would be delighted, here's everything you need to know about me.

INTERVIEWEE 5: It's very interesting. That's a lot of information to give away for free, in a way. Like knowing there are 281 individuals within the radius who you're searching for. You need to give this, maybe a sale, I just don't want to give away.

INTERVIEWER: So what you're saying is the way that you present the data product has to maintain the value of it. That could be another product where you allow companies to simply go on there and search for information. There's a subscription to allow ...

INTERVIEWEE 5: I would want to subscribe to that and then I just want to get there once a week to see there are ten more people who fit your criteria. I want the increments or... If I realise there are another ten people then maybe over time it's 20, it's 50, 100. When it gets to 100, you'll figure out more people are willing to buy what I've got. An extra 100.

INTERVIEWER: I really like that idea. It's a really good idea.

INTERVIEWEE 5: You already know what I've bought. You know what the query was.

INTERVIEWER: To be able to jump on there and say right, how many people are searching on my product in that area. I can do something about that. I can get my branch manager, I really like that idea. Because that's a lot less invasive. So the person will still be getting paid because their data is still showing up on their product and they're not being contacted. I really like that. That's fantastic. For things like planning store locations, branch locations to be able to get very accurate estimates of people's behaviours within areas at a time.

INTERVIEWEE 5: In the past, we had a service which allowed companies to upload their data to this portal. They did a match based on that the bureau which then allowed them to download attributes about the customers. Demographic information, age, all sorts of things. It didn't really work very well because I think there's a fair bit of setup to be able to upload data into it, different levels of matching. What we ended up finding was companies just wanted to send data somewhere, you guys figure it out, so we used a lot internally this facility. Because we knew how it worked.

INTERVIEWER: What were the customers getting from you when they get a profile of their existing customers?

INTERVIEWEE 5: I can see it's something a large corporation would use – what about smaller companies?

INTERVIEWER: That's a really big question is where to start because there's so many possibilities so one of the big challenges would be saying what particular data product that we go after, that's going to give us the biggest bang for the buck and simplifying it. I think it's way too complicated. That's some of the feedback I've got. They're looking more for a wizard, step by step thing they're trying to use by themselves, exactly where to start. That's a good question. Probably what I'm going to try to find out from these interviews is, gather all the things we've talked about. Number one on the wish list that would be most likely to pay for.

INTERVIEWEE 5: I think that if you're able to find out more information about your existing customer would be quite valuable. There's more value in trying to identify other customers and things like that. In the back of my mind, I'm thinking, how the hell would I then get that data into my existing system because I want to communicate via my CRM. I've got a central tool that manages all my communications, so if I've got 90% of it here and 10% over here, it's going to be really hard to match that up. The back of my mind, I'm thinking you have something great if I can identify all this how can I get it back to my system whereas the other one, I think would be quite easy. I just put a port into the CRM and that's it. And as soon as they come in, the API and I've got a page of all of this other information that I did not necessarily have before. And I'm not saving it on my system, because that's your information, which is fine, but it was used at that point in time.

INTERVIEWER: This personalised response product would be much easier to use – that makes a lot of sense. It will be easier to implement, integrate ...

INTERVIEWEE 5: You call the call centre or you walk into the branch, if you could incorporate things like next offer if it's relevant to that industry. If I'm looking at people who are buying a home right, would you allow me to pay more for it to be exclusive? Don't show it to competitors. I don't want my competitor to see that information but I'm happy to pay twice as much.

INTERVIEWER: Absolutely, that's the whole thing about what you're doing, the product itself are permission. If the permission is I'm buying exclusive use of the product for the next 12 months, yeah. That can be one of the parameters framing up what the permission is.

INTERVIEWEE 5: I guess the next question is how would you – have you thought about any reporting – to be able to say of customers that have purchased this information. How do we tie that to attribution back, I guess.

INTERVIEWER: I've thought about that. So you know I was thinking along the lines of one scenario that I wrote up in the research paper would say you go to a restaurant and you've got patronages down on Wednesday and Thursdays, so you say right, frame up my data product, I want people in Manly who eat out three times a month, who like Mexican food, who like jazz. So I buy that data product, I direct market them through Facebook ads. I know who came to my restaurant. So at the end of the month, I get a report saying you've paid this much, this is how much people actually came. That's a way of closing the loop. And you've got their payments data. And you know how much they spent.

INTERVIEWEE 5: When we were thinking about what we're going to do about the rating system, the eBay principle where you can rate things. So once you've bought it you can give it 3,4 stars. That would be great. We also thought about the Amazon recommendation type

approach. So people who bought and liked this, also found this useful. To be able to say you actually bought that, the people who bought that actually went on to buy other data products.

INTERVIEWER: So when you're doing this browsing, you'll get recommendations. Another issue we haven't talked about is the reliability of the data. We know people can game the system, get paid for the data the way we deal with that is rating individual based on their reliability

INTERVIEWEE 5: Because you can buy something like 10,000 likes on Instagram for \$1.

INTERVIEWER: Basically what we do is model people according to how realistic, what they're doing, and then their data would be rated as less valuable if it looks like they're gaming the system or displaying irregular behaviour or non-human behaviour, their data would be downgraded. You would pay less for it. Just like Facebook tries to exclude false profiles that aren't real people so that's how you manage that sort of thing.

INTERVIEWEE 5: I'd buy it.

INTERVIEWER: Great. Really appreciate it. Really valuable session as far as giving me different ways to think about and also prioritising what's easy to use, what's not easy to use. Really, really appreciate it. Thanks.

INTERVIEWEE 5: I guess I would have thought a smaller company would find this more important. It's just easier for us, the way companies work.

INTERVIEWER: I'm quite looking forward. One of the next things I'm going to do is prototype the app. Thanks again.

14.2.3 Transcript 6: Interview with Interviewee 6

INTERVIEWER: I'll give you an overview of the system first of all. So basically what I'm designing is a marketplace where people can capture their personal data and then we create data products out of that personal data and companies such as yours can buy them and use them. So you don't get the data, you get to use the data. That's what you're doing, you're purchasing permission so that's the very high level situation. So I'll show you two examples and then I'm going to ask you for your reaction to them. And then we're going to talk about other ones that you might use. I'll make some notes here. So basically what you've got is you've got a personal data store and individuals are clicking massive amounts of data from any type of sensors, could be off their cars, could be off their headsets, laptops, desktops, whatever, ok. And then you're the company, maybe not in a high rise building, but you can browse around in here and see what's available. You can then package it up and create it and then you can –

INTERVIEWEE 6: What do you mean package it –

INTERVIEWER: I'll show you. Create a personal data product and then you put it out and if anybody accepts the offer, you get to buy it and use it. So let's go into a demo. By the way, in a previous piece of research, I had these simulated offers for 205 people and approximately 80% of them accepted one or more. So people are willing to do this. The key point is a very high proportion of people, said yeah, we're happy to share our data. So I'm pretty confident the data will flow into the system. So the next question is what we can make out of it and will companies pay for it.

So this is your homepage, ok. You login there. What you can do is you can browse around and look at existing products that have already been created, ok. So there's one here that I've created for one guy that I'm going to interview. This one's called over 55 Northern beaches adults who are exploring self managed super funds. These ones are examples of targeted marketing. So basically you can create a very niche micro segment of people, and we know they are in this segment because their personal information puts them in this segment. So I've made two for you, adults in Eastern Sydney who are searching for a second car. If you want to see what that looks like you can click on View. So you can see there's 14 people in this product. It's going to cost you \$23 a month. You can see more or less where they are because it's a prototype, they are there, the little red dots. You get the concept, right? They will be all clustered around one postcode and you can see the parameters you've got in here, ages 21 and over so you don't want kids basically because they have to be drivers. They are prospects, so when I created this I specify that I want prospects, not existing customers, so exclude anyone who's already in your customer database. And I've put in some key search terms.

INTERVIEWEE 6: So I need to send you my database in order to use this?

INTERVIEWER: You need to upload a customer list. So basically it lists down here the parameters that you used or I've used to create the product, which in this case is things they are searching for, so go get their ages and location, right. Now if you go, yes, I'll go and buy that, you purchase it and then what we do is we go to Manage our personal data products so you can see these are the ones that you've purchased and you can view it. Now that you've bought it, you can go view it and now we're going to use it. So now we say, right, I want to send a message out to these people, I want to tell them about my new babysitting car service. So I attach a file, I attach my HTML file or plain text file that's got the information and I click Execute. Your data product has been executed, your message bla, bla, bla has been sent to the 14 individuals. So you don't get their data, you don't get their name or anything. But you're able to use the data product –

And you can see that the number of uses if I go back to Manage, the number of uses has gone down ... initially you bought 20 usages, it's gone down to 19. Because you've just used it. So that's the basic concept.

INTERVIEWEE 6: Why did I have 20?

INTERVIEWER: Because that's what you paid for. Let's go through the process of creating one.

INTERVIEWEE 6: What if I do another search and there's 1000 customers there? It's still \$22 or?

INTERVIEWER: Right so let's go create that process ok?

INTERVIEWEE 6: What's the point of this interview again?

INTERVIEWER: The aim is for you to give me feedback. Basically what's your opinion on it, useless, useful? I'm not selling it to you. It doesn't exist, it's just a prototype. The main point is to get your reaction. To find out if you see use in it or something like that and to make suggestions about what can be improved. So let's make one, ok? So we give it a name. Say you're looking at a new suburb –

INTERVIEWEE 6: What about if you're looking for car owners? Is that possible?

INTERVIEWER: Yep. OK, so car owners –

INTERVIEWEE 6: In Sydney. You want to go into a suburb.

INTERVIEWER: Yup. And what sort of car owners?

INTERVIEWEE 6: Depends what we can get really. You can say, people who've got 2 or more cars.

INTERVIEWER: In fact the one that I drafted up to you was people looking to buy 2nd cars. And you can approach them and ask do you really need a 2nd car?

INTERVIEWEE 6: We know our ideal customers, our ideal line have 2 cars, that would be –

INTERVIEWER: OK, so 2 or more cars. Now we go for the process. Do we care about the gender. Not really. Do you care about the age? Yep. We don't want anybody under 25.

INTERVIEWEE 6: 21 is the minimum.

INTERVIEWER: You can see as I move this it shows you the distribution of the people in the database so it's quite interactive. Where do I want them? In Sydney, anywhere between 0 and let's say 50 km.

INTERVIEWEE 6: Probably 10,12. Right there. Can you insert a field there?

INTERVIEWER: Yes. Income. Do you care about income?

INTERVIEWEE 6: No.

INTERVIEWER: That's fine. What we can do is preview it before we buy it. If there's too many people or too expensive we come back and narrow it down. So let's say any income. Do we care if they're married or single? Any children, not really. This is where we go, yep, we want prospects and then we upload a file of existing customers so they don't overlap. We don't want to market to them. Well, it could be email addresses, could be phone.

INTERVIEWEE 6: I'd be concerned about you then taking that.

INTERVIEWER: Sure, that's good feedback. I'm capturing all these. And then we go to Preferences, right? Companies or products they have liked on Facebook. So they are more likely to use our service, you know, sharing services or sharing whatever.

INTERVIEWEE 6: You can maybe refine that.

INTERVIEWER: Web browsing history and things they've been searching for. OK. Say you select those and then it starts to get more sophisticated. So you can say actually I only want people who are connected to my existing customers. Let's say, I want Facebook friends who are connected, say between 1 and 5 levels degrees of separation from existing customers. If they know someone who's using your company, they're more likely to sign up.

INTERVIEWEE 6: But 5 levels is a lot, isn't it? I would've thought that there would be very few people in Sydney for who I wouldn't be connected with. Probably 3 levels? Is that fair?

INTERVIEWER: Sure, we can make it 3. The thing is, what it will work with is once we've previewed it, it will tell you how many people you've found. Sometimes you do a preview and there's none. So you have to start adjusting it.

INTERVIEWEE 6: What you could do is to do show the preview as you update. I would find it valuable to be able to see the change as you go along.

INTERVIEWER: I started doing that up here you've got total individuals, when you click gender, changes it but I haven't built that in yet.

INTERVIEWEE 6: Hovering around the whole thing that's what I'm saying. Sit up on top of and stay in the screen.

INTERVIEWER: Absolutely and if you have the number of people with that price.

INTERVIEWEE 6: As you're doing that Facebook thing you should be able to loosen up with your – make it real-time.

INTERVIEWER: Then what we also have is in the background, because we've got all these data, we've done some analytics on it, calculated various propensities for customers like propensity to complain, propensity to self-service which is a big issue for a lot of customers.

INTERVIEWEE 6: I know the words but I don't know what 'propensity' means.

INTERVIEWER: That's between 0 and 1. 1 means definitely complaining. 0 means they've never complained. So you want customers with a low propensity.

INTERVIEWEE 6: Maybe just put a little tooltip in the bottom saying will 'compliant'.

INTERVIEWER: Plain English.

INTERVIEWEE 6: Yes, exactly. Probably get rid of the 0.2 even.

INTERVIEWER: Good idea. So I try to do a little bit of, if you'll hover over these, it gives you an idea of how good the model is. More plain English. And down here is the contact channels. So how do you want to contact them. Well, I want to email them. And I want to be able to SMS them. So we're all good. We go into the preview. So I only found 4 people. It's only going to cost us \$5 a month. And then we might go yes, that's good.

INTERVIEWEE 6: When you say a month, are we going to contact these people ongoing?

INTERVIEWER: So you've got a month to contact them 20 times. That parameter could be part of the specification, so you can say well, I want to pay for it monthly or pay for it a year or I want to get a job lot so I can use them when I want to. There're different ways to purchase.

What you're creating is a series of permissions, you're buying permissions and that permission could be the number of times you contact them in the period that you contact them.

Ok, so you're happy with it and you purchase it, or you can go back to refine it.

INTERVIEWEE 6: So how did you know they have 2 more cars? Was that one of the questions?

INTERVIEWER: So in that case, we haven't got that information in the dashboard, so if you were guided through the process, it's not in this one, I'll show you another one scenario that I created. If you were going through the process and there was a particular question that was not available, there's a custom box. You describe what it is you want, that would then go through the backend, and we'll say can we infer that?

INTERVIEWEE 6: So we're assuming you could infer it when you found some.

INTERVIEWER: Exactly, yeah. So that's one type of data product. Basically it helps you create a very niche target market.

INTERVIEWEE 6: Confusing what you say about -, I still don't understand what the monthly thing is. It's really unclear to me what so I'm going to pay \$4 a month to contact these people

20 times over the month, up to 20 times is it? Are these people going to want me to contact them 20 times a month? It seems really high.

INTERVIEWER: That's really good feedback, that's something I need to fix. I think what I need to do is say make it much clearer. Do you want to purchase a monthly for a maximum number of contacts or do you want to just buy 20 contacts which you can use whenever you want.

Alright, that's it. Any other feedback on that one, the targeted marketing?

INTERVIEWEE 6: So for \$4 you've 24 contacts, how much you're going to pay for that?

INTERVIEWER: For that one, we go back to your menu and we look at this one we just made which was car owners, review it. \$4.60, there's 4 people in the product. And 20 uses left. You can see up here, you've got how many uses remaining, their expiry date.

INTERVIEWEE 6: What does use mean? I'm, just trying to work it out, wouldn't have any idea-

INTERVIEWER: You understand the use now, choose a file and add the information, attach it and then execute it. And that would say 'your message has been sent to 4 individuals'.

INTERVIEWEE 6: So how would you choose it, what sort of file?

INTERVIEWER: The emails you send out to your existing customers, perhaps a new service just been launched.

INTERVIEWEE 6: So would there be a preview? How would you upload that message, know that it's going to work? When you say that, that file's an email. How do we send it, how do we send it through here?

INTERVIEWER: The way you send it through is by connecting it here. So you've got your email all set up. Simply attach it, the HTML file and then execute it. That's it. Then it goes to us and gets emailed out, or SMSed out.

INTERVIEWEE 6: So I guess I'd be expecting to see a portal that lets me upload a message properly and format it. We need to think about things like that.

INTERVIEWER: So that can preview the ad and everything.

INTERVIEWEE 6: Because it's going to be an email you're going to send out, like a mail merge.

INTERVIEWER: That's good feedback too.

INTERVIEWEE 6: And also the SMS. They can be pretty tricky SMS, basically be able to input the text that you want to write.

INTERVIEWER: Basically need a preview button here.

INTERVIEWEE 6: How many different ways they can contact? Email, sms?

INTERVIEWER: Let me go back to create. Email, Facebook ads, Google+ ads, Google search ads, internet, phone, sms. These are some of our products, you just want to analyse their data, you don't want –

INTERVIEWEE 6: So how are you going to do phone?

INTERVIEWER: I haven't built that interface where you can go in there and you can say I'm ready to talk to this person.

INTERVIEWEE 6: That would be quite complex, wouldn't it?

INTERVIEWER: And then there's another type of product which is called personalised service response. The idea here - so those are the targeted marketing ones, there's another one called personalised service response. Basically the idea is somebody contacts your company, either existing customer or a prospect, instead of taking hours interviewing them and asking them questions, you simply enter one piece of information you know about them and everything else you've bought access to comes up. So you can then personalise the service that you're providing them more quickly. So we just click and go through one of those. What I'll do is select all, turn off this custom one, preview and –

So now we're going to view it. Basically you provide this to your call centre staff. Somebody rings up, you know their email address which is this guy, instead of having to ask them, because you've bought access to one of those, you can see what type of car he uses, what subscription he has, what technology he uses, whatever's relevant, you would customise this to what's relevant to your business.

INTERVIEWEE 6: So this can only be used once it's plugged to live somewhere with people on the platform, is that right?

INTERVIEWER: Yes. It gives them a way of personalising the way they respond to customer and prospects.

INTERVIEWEE 6: This can only be if someone's opted in the platform.

INTERVIEWER: Everything is opt-in. When you create the product, when you click preview, only people who have agreed to share their data are included in the data product. If they only agree to share their web browsing history and that's being selected, you might come up. Absolutely everything is pre opt-in.

INTERVIEWEE 6: So if they do share their web browsing history, how is that information going to get through to us or to you?

INTERVIEWER: That's another part of the research, building a system that collects your information for you on your behalf, stores it and then exposes it into products that you've agreed to participate in. basically we're building that, you install on your phone and it captures the information that you want to capture and stores it on your behalf.

I'm meeting with another guy who has built a system that does that, it doesn't do the bit that captures it on your phone, but it stores anything, if you put a Java code in your app, then it will capture what happens on your app to a database. That part of things is already well done.

Any other questions about how it works?

INTERVIEWEE 6: I think a wizard, not a wizard but a step by step would be handy, like step one, you want to get this data, step two, do this. It's different when you're showing it to me, if you set me down and said, right, use it to send something out, it's in its current form, I wouldn't know what to do. I wouldn't know where to start. You'd be dancing around in circles. Basically lets you choose, put together a target market the same way you can do with Facebook or whatever. I'm just trying to explain what I think it is, you can then be able to work out if it's not clear. But all of these people have opted in to receiving direct messages from me but through your system. Put the list of people together and then send the message that I want to send and for \$4.50 I can send them up to 20 messages. I would guess, if I contacted my existing customers more than 3 times, they would unsubscribe you. I would be more restricted if I were you, how often they can contact them. Stretch out the time or something like that. Sending 20 messages out a month, you probably find a lot of people would be, because you know those people signed up to receive 20 messages a month, for the \$1.20 they've agreed to or whatever, they might've signed up to get one or maybe two, they'd unsubscribe. When you buy a mailing list of people, you're pretty sure you don't buy 20 people. The permission to send to 20 is like for 1 use. And if they want to get it again, they can. That's the decision they make then.

Would you be happy to get paid \$2 to be contacted by these people? And the answer is yes. Maybe they've agreed to be paid \$2 to be contacted twice by these people but contacted 20 times in a period of a month, I would say no.

The second product, anything that comes to my mind is like, it'd be pretty hard unless you started massive penetration like you know, 3 or 4 million people in there, there's going to be a lot of misses. But that could be a quite useful ID-ing tool for us as well. Like a trust verification thing. Extra confidence in who they are.

INTERVIEWER: That's my next question to you was, are there any ways you can see such a system could be used. I'm not getting too much down into the exact detail of the prototype. The concept of clicking and being allowed access to personal information, so things like that, using it as a verification tool. That's really interesting to hear about.

INTERVIEWEE 6: How much would you pay per usage, like pop up a person and you get the data, how much would you pay for that?

INTERVIEWER: The exact figure around pricing is unknown at the moment.

INTERVIEWEE 6: We're talking roughly like 20 cents or \$3?

INTERVIEWER: It would be minimal, a few cents.

INTERVIEWEE 6: So we pay \$2.50 every time we do a credit check, we do it because we want to check their credit record, we also do it because it lets us verify who they are. This could be an application.

You could ask them which telco they're with and you know because you guys know. Is there any identifying things you're going to ask them or you're able to get from that list?

INTERVIEWER: Really what it comes down to is how willing people are to share their data. We'd only know that in practice. So that would be - certainly we know there could be other systems. At the moment, we've got create data products, targeted marketing, personalised service response, personalised recommendations, that's going to be another one I haven't built yet. Based on your behaviour and your similarity to other customers, here are the recommendations for you. Probably wouldn't apply to you because of the limited number of products. But if you have a lot of products.

So thinking about other businesses you're familiar with, not necessarily your own, how do you see this working? Do you see it would have any benefit?

INTERVIEWEE 6: Yes, because just like us, as you know, if the cost of acquisition is below your lifetime value, it's worthwhile doing it. So I'm guessing for a lot of scenarios, it's like Google or AdWords. It's going to come in below or not, and where it comes in below, it's going to be a differentiator. There's a lot of similarities to Google. This is going to be treading on Google's toes a bit, I'm guessing?

INTERVIEWER: In fact, it's going to be - it's going to do something different to what Google does. Because Google just gets the Google bit. And the same with Facebook. What we're doing here is linking those multiple data streams on behalf of the individual because the individual's only one who has access and permissions to their data.

INTERVIEWEE 6: If I pay you \$10, you'd take a 10%, and give the \$9 to the individual.

INTERVIEWER: Exactly, yes. That's how it works. Basically we're the market makers, create a product. We've talked a bit about the changes, making it more user friendly. Say if we did solve the user interface issues, how do you see something like this incorporated into your workflow? Your internal backend systems. The issues with that?

INTERVIEWEE 6: I'm guessing our marketing would be hooked into it. Which is why the interface has to be good. I don't know if you're planning any APIs for it, if we work out, we can automate certain things, then we could start to automate it as well.

INTERVIEWER: It's a good point about the APIs. What we're looking at is creating 3 different types of APIs. One is for adding new data streams, so as new devices come online or when new services come up on the web, we can plug those in. The other one is APIs that would allow modellers to build models and extract meaningful information out of them. And every time a product is created or executed, they get a cut as well.

Basically, to make the personal information product, you need the data, that's the raw material. The individual, you need the model to extract information, that's the modeller. And you need a consumer who put those together. And purchases it. How do you see the API working on the consumer side?

INTERVIEWEE 6: I'm sort of thinking through potential scenarios, so let's say that we expand into a new suburb. Then the API can go right or it checks that suburb list and when a new suburb is added, so let's say now we're doing 2024 then that would send that to your system and your system will go, right, do this step of actions when a new suburb expanded.

INTERVIEWER: Right, so like a premade data product and you just need to know which suburb to move it to.

INTERVIEWEE 6: I would say that's a long way off. You just do it manually. So I'm not suggesting you go and build that into your first round, maybe down the track. Having an app sort of store for modellers would be very useful as well. I don't know how that's going to be built into your system.

INTERVIEWER: I think that's really key because if we're doing it all when you're able to do so many basically when you go through and create these things the more boxes there are to tick down here, the better for the user, right? Each modeller would build a new one, which would show up as a new choice or a new option for you to choose. I think that's quite a lot of potential. So we've talked about other data products and one was the identification, you know, verification product. Anything else that you think you guys would use if you had that repository of personal data and you want to extract insights from them? I was thinking things like driving patterns. You have to have your car, personal car logged into it. You could see if that person was a safe driver, whether they were driving in blind spots or car accidents.

INTERVIEWEE 6: We're building a system now that, we will when we have time, that tells us our members are shitty drivers, or kick them off but if you got any tool that could tell us before

that, before they even apply or when they apply that would be their personal data thing but I don't know how the hell you –

INTERVIEWER: I would have to have the telematics data from their car.

INTERVIEWEE 6: And have a GPS system.

INTERVIEWER: The other one would be insurance claims too. If you were to know how many times that person have made claims on accidents, or involved in accidents even if it wasn't their fault, it could be useful information.

INTERVIEWEE 6: I know insurance aren't allowed to ask other insurance that which I think's a stupid rule but the insurance can't compare rates. So how would you be able to do that? By their searches?

INTERVIEWER: First of all, no. the person has to agree, I want to share that data. I know it's valuable, I get paid for it. I want the company to know about me, because I'm a great driver. People who are willing to share.

INTERVIEWEE 6: Hypothetically, how would you find out that someone's had 6 insurance claims?

INTERVIEWER: That would be hard.

INTERVIEWEE 6: You don't get their individual Google searches or anything like that. Is there some special program will let you pretty much pick up every URL they've gone to?

INTERVIEWER: Yeah, it already exists.

INTERVIEWEE 6: For providing that they get extra money. So over a number of years they can't turn it on or off, they have to turn it on the whole time?

INTERVIEWER: They could opt in or out. There's no reason why they could say I've changed my preferences. How much you use it and how much you disclose back to them is a big issue.

INTERVIEWEE 6: Exactly. I'm sure there's endless stuff. What if there is anything you can use like, we use it as a referral generator. So, you could probably tell us who our users are by the amount of their searching for us, logging on or whatever. So I'm wondering if there's any application people who've logged on, but we know they are logged on from our end.

INTERVIEWER: It's more about prospects.

INTERVIEWEE 6: There's very little application once you've signed one on. Or we could use it to – here's one. We've been to Google on this one actually. We could use it to get much deeper understanding of the customer base. Who they really are, without them telling us. We did a customer survey and it'll tell us this. But if you matched that up with what they're really doing,

that would be very valuable. We will definitely use that. We've been talking about doing that with Google.

INTERVIEWER: What sort of things do you want to know about them?

INTERVIEWEE 6: What their interests are, what their interests really are, the example I told Google about was we might find a large number of people like outdoors and camping, you know, whatever. So we can start targeting our market towards that. And partnerships and things. Whereas if you survey the people, they might tell you they like going to the beach and they're actually –

INTERVIEWER: That's more around customer segmentation. Behavioural segmentation.

INTERVIEWEE 6: And on your pre-existing customers. We can match that back to how much they actually spend. So for our top performing customers we can give out their email addresses to match them up and you come back and give us back a customer profile. That would be very valuable.

INTERVIEWER: Then you can do targeted messaging to your own base. This group is going to get the outdoors message, this group is going to get the inner city restaurant message.

INTERVIEWEE 6: I haven't thought about that bit, I was more thinking like we know this group of 25 – 35 year old females who love the outdoors who are spending a disproportionately high amount of money. So let's go after that group because if we can have every one of them as our customers, lifetime value's going to be double because they spend.

INTERVIEWER: So that will actually inform the targeted marketing product that you –

INTERVIEWEE 6: In the same way that you do a survey, but this is real.

INTERVIEWER: Behaviour doesn't lie.

INTERVIEWEE 6: Exactly. Or minimum you could use that to cross reference against the survey and you double check the survey's right. Adds an element of richness to the survey. That would be a good one for you because that's your upsell. When you've already got a customer you can – I sat down with this girl, she does Google stuff. Whatever the Google products are, it's her job to get the businesses to use it as well as they can and to buy more stuff. That's obviously something that's going well for Google.

INTERVIEWER: Makes sense.

INTERVIEWEE 6: I'll have a think about any other applications.

INTERVIEWER: If you come across any other contacts who would agree to be interviewed, you think wouldn't mind sitting down, I've got quite a lot lined up already because I've got a bit of a matrix of different business sizes, tick them all off.

INTERVIEWEE 6: What verticals are you looking for? What business sizes?

INTERVIEWER: All of them. Yeah, I need to fill up a matrix of it. So thanks for your time – much appreciated.

14.3 Appendix 3

14.3.1 R code used to generate and analyse data with regards to individuals' appetites to participate on a PPIM.

```
install.packages('RPostgreSQL')
install.packages('plyr')
install.packages('sqldf')
install.packages('ggplot2')
install.packages('stringr')
install.packages('ROCR')
install.packages('glmulti')
install.packages('randomForest')
install.packages('ggmap')
install.packages('maps')
install.packages('stringi')
#-----
# set an option to make sqldf work properly on Mac
#-----
options(gsubfn.engine = "R")

#-----
# set up libraries
#-----
library(RPostgreSQL)
library(plyr)
library(sqldf)
library(ggplot2)
library(stringr)
```

```

library(ROCR)
library(glmulti)
library(randomForest)
library(ggmap)
library(stringi)

#-----
# connect to database
#-----
drv <- dbDriver("PostgreSQL")
con <- dbConnect(drv, host='XXXX', port='XXXX', dbname='XXXX', user='XXXX',
password='XXXX')

#-----
# read in data
#-----

results_query <- dbSendQuery(con, "select * from results")
results <- fetch(results_query,n=-1) ## fetch all elements from the result set

users_query <- dbSendQuery(con, "select * from users")
users <- fetch(users_query,n=-1) ## fetch all elements from the users set

uses_query <- dbSendQuery(con, "select * from uses")
uses <- fetch(uses_query,n=-1) ## fetch all elements from the users set

org_query <- dbSendQuery(con, "select * from organisations")
organisations <- fetch(org_query,n=-1) ## fetch all elements from the users set

dbDisconnect(con) ## Closes the connection
dbUnloadDriver(drv)

#detach RPostgres do sqldf will work on data frames
detach("package:RPostgreSQL", unload=TRUE)

#-----
#tidy up the uses and results

```

```

#-----

uses$name <- str_replace_all(uses$name, "personalised", "personalized")
results$use <- str_replace_all(results$use, "personalised", "personalized")

uses$name <- str_replace_all(uses$name, "revant", "relevant")
results$use <- str_replace_all(results$use, "revant", "relevant")

#-----
# join the tables to make the full results data set
#-----

results_1<- sqldf(
"select
users.id as user_id
, users.firstname as firstname
, users.lastname as lastname
, users.email as email
, users.age as age
, users.gender as gender
, users.highest_educational_attainment as highest_educational_attainment
, users.ethnic_background as ethnic_background
, users.state as state
, users.postcode as postcode
, results.id as result_id
, results.company as company
, results.use as use
, results.data as data
, results.device as device
, (case when trim(contact_channel) = 'would like to send you relevant messages via email'
then 'email'
when trim(contact_channel) = 'would like to send you relevant messages via facebook ads'
then 'facebook ads'
when trim(contact_channel) = 'would like to send you relevant messages via Four Square'
then 'Four Square'
when trim(contact_channel) = 'would like to send you relevant messages via Google plus ads'
then 'Google plus ads'
when trim(contact_channel) = 'would like to send you relevant messages via Google search
ads' then 'Google search ads'

```

```

when trim(contact_channel) = 'would like to send you relevant messages via linkedin ads'
then 'linkedin ads'

when trim(contact_channel) = 'would like to send you relevant messages via phone'
then 'phone'

when trim(contact_channel) = 'would like to send you relevant messages via sms' then 'sms'
else trim(contact_channel) end) as contact_channel
, results.price as price
, results.response as response
from results left join users
on results.user_id = users.id;")

results_2 <- sqldf("select
results_1.*
,uses.type as use_type
from results_1 left join uses
on results_1.use = uses.name;")

full_results<- sqldf(
"select
results_2.*
,organisations.id as org_id
,organisations.name as org_name
,organisations.type as org_type
,organisations.description as org_description
from results_2 left join organisations
on results_2.company = organisations.name")

#-----
# tidy up data ready for modelling
#-----
#include a numeric version of the respnse variable

full_results <- cbind(full_results, numerical_response =
as.numeric(full_results$response=='yes'))

#do a few sanity checks to make sure the joins are OK
sqldf("select use_type, count(*)
from full_results

```

```
group by 1
order by 2 desc")
```

```
sqldf("select use, count(*)
from full_results
group by 1
order by 2 desc")
```

```
sqldf("select distinct use
from full_results
where use_type is null")
```

```
#-- should return zero rows
```

```
#-----
```

```
# summary of respondents
```

```
#-----
```

```
#number of respondents
```

```
nrow(users)
```

```
#gender
```

```
table(users$gender)
```

```
#gender
```

```
table(users$age)
```

```
table(users$highest_educational_attainment)
```

```
table(users$ethnic_background)
```

```
table(users$state)
```

```
#-----
```

```
# summary of responses
```

```
#-----
```

```

nrow(full_results)

table(full_results$response)

#histogram of number of responses per person

num_responses_per_user <- sqldf(
  "select user_id, count(*) as freq
  from full_results
  where response = 'yes'
  group by 1"
)
nrow(num_responses_per_user)

m <- ggplot(num_responses_per_user, aes(x=freq))
m + geom_histogram(binwidth = 1)

#-----
# estimate the proportion of Australians who would participate in some way in the PPIM?
#-----
num_participants <- nrow(as.data.frame(table(full_results[full_results$response
== 'yes', 'user_id'])))
num_users <- nrow(users)

prop.test(num_participants, num_users, conf.level = 0.90)

num_participants/num_users

#-----
# estimate the proportion of proposed permissions which would be accepted?
#-----
prop.test(nrow(full_results[full_results$response == 'yes',]), nrow(full_results), conf.level =
0.90)

#-----
# identify the important factors which determine the price which participants will accept for an
offer
#-----

```

```
#explore individual factors
```

```
#plot the proportion of offers accepted by highest_educational_attainment
```

```
sqldf('select * from full_results limit 10')
```

```
sqldf("select age, response, case when response = 'yes' then 1 else 0 end as num_pos from  
full_results  
order by 1")
```

```
prop_by_age <- cbind(sqldf("select age, sum(numerical_response) as num_pos from  
full_results  
group by age  
order by 1"),  
sqldf("select age, count(*) as num_resp from full_results  
group by age  
order by 1")  
)
```

```
prop_by_age <- cbind(prop_by_age, percent =  
round((prop_by_age$num_pos/prop_by_age$num_resp)*100,2))
```

```
ggplot(prop_by_age, aes(x = age, y = percent))+  
geom_point(stat = "identity", size = 4)+  
xlab("Age")+  
ylab("Percent")+  
theme(axis.text=element_text(size=20),axis.title=element_text(size=30))+  
geom_smooth(method=lm, se=FALSE)
```

```
#check gender
```

```
prop_by_gender <- cbind(sqldf("select gender , sum(numerical_response) as num_pos from  
full_results  
group by gender  
order by 1"),  
sqldf("select gender , count(*) as num_resp from full_results  
group by gender
```

```
order by 1")
)
```

```
prop_by_gender <- cbind(prop_by_gender , prop =
round((prop_by_gender$num_pos/prop_by_gender$num_resp)*100,2))
```

```
ggplot(prop_by_gender , aes(x = factor(gender), y = prop)) + geom_bar(stat = "identity")
```

```
prop_by_highest_educational_attainment <- cbind(sqldf("select
highest_educational_attainment , sum(numerical_response) as num_pos from full_results
group by highest_educational_attainment
order by 1"),
sqldf("select highest_educational_attainment , count(*) as num_resp from full_results
group by highest_educational_attainment
order by 1")
)
```

```
prop_by_highest_educational_attainment <- cbind(prop_by_highest_educational_attainment ,
percent =
round((prop_by_highest_educational_attainment$num_pos/prop_by_highest_educational_atta
inment$num_resp)*100,2))
```

```
print.data.frame(prop_by_highest_educational_attainment)
```

```
ggplot(
prop_by_highest_educational_attainment
, aes(fill = highest_educational_attainment, x = factor(highest_educational_attainment), y =
prop)
) + geom_bar(stat = "identity" )
```

```
prop_by_ethnic_background <- cbind(sqldf("select ethnic_background ,
sum(numerical_response) as num_pos from full_results
group by ethnic_background
order by 1"),
sqldf("select ethnic_background , count(*) as num_resp from full_results
group by ethnic_background
order by 1")
)
```

```
prop_by_ethnic_background <- cbind(prop_by_ethnic_background , prop =  
prop_by_ethnic_background$num_pos/prop_by_ethnic_background$num_resp)
```

```
ggplot(  
  prop_by_ethnic_background  
  , aes(fill = ethnic_background, x = factor(ethnic_background), y = prop)  
  ) + geom_bar(stat = "identity" )
```

```
prop_by_use <- cbind(sqldf("select use , sum(numerical_response) as num_pos from  
full_results
```

```
  group by use  
  order by 1"),  
  sqldf("select use , count(*) as num_resp from full_results  
  group by use  
  order by 1")  
  )
```

```
prop_by_use <- cbind(prop_by_use , prop = prop_by_use$num_pos/prop_by_use$num_resp)
```

```
ggplot(  
  prop_by_use  
  , aes(fill = use, x = factor(use), y = prop)  
  ) + geom_bar(stat = "identity" )
```

```
prop_by_use_type <- cbind(sqldf("select use_type , sum(numerical_response) as num_pos  
from full_results
```

```
  group by use_type  
  order by 1"),  
  sqldf("select use_type , count(*) as num_resp from full_results  
  group by use_type  
  order by 1")  
  )
```

```
prop_by_use_type <- cbind(prop_by_use_type , prop =  
prop_by_use_type$num_pos/prop_by_use_type$num_resp)
```

```
ggplot(  
  prop_by_use_type
```

```
, aes(fill = use_type, x = factor(use_type), y = prop)
) + geom_bar(stat = "identity" )
```

```
prop_by_contact_channel <- cbind(sqldf("select contact_channel , sum(numerical_response)
as num_pos from full_results
  group by contact_channel
  order by 1"),
sqldf("select contact_channel , count(*) as num_resp from full_results
  group by contact_channel
  order by 1")
)
```

```
prop_by_contact_channel <- cbind(prop_by_contact_channel , prop =
prop_by_contact_channel$num_pos/prop_by_contact_channel$num_resp)
```

```
ggplot(
  prop_by_contact_channel
  , aes(fill = contact_channel, x = factor(contact_channel), y = prop)
) + geom_bar(stat = "identity" )
```

```
full_results <- cbind (full_results,
  price_cat = sqldf("select (case when price < 0.2 then'0.00 - 0.19'
  when price >= 0.2 and price < 0.4 then'0.20 - 0.39'
  when price >= 0.4 and price < 0.6 then'0.40 - 0.59'
  when price >= 0.6 and price < 0.8 then'0.60 - 0.79'
  when price >= 0.8 and price < 1.0 then'0.80 - 0.99'

  when price >= 1.0 and price < 1.2 then'1.00 - 1.19'
  when price >= 1.2 and price < 1.4 then'1.20 - 1.39'
  when price >= 1.4 and price < 1.6 then'1.40 - 1.59'
  when price >= 1.6 and price < 1.8 then'1.60 - 1.79'
  when price >= 1.8 and price < 2.0 then'1.80 - 1.99'

  when price >= 2.0 and price < 2.2 then'2.00 - 2.19'
  when price >= 2.2 and price < 2.4 then'2.20 - 2.39'
  when price >= 2.4 and price < 2.6 then'2.40 - 2.59'
  when price >= 2.6 and price < 2.8 then'2.60 - 2.79'
```

```

when price >= 2.8 and price < 3.0 then '2.80 - 2.99'

when price >= 3.0 and price < 3.2 then '3.00 - 3.19'
when price >= 3.2 and price < 3.4 then '3.20 - 3.39'
when price >= 3.4 and price < 3.6 then '3.40 - 3.59'
when price >= 3.6 and price < 3.8 then '3.60 - 3.79'

else '3.80 - 4.00'

end) as price_cat from full_results")
)

prop_by_price_cat <- cbind(sqldf("select price_cat , sum(numerical_response) as num_pos
from full_results
  group by price_cat
  order by 1"),
sqldf("select price_cat , count(*) as num_resp from full_results
  group by price_cat
  order by 1"))
)

prop_by_price_cat <- cbind(prop_by_price_cat , percent =
round((prop_by_price_cat$num_pos/prop_by_price_cat$num_resp)*100,2))

ggplot(prop_by_price_cat, aes(x = price_cat, y = percent))+
  geom_point(stat = "identity", size = 4)+
  xlab("Price")+
  ylab("Percent")+
  theme(axis.text=element_text(size=20),axis.title=element_text(size=30))

ggplot(
  prop_by_price_cat
  , aes(fill = price_cat, x = factor(price_cat), y = prop)
) + geom_bar(stat = "identity" )

prop_by_org_type <- cbind(sqldf("select org_type , sum(numerical_response) as num_pos
from full_results
  group by org_type

```

```

    order by 1"),
  sqldf("select org_type , count(*) as num_resp from full_results
  group by org_type
  order by 1")
)

prop_by_org_type <- cbind(prop_by_org_type , prop =
prop_by_org_type$num_pos/prop_by_org_type$num_resp)

ggplot(
  prop_by_org_type
  , aes(fill = org_type, x = factor(org_type), y = prop)
) + geom_bar(stat = "identity" )

full_results <- subset(full_results, select = -c(org_type_reduced) )

#create a factor which distinguished health organisations from all other
full_results <- cbind(full_results,
  org_type_reduced = sqldf("select (case when org_type = 'Health and Community Services'
then 'Health and Community Services'
  else 'other' end
) as org_type_reduced from full_results"))

full_results <- subset(full_results, select = -c(use_type_reduced) )

full_results <- cbind(full_results,
  use_type_reduced = sqldf("select (case when use_type like '%personalized medicine%'
then 'personalized medicine'
  else 'other' end
) as use_type_reduced from full_results"))

full_results <- cbind(full_results,
  use_type_reduced = sqldf("select (
case when use_type like '%personalized medicine%' then 'personalized medicine'
else 'other' end
) as use_type_reduced from full_results"))

```

```

full_results <- cbind(full_results,
  use_type_reduced = sqldf("select (
    case when contact_channel like '%facebook ads%' then 'facebook ads'
    when contact_channel like '%Google search ads%' then 'Google search ads'
    when contact_channel like '%sms%' then 'sms'
    when contact_channel like '%will not contact you directly through any media%'
then 'will not contact'
    else 'other' end
  ) as contact_channel_reduced from full_results"))

```

```

fit <- glm(numerical_response ~ age + price + use_type , data = full_results, family =
binomial())

```

```

fit <- glm(numerical_response ~ . , data = full_results, family = binomial())

```

```

summary(fit) # display full_results

```

```

confint(fit) # 95% CI for the coefficients

```

```

exp(coef(fit)) # exponentiated coefficients

```

```

exp(confint(fit)) # 95% CI for exponentiated coefficients

```

```

pred <- predict(fit, type="response") # predicted values

```

```

residuals(fit, type="deviance") # residuals

```

```

make quantile-quantile plot of residuals

```

```

qqnorm(residuals(fit, type="deviance"))

```

```

abline(a=0,b=1)

```

```

# assess the goodness of fit and predictive power of the model

```

```

full_results <- subset(full_results, select = -c(data_reduced) )

```

```

full_results <- cbind(full_results,
  data_reduced = sqldf("select (
    case when data like '%bank statements%' then 'bank statements'
    else 'other' end
  ) as data_reduced from full_results"))

```

```

#-----

```

```

# stepwise regression for all factors

```

```

#-----

```

```

# All Subsets Regression

```

```

glmulti.logistic.out <-
  glmulti(numerical_response ~
age+price+contact_channel+use_type+ethnic_background+org_type+gender
  ,data = full_results,
  level = 1, # No interaction considered
  method = "h", # Exhaustive approach
  crit = "aic", # AIC as criteria
  confsetsize = 5, # Keep 5 best models
  plotty = F, report = F, # No plot or interim reports
  fitfunction = "glm", # glm function
  family = binomial) # binomial family for logistic regression

## Show result for the best model
summary(glmulti.logistic.out@objects[[1]])

## Show 5 best models (Use @ instead of $ for an S4 object)
glmulti.logistic.out@formulas

summary(glmulti.logistic.out)

#-----
# do some likelihood ratio tests to check that overall the categorical variables are significant
#-----

my.mod1 <- glm(numerical_response ~ age, data = full_results, family = "binomial")
my.mod2 <- glm(numerical_response ~ age+price, data = full_results, family = "binomial")
anova(my.mod1, my.mod2, test="LRT")

my.mod1 <- glm(numerical_response ~ age+price, data = full_results, family = "binomial")
my.mod2 <- glm(numerical_response ~ age+price+contact_channel, data = full_results, family
= "binomial")
anova(my.mod1, my.mod2, test="LRT")

my.mod1 <- glm(numerical_response ~ age+price+contact_channel, data = full_results, family
= "binomial")

```

```

my.mod2 <- glm(numerical_response ~ age+price+contact_channel+use_type, data =
full_results, family = "binomial")

anova(my.mod1, my.mod2, test="LRT")

my.mod1 <- glm(numerical_response ~ age+price+contact_channel+use_type, data =
full_results, family = "binomial")

my.mod2 <- glm(numerical_response ~
age+price+contact_channel+use_type+ethnic_background, data = full_results, family =
"binomial")

anova(my.mod1, my.mod2, test="LRT")

#device adds nothing

my.mod1 <- glm(numerical_response ~
age+price+contact_channel+use_type+ethnic_background, data = full_results, family =
"binomial")

my.mod2 <- glm(numerical_response ~
age+price+contact_channel+use_type+ethnic_background+device, data = full_results, family
= "binomial")

anova(my.mod1, my.mod2, test="LRT")

my.mod1 <- glm(numerical_response ~
age+price+contact_channel+use_type+ethnic_background, data = full_results, family =
"binomial")

my.mod2 <- glm(numerical_response ~
age+price+contact_channel+use_type+ethnic_background+org_type, data = full_results,
family = "binomial")

anova(my.mod1, my.mod2, test="LRT")

#####

# refit the model setting the reference categories
#####

sqldf("select distinct ethnic_background from full_results")

# set referebce levels

full_results$contact_channel <- relevel(as.factor(full_results$contact_channel), "will not
contact you directly through any media")

full_results$use_type <- relevel(as.factor(full_results$use_type), "personalized medicine ")

full_results$org_type <- relevel(as.factor(full_results$org_type), "Health and Community
Services")

```

```
full_results$ethnic_background <- relevel(as.factor(full_results$ethnic_background), "Non-
Hispanic White or Euro")
```

```
fit <- glm(numerical_response ~ age + price + contact_channel + use_type +
ethnic_background + org_type, data = full_results, family = binomial())
```

```
summary(fit) # display full_results
```

```
#check number of users in each category
```

```
table(full_results$ethnic_background)
```

```
table(users$state)
```

```
#-----
```

```
# assess the predictive power of the best model
```

```
#-----
```

```
floor(2*nrow(full_results)/3)
```

```
train <- sqldf("select * from full_results limit 2465")
```

```
test <- sqldf("select * from full_results where result_id not in (select result_id from train)")
```

```
#---- logistic regression
```

```
lr_model <- glm(numerical_response ~ age + gender + price + contact_channel + use_type +
org_type, family=binomial(link='logit'), data=train)
```

```
fitted.results <- predict(lr_model, newdata=test, type='response')
```

```
fitted.results <- ifelse(fitted.results > 0.5, 1, 0)
```

```
misClassificationError <- mean(fitted.results != test$numerical_response)
```

```
print(paste('Accuracy', 1-misClassificationError))
```

```
p <- predict(lr_model, newdata=test, type="response")
```

```
pr <- prediction(p, test$numerical_response)
```

```
prf <- performance(pr, measure = "tpr", x.measure = "fpr")
```

```
plot(prf)
```

```
auc <- performance(pr, measure = "auc")
```

```
auc <- auc@y.values[[1]]
```

```
auc
```

14.4 R code used to simulate population data in section 8.5

```
full_results$income <- paste('$',formatC(round(runif(nrow(full_results), min = 10000, max =
400000), 0), format="d", big.mark=','), sep = "")
setwd('/Users/rf186004/Documents/Dropbox/___PhD/prototype/business facing');
au_nz_cities <- read.csv('AU_NZ_cities.csv', header = TRUE)
common_words_4000 <- read.csv('4000-most-common-english-words-csv.csv', header =
FALSE)
full_results <- data.frame(
  full_results,
  au_nz_cities[sample(1:nrow(au_nz_cities), nrow(full_results), replace = TRUE), c(2, 4, 6, 7)]
)

full_results <- data.frame(
  full_results
  , relationship_status = sample(c('single', 'married', 'de-facto', 'divorced'), nrow(full_results),
replace = TRUE)
)

full_results <- data.frame(
  full_results
  , num_children = sample(c(0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 4, 5, 6),
nrow(full_results), replace = TRUE)
)

full_results <- data.frame(
  full_results
  , num_connections = sample(c(0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 4, 5,
6), nrow(full_results), replace = TRUE)
)

full_results <- data.frame(
  full_results
  , num_siblings = sample(c(0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 4, 5, 6),
nrow(full_results), replace = TRUE)
)

full_results <- data.frame(
  full_results
  , num_parents = sample(c(0, 1, 2), nrow(full_results), replace = TRUE)
```

```

)
full_results <- data.frame(
  full_results
  , type_of_customer = sample(c('existing_customer', 'prospect'), nrow(full_results), replace =
TRUE)
)
full_results <- data.frame(
  full_results
  , preferences = tolower(replicate(nrow(full_results),
paste(as.character(sample(common_words_4000$V1, 100, replace = TRUE)), collapse = ' ')))
)
full_results <- data.frame(
  full_results
  , web_browsing_history = tolower(replicate(nrow(full_results),
paste(as.character(sample(common_words_4000$V1, 100, replace = TRUE)), collapse = ' ')))
)
full_results <- data.frame(
  full_results
  , search_terms = tolower(replicate(nrow(full_results),
paste(as.character(sample(common_words_4000$V1, 100, replace = TRUE)), collapse = ' ')))
)
full_results <- data.frame(
  full_results
  , degrees = sample(c(1:40), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
  full_results
  , maven = sample(c(1:2000), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
  full_results
  , influence = sample(seq(from = 0, to = 1, by = 0.01), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
  full_results
  , embedded = sample(seq(from = 0, to = 10, by = 0.1), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(

```

```

full_results
, propensity_to_buy = sample(seq(from = 0, to = 1, by = 0.01), nrow(full_results), replace =
TRUE)
, propensity_to_complain = sample(seq(from = 0, to = 1, by = 0.01), nrow(full_results),
replace = TRUE)
, propensity_to_churn = sample(seq(from = 0, to = 1, by = 0.01), nrow(full_results), replace =
TRUE)
, propensity_to_self_service = sample(seq(from = 0, to = 1, by = 0.01), nrow(full_results),
replace = TRUE)
)
full_results <- data.frame(
full_results
, contact_frequency = sample(seq(from = 0, to = 15, by = 1), nrow(full_results), replace =
TRUE)
)
full_results <- data.frame(
full_results
, religion = sample(c('Christian', 'Jewish', 'Islam', 'Atheist', 'Agnostic', 'Buddist'),
nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
full_results
, current_employment = sample(c('Academic librarian',
'Accountant (public finance, chartered)',
'Accountant (tax, chartered)',
'Accountant',
'Accounting technician',
'Actuary',
'Adult nurse',
'Advertising account executive',
'Advertising account planner',
'Advertising copywriter',
'Advice worker',
'Advocate (Scotland)',
'Aeronautical engineer',
'Agricultural consultant',
'Agricultural manager',
'Aid worker/humanitarian worker',
'Air traffic controller',

```

‘Airline cabin crew’,
‘Amenity horticulturist’,
‘Analytical chemist’,
‘Animal nutritionist’,
‘Animator’,
‘Applications programmer’,
‘Archaeologist’,
‘Architect’,
‘Architectural technologist’,
‘Archivist’,
‘Armed forces officer’,
‘Aromatherapist’,
‘Art therapist’,
‘Arts administrator’,
‘Auditor’,
‘Automotive engineer’,
‘Barristers clerk’,
‘Barrister’,
‘Bilingual secretary’,
‘Biomedical engineer’,
‘Biomedical scientist’,
‘Broadcasting presenter’,
‘Building control officer/surveyor’,
‘Building services engineer’,
‘Building surveyor’,
‘Camera operator’,
‘Careers adviser (higher education)’,
‘Careers adviser’,
‘Careers consultant’,
‘Cartographer’,
‘Catering manager’,
‘Charities administrator’,
‘Charities fundraiser’,
‘Chemical (process) engineer’,
‘Child psychotherapist’,
‘Childrens nurse’,
‘Chiropractor’,

'Civil engineer',
'Civil Service administrator',
'Clinical biochemist',
'Clinical cytogeneticist',
'Clinical microbiologist',
'Clinical molecular geneticist',
'Clinical research associate',
'Clinical scientist - tissue typing',
'Clothing and textile technologist',
'Colour technologist',
'Commercial horticulturist',
'Commissioning editor',
'Commissioning engineer',
'Commodity broker',
'Communications engineer',
'Community arts worker',
'Community education officer',
'Community worker',
'Company secretary',
'Computer sales support',
'Conference organiser',
'Construction manager',
'Consumer rights adviser',
'Control and instrumentation engineer',
'Corporate banker',
'Corporate treasurer',
'Counsellor',
'Courier/tour guide',
'Court reporter/verbatim reporter',
'Credit analyst',
'Crown Prosecution Service lawyer',
'Curator',
'Customs and excise officer',
'Customs officer',
'Dance movement therapist',
'Database administrator',
'Debt/finance adviser',

‘Dental hygienist’,
‘Dentist’,
‘Dietitian’,
‘Diplomatic service’,
‘Doctor (general practitioner, GP)’,
‘Doctor (hospital)’,
‘Dramatherapist’,
‘Economist’,
‘Editorial assistant’,
‘Education administrator’,
‘Electrical engineer’,
‘Employment advice worker’,
‘Energy conservation officer’,
‘Engineering geologist’,
‘Environmental education officer’,
‘Environmental health officer’,
‘Environmental manager’,
‘Equal opportunities officer’,
‘Ergonomist’,
‘Estate agent’,
‘European Commission administrators’,
‘Exhibition display designer’,
‘Exhibition organiser’,
‘Exploration geologist’,
‘Facilities manager’,
‘Field trials officer’,
‘Financial manager’,
‘Firefighter’,
‘Fisheries officer’,
‘Fitness centre manager’,
‘Food scientist’,
‘Food technologist’,
‘Forensic scientist’,
‘Geographical information systems manager’,
‘Geomatics/land surveyor’,
‘Government lawyer’,
‘Government research officer’,

‘Graphic designer’,
‘Health and safety adviser’,
‘Health and safety inspector’,
‘Health promotion specialist’,
‘Health service manager’,
‘Health visitor’,
‘Herbalist’,
‘Heritage manager’,
‘Higher education administrator’,
‘Higher education advice worker’,
‘Homeless worker’,
‘Horticultural consultant’,
‘Hotel manager’,
‘Housing adviser’,
‘Hydrologist’,
‘Illustrator’,
‘Immigration officer’,
‘Immunologist’,
‘Industrial/product designer’,
‘Information scientist’,
‘Information systems manager’,
‘Information technology/software trainers’,
‘Insurance broker’,
‘Insurance claims inspector’,
‘Insurance risk surveyor’,
‘Insurance underwriter’,
‘Interpreter’,
‘Investment analyst’,
‘Investment banker - corporate finance’,
‘Investment banker - operations’,
‘Investment banker’,
‘Investment fund manager’,
‘IT consultant’,
‘IT technical support officer’,
‘Journalist’,
‘Land-based engineer’,
‘Landscape architect’,

'Learning disability nurse',
'Learning mentor',
'Lecturer (adult education)',
'Lecturer (further education)',
'Lecturer (higher education)',
'Legal executive',
'Leisure centre manager',
'Licensed conveyancer',
'Local government administrator',
'Local government lawyer',
'Logistics/distribution manager',
'Loss adjuster (chartered)',
'Magazine features editor',
'Magazine journalist',
'Maintenance engineer',
'Management accountant',
'Management consultant (strategic)',
'Management consultant',
'Manufacturing engineer',
'Marine scientist',
'Market research executive',
'Marketing account manager',
'Marketing assistant',
'Marketing executive',
'Materials engineer',
'Materials specialist',
'Mechanical engineer',
'Media analyst',
'Media buyer',
'Media planner',
'Medical physicist',
'Medical representative',
'Mental health nurse',
'Metallurgist',
'Meteorologist',
'Microbiologist',
'Midwife',

‘Mining engineers’,
‘Multimedia programmer’,
‘Multimedia specialists’,
‘Museum education officer’,
‘Museum/gallery exhibition officer’,
‘Music therapist’,
‘Nature conservation officer’,
‘Naval architect’,
‘Network administrator’,
‘Nurse’,
‘Nutritional therapist’,
‘Nutritionist’,
‘Occupational therapist’,
‘Oceanographer’,
‘Office manager’,
‘Operational researcher’,
‘Orthoptist’,
‘Outdoor pursuits manager’,
‘Packaging technologist’,
‘Paediatric nurse’,
‘Paramedic’,
‘Patent attorney’,
‘Patent examiner’,
‘Pension scheme manager’,
‘Personal assistant’,
‘Personnel officer’,
‘Petroleum engineer’,
‘Pharmacologist’,
‘Photographer’,
‘Physiotherapist’,
‘Picture researcher’,
‘Planning surveyor’,
‘Planning technician’,
‘Plant breeder’,
‘Police officer’,
‘Political party agent’,
‘Political party research officer’,

'Practice nurse',
'Press photographer',
'Press sub-editor',
'Prison officer',
'Private music teacher',
'Probation officer',
'Product development scientist',
'Production manager',
'Programme researcher',
'Project manager',
'Psychologist (clinical)',
'Psychologist (educational)',
'Psychotherapist',
'Public affairs consultant (lobbyist)',
'Public affairs consultant (research)',
'Public house manager',
'Public librarian',
'Public relations (PR) officer',
'Quality assurance manager',
'Quantity surveyor',
'Race relations worker',
'Records manager',
'Recruitment consultant',
'Recycling officer',
'Regulatory affairs officer',
'Research scientist',
'Restaurant manager',
'Retail banker',
'Retail buyer',
'Retail manager',
'Retail merchandiser',
'Retail pharmacist',
'Sales executive',
'Scene of crime officer',
'Scientific laboratory technician',
'Secretary',
'Seismic interpreter',

‘Site engineer’,
‘Site manager’,
‘Social researcher’,
‘Social worker’,
‘Software engineer’,
‘Soil scientist’,
‘Solicitor’,
‘Speech and language therapist’,
‘Sports coach’,
‘Sports development officer’,
‘Sports therapist’,
‘Statistician’,
‘Stockbroker’,
‘Structural engineer’,
‘Surveyor’,
‘Systems analyst’,
‘Systems designer/builder’,
‘Systems developer’,
‘Systems programmer’,
‘Tax adviser’,
‘Tax inspector’,
‘Teacher (nursery/pre-school)’,
‘Teacher (primary)’,
‘Teacher (secondary)’,
‘Teacher (special educational needs)’,
‘Teaching/classroom assistant’,
‘Technical author’,
‘Technical sales engineer’,
‘TEFL/TESL teacher’,
‘Television production assistant’,
‘Tour operator’,
‘Tour/holiday representative’,
‘Tourism officer’,
‘Tourist information manager’,
‘Town and country planner’,
‘Toxicologist’,
‘Trade union research officer’,

```

‘Trader’,
‘Trading standards officer’,
‘Training and development officer’,
‘Translator’,
‘Transportation planner’,
‘Travel agent’,
‘TV/film/theatre set designer’,
‘Urban general practice surveyor’,
‘Veterinary nurse’,
‘Veterinary surgeon’,
‘Volunteer work organiser’,
‘Warehouse manager’,
‘Waste disposal officer’,
‘Water conservation officer’,
‘Water engineer’,
‘Web designer’,
‘Welfare rights adviser’,
‘Writer’,
‘Youth worker’), nrow(full_results), replace = TRUE)
)

```

```

full_results <- data.frame(
  full_results
  , employment_history = sample(c(
‘Academic librarian,Social worker,Television production assistant’,
‘Accountant (public finance, chartered),Archivist,Translator’,
‘Accountant (tax, chartered),Courier/tour guide,Production manager’,
‘Accountant,Tax adviser,Broadcasting presenter’,
‘Accounting technician,Meteorologist,Photographer’,
‘Actuary,Exhibition display designer,Oceanographer’,
‘Adult nurse,Trade union research officer,Advice worker’,
‘Advertising account executive,Editorial assistant,Network administrator’,
‘Advertising account planner,Political party research officer,Cartographer’,
‘Advertising copywriter,Practice nurse,Accountant’,
‘Advice worker,Toxicologist,Horticultural consultant’,
‘Advocate (Scotland),Investment fund manager,Medical physicist’,
‘Aeronautical engineer,Information scientist,Recycling officer’,

```

'Agricultural consultant,Legal executive,Hydrologist',
 'Agricultural manager,Environmental education officer,Transportation planner',
 'Aid worker/humanitarian worker,Town and country planner,Learning mentor',
 'Air traffic controller,Consumer rights adviser,Fitness centre manager',
 'Airline cabin crew,Civil Service administrator,Psychologist (educational)',
 'Amenity horticulturist,Clinical molecular geneticist,Information scientist',
 'Analytical chemist,Insurance broker,Building services engineer',
 'Animal nutritionist,Nutritionist,Sports therapist',
 'Animator,Charities administrator,Product development scientist',
 'Applications programmer,Learning mentor,Social researcher',
 'Archaeologist,Outdoor pursuits manager,Automotive engineer',
 'Architect,Sports development officer,Tourist information manager',
 'Architectural technologist,Dance movement therapist,Health visitor',
 'Archivist,Scene of crime officer,Probation officer',
 'Armed forces officer,Environmental health officer,Tour operator',
 'Aromatherapist,Fisheries officer,Engineering geologist',
 'Art therapist,Child psychotherapist,Local government administrator',
 'Arts administrator,Oceanographer,Accounting technician',
 'Auditor,Construction manager,Travel agent',
 'Automotive engineer,Multimedia specialists,Facilities manager',
 'Barristers clerk,Biomedical engineer,Environmental education officer',
 'Barrister,Health service manager,Market research executive',
 'Bilingual secretary,Recruitment consultant,Doctor (hospital)',
 'Biomedical engineer,Packaging technologist,Tour/holiday representative',
 'Biomedical scientist,IT technical support officer,Investment fund manager',
 'Broadcasting presenter,Adult nurse,Commissioning engineer',
 'Building control officer/surveyor,Doctor (hospital),Courier/tour guide',
 'Building services engineer,Nurse,Tax adviser',
 'Building surveyor,Dental hygienist,Childrens nurse',
 'Camera operator,Facilities manager,Teaching/classroom assistant',
 'Careers adviser (higher education),Retail pharmacist,Exploration geologist',
 'Careers adviser,Project manager,Mining engineers',
 'Careers consultant,Materials engineer,Architect',
 'Cartographer,Firefighter,Illustrator',
 'Catering manager,Crown Prosecution Service lawyer,Marketing assistant',
 'Charities administrator,Race relations worker,Marketing executive',
 'Charities fundraiser,Commodity broker,Stockbroker',

'Chemical (process) engineer,Political party agent,Web designer',
 'Child psychotherapist,Barristers clerk,Forensic scientist',
 'Childrens nurse,Software engineer,Speech and language therapist',
 'Chiropractor,Corporate banker,Project manager',
 'Civil engineer,Secretary,Picture researcher',
 'Civil Service administrator,Journalist,Air traffic controller',
 'Clinical biochemist,Airline cabin crew,Public relations (PR) officer',
 'Clinical cytogeneticist,Dentist,Outdoor pursuits manager',
 'Clinical microbiologist,Patent examiner,Volunteer work organiser',
 'Clinical molecular geneticist,Site manager,Personnel officer',
 'Clinical research associate,Colour technologist,Public house manager',
 'Clinical scientist - tissue typing,Information systems manager,Auditor',
 'Clothing and textile technologist,Hotel manager,Energy conservation officer',
 'Colour technologist,Marketing executive,Software engineer',
 'Commercial horticulturist,Mental health nurse,Investment banker & operations',
 'Commissioning editor,Advertising copywriter,Database administrator',
 'Commissioning engineer,Licensed conveyancer,Community arts worker',
 'Commodity broker,Customs officer,Management accountant',
 'Communications engineer,Immunologist,Catering manager',
 'Community arts worker,Magazine features editor,Information technology/software trainers',
 'Community education officer,TEFL/TESL teacher,Regulatory affairs officer',
 'Community worker,Building services engineer,Health and safety inspector',
 'Company secretary,Auditor,Trader',
 'Computer sales support,Engineering geologist,Retail banker',
 'Conference organiser,Aromatherapist,Camera operator',
 'Construction manager,Careers adviser,Advertising copywriter',
 'Consumer rights adviser,Sports coach,Solicitor',
 'Control and instrumentation engineer,Hydrologist,Retail merchandiser',
 'Corporate banker,Production manager,TEFL/TESL teacher',
 'Corporate treasurer,Orthoptist,Water engineer',
 'Counsellor,Metallurgist,Credit analyst',
 'Courier/tour guide,Youth worker,Electrical engineer',
 'Court reporter/verbatim reporter,Learning disability nurse,Immigration officer',
 'Credit analyst,Cartographer,Technical sales engineer',
 'Crown Prosecution Service lawyer,Television production assistant,Child psychotherapist',
 'Curator,Higher education administrator,Town and country planner',
 'Customs and excise officer,Probation officer,Clinical scientist - tissue typing',

‘Customs officer,Social researcher,Editorial assistant’,
 ‘Dance movement therapist,Insurance risk surveyor,Orthoptist’,
 ‘Database administrator,Stockbroker,Civil engineer’,
 ‘Debt/finance adviser,Product development scientist,Microbiologist’,
 ‘Dental hygienist,Estate agent,Health service manager’,
 ‘Dentist,Tax inspector,Marketing account manager’,
 ‘Dietitian,Mining engineers,Environmental health officer’,
 ‘Diplomatic service,Volunteer work organiser,Government research officer’,
 ‘Doctor (general practitioner, GP),Landscape architect,Construction manager’,
 ‘Doctor (hospital),Investment banker - corporate finance,Housing adviser’,
 ‘Dramatherapist,Retail buyer,Environmental manager’,
 ‘Economist,Loss adjuster (chartered),Court reporter/verbatim reporter’,
 ‘Editorial assistant,Health visitor,Applications programmer’,
 ‘Education administrator,Urban general practice surveyor,Health and safety adviser’,
 ‘Electrical engineer,Systems developer,Counsellor’,
 ‘Employment advice worker,Restaurant manager,Archaeologist’,
 ‘Energy conservation officer,Psychologist (clinical)s,Clinical research associate’,
 ‘Engineering geologist,Advertising account planner,Planning technician’,
 ‘Environmental education officer,Computer sales support,Crown Prosecution Service lawyer’,
 ‘Environmental health officer,Translator,Communications engineer’,
 ‘Environmental manager,Office manager,Graphic designer’,
 ‘Equal opportunities officer,Museum education officer,Education administrator’,
 ‘Ergonomist,Personnel officer,Management consultant’,
 ‘Estate agent,Operational researcher,Financial manager’,
 ‘European Commission administrators,Aeronautical engineer,Materials engineer’,
 ‘Exhibition display designer,Marketing account manager,Clinical biochemist’,
 ‘Exhibition organiser,Community education officer,Tourism officer’,
 ‘Exploration geologist,Seismic interpreter,Dental hygienist’,
 ‘Facilities manager,Music therapist,Magazine features editor’,
 ‘Field trials officer,Housing adviser,Investment banker’,
 ‘Financial manager,Waste disposal officer,Herbalist’,
 ‘Firefighter,Electrical engineer,Public affairs consultant (research)’,
 ‘Fisheries officer,Credit analyst,Political party research officer’,
 ‘Fitness centre manager,Mechanical engineer,Careers consultant’,
 ‘Food scientist,Barrister,Consumer rights adviser’,
 ‘Food technologist,Retail banker,Licensed conveyancer’,
 ‘Forensic scientist,Systems designer/builder,Government lawyer’,

'Geographical information systems manager,Maintenance engineer,Commercial horticulturist',
 'Geomatics/land surveyor,Teacher (nursery/pre-school),Mechanical engineer',
 'Government lawyer,Accounting technician,Lecturer (adult education)',
 'Government research officer,Soil scientist,Loss adjuster (chartered)',
 'Graphic designer,Armed forces officer,Research scientist',
 'Health and safety adviser,Building surveyor,Geomatics/land surveyor',
 'Health and safety inspector,Quantity surveyor,Doctor (general practitioner, GP)',
 'Health promotion specialist,Teacher (secondary),Dietitian',
 'Health service manager,Lecturer (further education),Geographical information systems manager',
 'Health visitor,Leisure centre manager,Pension scheme manager',
 'Herbalist,Economist,Community worker',
 'Heritage manager,Museum/gallery exhibition officer,Debt/finance adviser',
 'Higher education administrator,Retail merchandiser,Higher education advice worker',
 'Higher education advice worker,Dramatherapist,Social worker',
 'Homeless worker,Land-based engineer,Immunologist',
 'Horticultural consultant,Careers adviser (higher education),Adult nurse',
 'Hotel manager,Chemical (process) engineer,Advertising account executive',
 'Housing adviser,Management consultant (strategic),Company secretary',
 'Hydrologist,Agricultural consultant,Dance movement therapist',
 'Illustrator,Paediatric nurse,Retail pharmacist',
 'Immigration officer,Physiotherapist,Veterinary surgeon',
 'Immunologist,Air traffic controller,Amenity horticulturist',
 'Industrial/product designer,Teacher (primary),Veterinary nurse',
 'Information scientist,Welfare rights adviser,Careers adviser (higher education)',
 'Information systems manager,Pharmacologist,Health promotion specialist',
 'Information technology/software trainers,Diplomatic service,Heritage manager',
 'Insurance broker,Tour operator,Sales executive',
 'Insurance claims inspector,Midwife,Interpreter',
 'Insurance risk surveyor,Press sub-editor,Armed forces officer',
 'Insurance underwriter,Clinical microbiologist,Paramedic',
 'Interpreter,Geographical information systems manager,Tax inspector',
 'Investment analyst,Control and instrumentation engineer,Midwife',
 'Investment banker - corporate finance,Catering manager,Systems developer',
 'Investment banker - operations,Camera operator,Bilingual secretary',
 'Investment banker,Architect,Sports coach',
 'Investment fund manager,Exhibition organiser,Race relations worker',

'IT consultant,Medical physicist,Pharmacologist',
 'IT technical support officer,Biomedical scientist,Animator',
 'Journalist,Heritage manager,Marine scientist',
 'Land-based engineer,Government lawyer,Nutritional therapist',
 'Landscape architect,Press photographer,Medical representative',
 'Learning disability nurse,Trader,Aeronautical engineer',
 'Learning mentor,Childrens nurse,Economist',
 'Lecturer (adult education),Teaching/classroom assistant,Clinical microbiologist',
 'Lecturer (further education),Corporate treasurer,Meteorologist',
 'Lecturer (higher education),Travel agent,Plant breeder',
 'Legal executive,European Commission administrators,Charities administrator',
 'Leisure centre manager,Patent attorney,Insurance risk surveyor',
 'Licensed conveyancer,Public relations (PR) officer,Food technologist',
 'Local government administrator,Magazine journalist,Actuary',
 'Local government lawyer,Regulatory affairs officer,Information systems manager',
 'Logistics/distribution manager,Graphic designer,Animal nutritionist',
 'Loss adjuster (chartered),Accountant (tax, chartered),Seismic interpreter',
 'Magazine features editor,Financial manager,Recruitment consultant',
 'Magazine journalist,Clinical biochemist,Arts administrator',
 'Maintenance engineer,Homeless worker,Hotel manager',
 'Management accountant,Transportation planner,Media buyer',
 'Management consultant (strategic),Technical sales engineer,Firefighter',
 'Management consultant,Tourist information manager,Commissioning editor',
 'Manufacturing engineer,Community worker,Mental health nurse',
 'Marine scientist,Prison officer,Academic librarian',
 'Market research executive,Trading standards officer,Media planner',
 'Marketing account manager,Technical author,Quantity surveyor',
 'Marketing assistant,Academic librarian,IT technical support officer',
 'Marketing executive,Teacher (special educational needs),European Commission administrators',
 'Materials engineer,Site engineer,Psychotherapist',
 'Materials specialist,Surveyor,Quality assurance manager',
 'Mechanical engineer,Doctor (general practitioner, GP),Chiropractor',
 'Media analyst,Art therapist,Programme researcher',
 'Media buyer,Management accountant,Computer sales support',
 'Media planner,Advertising account executive,Sports development officer',
 'Medical physicist,Fitness centre manager,Petroleum engineer',

'Medical representative,Geomatics/land surveyor,Paediatric nurse',
 'Mental health nurse,Database administrator,Dentist',
 'Metallurgist,Solicitor,Building surveyor',
 'Meteorologist,Commissioning editor,Leisure centre manager',
 'Microbiologist,Architectural technologist,Employment advice worker',
 'Midwife,Higher education advice worker,Fisheries officer',
 'Mining engineers,Manufacturing engineer,Estate agent',
 'Multimedia programmer,Systems analyst,Scene of crime officer',
 'Multimedia specialists,Scientific laboratory technician,Soil scientist',
 'Museum education officer,TV/film/theatre set designer,Analytical chemist',
 'Museum/gallery exhibition officer,Actuary,Diplomatic service',
 'Music therapist,Archaeologist,Physiotherapist',
 'Nature conservation officer,Marketing assistant,Agricultural consultant',
 'Naval architect,Public librarian,Manufacturing engineer',
 'Network administrator,Nutritional therapist,Private music teacher',
 'Nurse,Media analyst,Biomedical scientist',
 'Nutritional therapist,Training and development officer,TV/film/theatre set designer',
 'Nutritionist,Personal assistant,Museum education officer',
 'Occupational therapist,Forensic scientist,Metallurgist',
 'Oceanographer,Horticultural consultant,Legal executive',
 'Office manager,Animator,Lecturer (further education)',
 'Operational researcher,Speech and language therapist,Learning disability nurse',
 'Orthoptist,Water conservation officer,Curator',
 'Outdoor pursuits manager,Investment banker & operations,Statistician',
 'Packaging technologist,Media planner,Political party agent',
 'Paediatric nurse,Education administrator,Exhibition organiser',
 'Paramedic,Communications engineer,Corporate treasurer',
 'Patent attorney,Warehouse manager,Insurance underwriter',
 'Patent examiner,Field trials officer,Biomedical engineer',
 'Pension scheme manager,Psychologist (educational),Press sub-editor',
 'Personal assistant,Writer,Airline cabin crew',
 'Personnel officer,Lecturer (adult education),Prison officer',
 'Petroleum engineer,Statistician,Logistics/distribution manager',
 'Pharmacologist,Clothing and textile technologist,Conference organiser',
 'Photographer,Records manager,Psychologist (clinical)s',
 'Physiotherapist,Food scientist,Multimedia programmer',
 'Picture researcher,Amenity horticulturist,Systems programmer',

'Planning surveyor,Health promotion specialist,Architectural technologist',
 'Planning technician,Medical representative,Homeless worker',
 'Plant breeder,Private music teacher,Advertising account planner',
 'Police officer,Charities fundraiser,Museum/gallery exhibition officer',
 'Political party agent,Public house manager,Control and instrumentation engineer',
 'Political party research officer,Company secretary,Waste disposal officer',
 'Practice nurse,Market research executive,Field trials officer',
 'Press photographer,Retail manager,Youth worker',
 'Press sub-editor,Automotive engineer,Charities fundraiser',
 'Prison officer,Occupational therapist,Press photographer',
 'Private music teacher,Materials specialist,Magazine journalist',
 'Probation officer,Food technologist,Patent attorney',
 'Product development scientist,Industrial/product designer,Civil Service administrator',
 'Production manager,Accountant (public finance, chartered),Exhibition display designer',
 'Programme researcher,Curator,Welfare rights adviser',
 'Project manager,Quality assurance manager,Land-based engineer',
 'Psychologist (clinical)s,Local government administrator,Investment analyst',
 'Psychologist (educational),Ergonomist,Food scientist',
 'Psychotherapist,Recycling officer,Trade union research officer',
 'Public affairs consultant (lobbyist),Veterinary nurse,Insurance claims inspector',
 'Public affairs consultant (research),Analytical chemist,Archivist',
 'Public house manager,Counsellor,Corporate banker',
 'Public librarian,Programme researcher,Music therapist',
 'Public relations (PR) officer,Network administrator,Community education officer',
 'Quality assurance manager,Nature conservation officer,Customs officer',
 'Quantity surveyor,Planning technician,Systems designer/builder',
 'Race relations worker,Petroleum engineer,Retail buyer',
 'Records manager,Veterinary surgeon,Insurance broker',
 'Recruitment consultant,Customs and excise officer,Office manager',
 'Recycling officer,Chiropractor,Clothing and textile technologist',
 'Regulatory affairs officer,Agricultural manager,Packaging technologist',
 'Research scientist,Accountant,Site manager',
 'Restaurant manager,Psychotherapist,Advocate (Scotland)',
 'Retail banker,Water engineer,Chemical (process) engineer',
 'Retail buyer,Sports therapist,Occupational therapist',
 'Retail manager,Civil engineer,Water conservation officer',
 'Retail merchandiser,IT consultant,Site engineer',

'Retail pharmacist,Clinical research associate,Art therapist',
 'Sales executive,Plant breeder,Equal opportunities officer',
 'Scene of crime officer,Insurance underwriter,Records manager',
 'Scientific laboratory technician,Systems programmer,Media analyst',
 'Secretary,Court reporter/verbatim reporter,Building control officer/surveyor',
 'Seismic interpreter,Broadcasting presenter,Secretary',
 'Site engineer,Aid worker/humanitarian worker,Retail manager',
 'Site manager,Marine scientist,Barristers clerk',
 'Social researcher,Logistics/distribution manager,Nutritionist',
 'Social worker,Sales executive,Barrister',
 'Software engineer,Energy conservation officer,Lecturer (higher education)',
 'Soil scientist,Equal opportunities officer,Patent examiner',
 'Solicitor,Picture researcher,Dramatherapist',
 'Speech and language therapist,Lecturer (higher education),Police officer',
 'Sports coach,Commissioning engineer,Public librarian',
 'Sports development officer,Government research officer,Public affairs consultant (lobbyist)',
 'Sports therapist,Tour/holiday representative,Restaurant manager',
 'Statistician,Investment analyst,Commodity broker',
 'Stockbroker,Applications programmer,Writer',
 'Structural engineer,Building control officer/surveyor,Landscape architect',
 'Surveyor,Microbiologist,Colour technologist',
 'Systems analyst,Careers consultant,Maintenance engineer',
 'Systems designer/builder,Public affairs consultant (lobbyist),Practice nurse',
 'Systems developer,Animal nutritionist,Nurse',
 'Systems programmer,Pension scheme manager,Training and development officer',
 'Tax adviser,Arts administrator,Urban general practice surveyor',
 'Tax inspector,Insurance claims inspector,Systems analyst',
 'Teacher (nursery/pre-school),Investment banker,Toxicologist',
 'Teacher (primary),Media buyer,Nature conservation officer',
 'Teacher (secondary),Police officer,Teacher (special educational needs)',
 'Teacher (special educational needs),Information technology/software trainers,Operational researcher',
 'Teaching/classroom assistant,Naval architect,Structural engineer',
 'Technical author,Clinical scientist - tissue typing,Scientific laboratory technician',
 'Technical sales engineer,Herbalist,Management consultant (strategic)',
 'TEFL/TESL teacher,Conference organiser,Agricultural manager',
 'Television production assistant,Web designer,Personal assistant',

```

'Tour operator,Exploration geologist,Warehouse manager',
'Tour/holiday representative,Photographer,Surveyor',
'Tourism officer,Employment advice worker,Local government lawyer',
'Tourist information manager,Multimedia programmer,Investment banker - corporate finance',
'Town and country planner,Bilingual secretary,Accountant (tax, chartered)',
'Toxicologist,Structural engineer,Higher education administrator',
'Trade union research officer,Dietitian,Careers adviser',
'Trader,Interpreter,Multimedia specialists',
'Trading standards officer,Public affairs consultant (research),Journalist',
'Training and development officer,Advocate (Scotland),Trading standards officer',
'Translator,Commercial horticulturist,Teacher (primary)',
'Transportation planner,Paramedic,Materials specialist',
'Travel agent,Advice worker,Customs and excise officer',
'TV/film/theatre set designer,Health and safety adviser,Naval architect',
'Urban general practice surveyor,Health and safety inspector,Clinical molecular geneticist',
'Veterinary nurse,Immigration officer,Technical author',
'Veterinary surgeon,Debt/finance adviser,Teacher (nursery/pre-school)',
'Volunteer work organiser,Management consultant,Teacher (secondary)',
'Warehouse manager,Environmental manager,IT consultant',
'Waste disposal officer,Community arts worker,Ergonomist',
'Water conservation officer,Illustrator,Aromatherapist',
'Water engineer,Research scientist,Industrial/product designer',
'Web designer,Clinical cytogeneticist,Accountant (public finance, chartered)',
'Welfare rights adviser,Planning surveyor,Aid worker/humanitarian worker',
'Writer,Local government lawyer,Clinical cytogeneticist',
'Youth worker,Tourism officer,Planning surveyor'), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
  full_results
  , accommodation_type
  = sample(c('owner', 'renting', 'boarding', 'shared_accommodation', 'living with parents'),
nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
  full_results
  , holiday_destinations
  = sample(c('Peru',

```

'Togo',
'Aruba',
'Benin',
'Burma',
'Chile',
'China',
'Costa',
'Czech',
'Egypt',
'Gabon',
'Ghana',
'Haiti',
'India',
'Italy',
'Japan',
'Kenya',
'Libya',
'Macau',
'Malta',
'Nauru',
'Nepal',
'Niger',
'Palau',
'Papua',
'Qatar',
'Saint',
'Saint',
'Saint',
'Samoa',
'Saudi',
'South',
'South',
'South',
'Spain',
'Sudan',
'Syria',
'Tonga',

'Yemen',
'Angola',
'Belize',
'Bhutan',
'Bosnia',
'Brazil',
'Brunei',
'Canada',
'Congo',
'Congo',
'Cyprus',
'France',
'Greece',
'Guinea',
'Guyana',
'Israel',
'Jordan',
'Korea',
'Korea',
'Kosovo',
'Kuwait',
'Latvia',
'Malawi',
'Mexico',
'Monaco',
'Norway',
'Panama',
'Poland',
'Russia',
'Rwanda',
'Serbia',
'Sierra',
'Sweden',
'Taiwan',
'Turkey',
'Tuvalu',
'Uganda',

'Zambia',
'Albania',
'Algeria',
'Andorra',
'Antigua',
'Armenia',
'Austria',
'Bahrain',
'Belarus',
'Belgium',
'Bolivia',
'Burkina',
'Burundi',
'Central',
'Comoros',
'Croatia',
'Curacao',
'Denmark',
'Ecuador',
'Eritrea',
'Estonia',
'Finland',
'Gambia',
'Georgia',
'Germany',
'Grenada',
'Hungary',
'Iceland',
'Ireland',
'Jamaica',
'Lebanon',
'Lesotho',
'Liberia',
'Moldova',
'Morocco',
'Namibia',
'Nigeria',

'Romania',
'Senegal',
'Solomon',
'Somalia',
'Tunisia',
'Ukraine',
'Uruguay',
'Vanuatu',
'Vietnam',
'Bahamas',
'Barbados',
'Botswana',
'Bulgaria',
'Cambodia',
'Cameroon',
'Colombia',
'Djibouti',
'Dominica',
'Ethiopia',
'Honduras',
'Kiribati',
'Malaysia',
'Maldives',
'Marshall',
'Mongolia',
'Pakistan',
'Paraguay',
'Portugal',
'Slovakia',
'Slovenia',
'Suriname',
'Tanzania',
'Thailand',
'Trinidad',
'Zimbabwe',
'Argentina',
'Australia',

```

'Dominican',
'Guatemala',
'Indonesia',
'Lithuania',
'Macedonia',
'Mauritius',
'Nicaragua',
'Singapore',
'Swaziland',
'Venezuela',
'Azerbaijan',
'Bangladesh',
'Equatorial',
'Kazakhstan',
'Kyrgyzstan',
'Luxembourg',
'Madagascar',
'Mauritania',
'Micronesia',
'Montenegro',
'Mozambique',
'Seychelles',
'Tajikistan',
'Uzbekistan',
'Afghanistan',
'Netherlands',
'Netherlands',
'Palestinian',
'Philippines',
'Switzerland',
'Timor-Leste',
'Turkmenistan',
'Guinea-Bissau',
'Liechtenstein'), nrow(full_results), replace = TRUE)
)
full_results <- subset(full_results, select = -c(insurances.1) )
full_results <- data.frame(

```

```

full_results
, mode_of_transport
= sample(c('drive', 'bus', 'train', 'walk', 'bike'), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
full_results
, passtimes
=
sample(c('rugby', 'cricket', 'ballet', 'gaming', 'tennis', 'netball', 'golf', 'painting', 'reading', 'cooking', 'cycling', 'sailing', 'triathlon', 'swimming', 'bushwalking'), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
full_results
, life_events_recent
= sample(c('1st child', '2nd child', '3rd child', 'engaged', 'married', 'divorced', 'separated', '1st job', 'promotion', 'bought 1st home', '1st grandchild', 'retired'), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
full_results
, life_events_forthcoming
= sample(c('1st child', '2nd child', '3rd child', 'engaged', 'married', 'divorced', 'separated', '1st job', 'promotion', 'bought 1st home', '1st grandchild', 'retired'), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
full_results
, life_events_excluded
= sample(c('1st child', '2nd child', '3rd child', 'engaged', 'married', 'divorced', 'separated', '1st job', 'promotion', 'bought 1st home', '1st grandchild', 'retired'), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
full_results
, vehicles
= sample(c('Ford',
'Vauxhall',
'Volkswagen',
'Nissan',

```

```

'Audi',
'Mercedes-Benz',
'BMW',
'Peugeot',
'Toyota',
'Citroen',
'Hyundai',
'Kia',
'Land',
'Renault',
'Skoda',
'Fiat',
'Honda',
'MINI',
'Mazda',
'SEAT',
'Volvo',
'Suzuki',
'Mitsubishi',
'Dacia',
'Jaguar',
'Lexus',
'Porsche',
'Jeep',
'Alfa'), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
full_results
, insurances
= sample(c(
'Compulsory Third Party Insurance (CTP)',
'General Property, Home and Contents',
'Pet Insurance',
'Travel',
'Marine Insurance',
'Medical Indemnity Insurance',
'Motor Vehicle'

```

```

), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
  full_results
  , financial_institution
  = sample(c(
    'Westpac',
    'CBA',
    'ANZ',
    'NAB'
  ), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
  full_results
  , super
  = sample(c(
    'Colonial First State FirstChoice Superannuation Trust',
    'State Public Sector Superannuation Scheme',
    'Retirement Wrap (BT)',
    'First State Superannuation Scheme',
    'AMP Superannuation Savings Trust',
    'UniSuper',
    'The Universal Super Scheme (MLC)',
    'Retail Employees Superannuation Trust (REST)',
    'Sunsuper Superannuation Fund',
    'OnePath Masterfund (ANZ)',
    'HESTA Industry',
    'Construction & Building Unions Superannuation (CBUS)',
    'Wealth Personal Superannuation & Pension Fund',
    'ASGARD Independence Plan Division Two',
    'Mercer Super Trust',
    'HOSTPLUS',
    'IOOF Portfolio Service Superannuation Fund',
    'Telstra Superannuation Scheme',
    'Public Sector Superannuation Scheme'
  ), nrow(full_results), replace = TRUE)
)

```

```

full_results <- data.frame(
  full_results
  , technology
  = sample(c(
'Apple',      'Samsung      LG',  'Motorola', 'HTC'
  ), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
  full_results
  , utilities
  = sample(c(
'Origin', 'Energy Australia', 'AGL', 'Genesis'
  ), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
  full_results
  , financial_advisor
  = sample(c(
'Quantum Financial',
'Fortnum Financial Advisers',
'Cartwright Brown & Co',
'Westpac Financial Planner',
'Navigate Wealth'
  ), nrow(full_results), replace = TRUE)
)
full_results <- data.frame(
  full_results
  , subscriptions
  = sample(c(
'The Australian Womens Weekly,Australian House and Garden,Australian House and Garden',
'Better Homes And Gardens,Club Marine,Thats Life!',
'Womans Day,TV Week,Woolworths Good Taste',
'New Idea,Famous,Who',
'Thats Life!,Recipes+,Australian Geographic',
'Readers Digest,Better Homes And Gardens,Readers Digest',
'Take 5,Woolworths Good Taste,Take 5',
'Super Food Ideas,Womans Day,Womens Health',

```

'TV Week,Marie Claire,Super Food Ideas',
 'Who,Womens Health,Delicious',
 'Delicious,Who,Womans Day',
 'Australian House and Garden,New Idea,The Australian Womens Weekly',
 'Recipes+,Thats Life!,Dolly',
 'Cosmopolitan,Dolly,Club Marine',
 'NW,OK!,OK!',
 'Donna Hay,Australian Geographic,Donna Hay',
 'Womens Health,Donna Hay,New Idea',
 'Marie Claire,Family Circle,Family Circle',
 'Dolly,Readers Digest,Recipes+',
 'OK!,Cosmopolitan,Better Homes And Gardens',
 'Famous,Super Food Ideas,Marie Claire',
 'Club Marine,The Australian Womens Weekly,NW',
 'Family Circle,NW,Famous',
 'Woolworths Good Taste,Delicious,TV Week',
 'Australian Geographic,Take 5,Cosmopolitan'

), nrow(full_results), replace = TRUE)

)

full_results <- data.frame(

full_results

, real_estate

= sample(c(

'owner occupier',

'investment property',

'none'

), nrow(full_results), replace = TRUE)

)

full_results <- data.frame(

full_results

, pay_tv

= sample(c(

'Foxtel',

'Fetch TV',

'Netflix',

'none'

), nrow(full_results), replace = TRUE)

)

```
full_results <- data.frame(  
  full_results  
  , share_portfolio  
  = sample(c(  
    'Australian',  
    'International',  
    'diverserfied',  
    'none'  
  ), nrow(full_results), replace = TRUE)  
)
```

)

```
full_results <- data.frame(  
  full_results  
  , telco  
  = sample(c(  
    'Telstra',  
    'Optus',  
    'Vodafone'  
  ), nrow(full_results), replace = TRUE)  
)
```

)

```
full_results <- data.frame(  
  full_results  
  , political_leanings  
  = sample(c(  
    'Labor',  
    'Liberal',  
    'Greens',  
    'Coalition',  
    'National',  
    'Other'  
  ), nrow(full_results), replace = TRUE)  
)
```

)

```
full_results <- data.frame(  
  full_results  
  , issues  
  = sample(c(  
    'Labor',  
    'Liberal',  
    'Greens',  
    'Coalition',  
    'National',  
    'Other'  
  ), nrow(full_results), replace = TRUE)  
)
```

```

'timeliness',
'politeness',
'low fees',
'honesty'
), nrow(full_results), replace = TRUE)
)
#write full_results back to the postgres database

```

15 References

- Abel, F., Henze, N., Herder, E. & Krause, D. 2010, 'Interweaving Public User Profiles on the Web', in P. Bra, A. Kobsa & D. Chin (eds), *User Modeling, Adaptation, and Personalization*, vol. 6075, Springer Berlin Heidelberg, pp. 16-27.
- Abra 2016, viewed 27th October 2016, <<https://www.goabra.com/>>.
- ABS 1993, *Australian and New Zealand Standard Industrial Classification*, Australian Bureau of Statistics, viewed 15th Sept 2015, <<http://www.abs.gov.au/AUSSTATS/abs@.nsf/66f306f503e529a5ca25697e0017661f/7cd8aebba7225c4eca25697e0018faf3!OpenDocument>>.
- Acquisti, A. 2010, 'The economics of personal data and the economics of privacy', *Background Paper for OECD Joint WPISP-WPIE Roundtable*, vol. 1.
- Adomavicius, G., Bockstedt, J.C., Gupta, A. & Kauffman, R.J. 2008, 'Making sense of technology trends in the information technology landscape: A design science approach', *MIS Quarterly*, pp. 779-809.
- Ahrar, N. & Rahman, A.A. 2014, 'Review on DART model used in value co-creation for evaluation of e-service system', *International Journal of Management, IT and Engineering*, vol. 4, no. 2, pp. 409-15.
- Aitken, R., Ballantyne, D., Osborne, P. & Williams, J. 2006, 'Introduction to the special issue on the service-dominant logic of marketing: insights from The Otago Forum', *Marketing Theory*, vol. 6, no. 3, p. 275.
- Akaka, M.A., Vargo, S.L. & Lusch, R.F. 2013, 'The complexity of context: A service ecosystems approach for international marketing', *Journal of Marketing Research*, vol. 21, no. 4, pp. 1-20.
- Albatal, R., Gurrin, C., Jiang, Z., Yang, Y., Carthy, D. & Na, L. 2013, 'SenseSeer mobile-cloud-based Lifelogging framework', *Technology and Society (ISTAS), 2013 IEEE International Symposium on*, pp. 144-6.
- Aljafari, R. & Khazanachi, D. 2013, 'On the veridicality of claims in design science research', paper presented to the *46th Hawaii International Conference on System Sciences (HICSS)*, Hawaii.
- Almirall, E. & Casadesus-Masanell, R. 2010, 'Open versus closed innovation: A model of discovery and divergence', *Academy of management review*, vol. 35, no. 1, pp. 27-47.
- Alvesson, M. & Sandberg, J. 2011, 'Generating research questions through problematization', *Academy of Management Review*, vol. 36, no. 2, pp. 247-71.
- Anderson, R. 2001, 'Why information security is hard-an economic perspective', *Computer Security Applications Conference, 2001. ACSAC 2001. Proceedings 17th Annual*, IEEE, pp. 358-65.
- Aperjis, C. & Huberman, B. 2012, 'A Market for Unbiased Private Data: Paying Individuals According to their Privacy Attitudes', *Available at SSRN 2046861 (2012)*.

- Aral, S., Dellarocas, C. & Godes, D. 2013, 'Introduction to the special issue-social media and business transformation: A framework for research', *Information Systems Research*, vol. 24, no. 1, pp. 3-13.
- Ayers, D.J., Menachemi, N., Ramamonjiarivelo, Z., Matthews, M. & Brooks, R.G. 2009, 'Adoption of electronic medical records: the role of network effects', *Journal of Product & Brand Management*, vol. 18, no. 2, pp. 127-35.
- Bagust, P. 2014, 'Honey, They Stole My Flickr!' Social Network Typologies, Online Trust and Dissent, and the Monetisation of Immaterial Labour', *Media International Australia*, vol. 151, no. 1, pp. 89-96.
- Bakos, Y. & Bailey, J. 1997, 'An exploratory study of the emerging role of electronic intermediaries', *International Journal of Electronic Commerce*, vol. 1, no. 3, pp. 7-20.
- Baldwin, C.Y. & Clark, K.B. 2000, *Design rules: The power of modularity*, vol. 1, MIT press.
- Baldwin, C.Y. & Woodard, C.J. 2008, 'The architecture of platforms: a unified view', *Harvard Business School Finance Working Paper*, no. 09-034.
- Ballantyne, D. & Varey, R.J. 2006, 'Creating value-in-use through marketing interaction: the exchange logic of relating, communicating and knowing', *Marketing Theory*, vol. 6, no. 3, pp. 335-48.
- Baltar, F. & Brunet, I. 2012, 'Social research 2.0: virtual snowball sampling method using Facebook', *Internet Research*, vol. 22, no. 1, pp. 57-74.
- Bask, A., Merisalo-Rantanen, H. & Tuunanen, T. 2014, 'Developing a modular service architecture for E-store supply chains: The small-and medium-sized enterprise perspective', *Service Science*, vol. 6, no. 4, pp. 251-73.
- Basole, R.C. & Karla, J. 2011, 'On the evolution of mobile platform ecosystem structure and strategy', *Business & Information Systems Engineering*, vol. 3, no. 5, pp. 313-22.
- Bendapudi, N. & Leone, R.P. 2003, 'Psychological implications of customer participation in co-production', *Journal of Marketing*, vol. 67, no. 1, pp. 14-28.
- Benjamin, R. & Wigand, R. 1995, 'Electronic markets and virtual value chains on the information superhighway', *Sloan Management Review (Winter, 1995)*.
- Benndorf, V. & Normann, H.-T. 2014, 'The willingness to sell personal data', <http://hdl.handle.net/10419/100663>.
- Berinato, S. 2014, 'With Big Data Comes Big Responsibility', *Harvard Business Review*, vol. 92, no. 11, pp. 100-4.
- Berry, L.L. & Parasuraman, A. 1993, 'Building a new academic field—The case of services marketing', *Journal of Retailing*, vol. 69, no. 1, pp. 13-60.
- Blanchette, J. 2008, 'The little manual of API design', *Trolltech, Nokia*.
- Borlund, P. 2016, 'A study of the use of simulated work task situations in interactive information retrieval evaluations: A meta-evaluation', *Journal of Documentation*, vol. 72, no. 3, pp. 394-413.
- Bose, S. & Goswami, S. 2017, *RSentiment*, <https://CRAN.R-project.org/package=RSentiment>.
- Bowman, C. & Ambrosini, V. 2000, 'Value creation versus value capture: towards a coherent definition of value in strategy', *British Journal of Management*, vol. 11, no. 1, pp. 1-15.
- Brax, S.A., Bask, A., Hsuan, J. & Voss, C. 2017, 'Service modularity and architecture—an overview and research agenda', *International Journal of Operations & Production Management*, vol. 37, no. 6, pp. 686-702.
- Brickman-Bhutta, C. 2012, 'Not by the book: Facebook as a sampling frame', *Sociological Methods & Research*, p. 0049124112440795.
- Broekhuis, M., van Offenbeek, M. & van der Laan, M. 2017, 'What professionals consider when designing a modular service architecture', *International Journal of Operations & Production Management*, vol. 37, no. 6.
- Brustein, J. 2012, *Start-ups seek to help users put a price on their personal data*, http://www.nytimes.com/2012/02/13/technology/start-ups-aim-to-help-users-put-a-price-on-their-personal-data.html?_r=0.

- Calcagno, V. 2013, 'glmulti, Model selection and multimodel inference made easy', 1.0.7. edn, R package version 1.0.7. <http://CRAN.R-project.org/package=glmulti>.
- Cao, J., Basoglu, K.A., Sheng, H. & Lowry, P.B. 2015, 'A Systematic Review of Social Networks Research in Information Systems: Building a Foundation for Exciting Future Research', *Information Systems*, vol. 36, p. 37.
- Carrascal, J.P., Riederer, C., Erramilli, V., Cherubini, M. & de Oliveira, R. 2013, 'Your browsing behavior for a big mac: Economics of personal information online', *Proceedings of the 22nd international conference on World Wide Web*, International World Wide Web Conferences Steering Committee, pp. 189-200.
- Carù, A. & Cova, B. 2003, 'Revisiting consumption experience a more humble but complete view of the concept', *Marketing Theory*, vol. 3, no. 2, pp. 267-86.
- Cenamor, J., Usero, B. & Fernández, Z. 2013, 'The role of complementary products on platform adoption: Evidence from the video console market', *Technovation*, vol. 33, no. 12, pp. 405-16.
- Chandler, J.D. & Vargo, S.L. 2011, 'Contextualization and value-in-context: How context frames exchange', *Marketing Theory*, vol. 11, no. 1, pp. 35-49.
- Chang, H.H. & Chuang, S.-S. 2011, 'Social capital and individual motivations on knowledge sharing: Participant involvement as a moderator', *Information & Management*, vol. 48, no. 1, pp. 9-18.
- Checkland, P. 2013, 'Soft systems methodology', *Encyclopedia of Operations Research and Management Science*, Springer, pp. 1430-6.
- Chen, C.-J. & Hung, S.-W. 2010, 'To give or to receive? Factors influencing members' knowledge sharing and community promotion in professional virtual communities', *Information & Management*, vol. 47, no. 4, pp. 226-36.
- Chen, Y., Xu, X. & Zhu, L. 2012, 'Web Platform API Design Principles and Service Contract', *Software Engineering Conference (APSEC), 2012 19th Asia-Pacific*, vol. 1, IEEE, pp. 877-86.
- Chew, E. 2013a, 'Value Co-Creation in the Organizations of the Future', *Proceedings For the 9th European Conference on Management Leadership and Governance: ECMLG 2013*, Academic Conferences Limited, p. 16.
- Chew, E.K. 2013b, 'An Integrative Design Framework for New Service Development', paper presented to the *Australian Symposium on Service Research and Innovation*.
- Chew, E.K. 2015, 'iSIM: An integrated design method for commercializing service innovation', *Information Systems Frontiers*, pp. 1-22.
- Clarke, R. 2014, 'Persona missing, feared drowned: the digital persona concept, two decades later', *Information Technology & People*, vol. 27, no. 2, pp. 182-207.
- Claussen, J., Kretschmer, T. & Mayrhofer, P. 2013, 'The effects of rewarding user engagement: the case of facebook apps', *Information Systems Research*, vol. 24, no. 1, pp. 186-200.
- Coase, R.H. 1937, 'The nature of the firm', *economica*, vol. 4, no. 16, pp. 386-405.
- Conger, S., Pratt, J.H. & Loch, K.D. 2013, 'Personal information privacy and emerging technologies', *Information Systems Journal*, vol. 23, no. 5, pp. 401-17.
- Conner-Simons, A. 2015, *Web Inventor Tim Berners-Lee's Next Project: A Platform That Gives Users Control Of Their Data*, viewed 2nd May 2017, <http://www.csail.mit.edu/solid_mastercard_gift>.
- Constantinides, P., Henfridsson, O. & Parker, G.G. 2018, 'Introduction - Platforms and Infrastructures in the Digital Age', *Information Systems Research*, vol. 29, no. 2.
- Copeland, M.T. 1923, 'Relation of consumers' buying habits to marketing methods', *Harvard Business Review*, vol. 1, no. 3, pp. 282-9.
- Cormack, L. 2017, *We-Vibe sex toy manufacturer settles US class action for almost \$4 million*, Sydney Morning Herald, viewed 15th March 2017, <<http://www.smh.com.au/business/consumer-affairs/wevibe-sex-toy-manufacturer-settles-us-class-action-for-almost-4-million-20170315-guyely.html>>.

- Craig, R., Spyridopoulos, T., Tryfonas, T. & May, J. 2014, 'Soft systems methodology in net-centric cyber defence system development', *Systems, Man and Cybernetics (SMC), 2014 IEEE International Conference on*, IEEE, pp. 672-7.
- Crawford, K. & Gillespie, T. 2016, 'What is a flag for? Social media reporting tools and the vocabulary of complaint', *New Media & Society*, vol. 18, no. 3, pp. 410-28.
- Cross, N. 2001, 'Designerly Ways of Knowing: Design Discipline versus Design Science', *Design Issues*, vol. 17, no. 3, pp. 49-55.
- Crosscloud 2017, *Crosscloud*, viewed 25th Feb 2017, <<http://crosscloud.org/>>.
- Cusumano, M. 2010a, 'Cloud computing and SaaS as new computing platforms', *Communications of the ACM*, vol. 53, no. 4, pp. 27-9.
- Cusumano, M. 2010b, 'Technology strategy and management, the evolution of platform thinking', *Communications of the ACM*, vol. 53, no. 1, pp. 32-4.
- Cvrcek, D., Kumpost, M., Matyas, V. & Danezis, G. 2006, 'A study on the value of location privacy', *Proceedings of the 5th ACM workshop on privacy in electronic society*, ACM, pp. 109-18.
- Dalgic, T. & Leeuw, M. 1994, 'Niche Marketing Revisited: Concept, Applications and Some European Cases', *European Journal of Marketing*, vol. 28, no. 4, pp. 39-55.
- Danezis, G., Lewis, S. & Anderson, R.J. 2005, 'How much is location privacy worth?', *WEIS*, vol. 5, Citeseer.
- Day, G.S., Deighton, J., Narayandas, D., Gummesson, E., Hunt, S.D., Prahalad, C., Rust, R.T. & Shugan, S.M. 2004, 'Invited commentaries on "Evolving to a new dominant logic for marketing"', *Journal of Marketing*, vol. 68, no. 1, pp. 18-27.
- de Blok, C., Meijboom, B., Luijkx, K., Schols, J. & Schroeder, R. 2014, 'Interfaces in service modularity: a typology developed in modular health care provision', *Journal of Operations Management*, vol. 32, no. 4, pp. 175-89.
- de Montjoye, Y.A., Hidalgo, C.A., Verleysen, M. & Blondel, V.D. 2013, 'Unique in the Crowd: The privacy bounds of human mobility', *Scientific Reports*, vol. 3.
- de Montjoye, Y.A., Shmueli, E., Wang, S.S. & Pentland, A.S. 2014, 'openPDS: Protecting the Privacy of Metadata through SafeAnswers', *PloS one*, vol. 9, no. 7, p. e98790.
- de Montjoye, Y.A., Wang, S.S., Pentland, A., Anh, D.T.T. & Datta, A. 2012, 'On the Trusted Use of Large-Scale Personal Data', *IEEE Data Eng. Bull.*, vol. 35, no. 4, pp. 5-8.
- de Reuver, M., Haaker, T., Nikayin, F. & Kosman, R. 2015, 'Designing Viable Multi-sided Data Platforms: The Case of Context-Aware Mobile Travel Applications', *Open and Big Data Management and Innovation*, Springer, pp. 354-65.
- de Reuver, M., Sørensen, C. & Basole, R.C. 2017, 'The digital platform: a research agenda', *Journal of Information Technology*, pp. 1-12.
- Debatin, B., Lovejoy, J.P., Horn, A.K. & Hughes, B.N. 2009, 'Facebook and online privacy: Attitudes, behaviors, and unintended consequences', *Journal of Computer-Mediated Communication*, vol. 15, no. 1, pp. 83-108.
- Doyle, K. 2015, 'Facebook, Whatsapp and the Commodification of Affective labour', *Communication, Politics & Culture*, vol. 48, no. 1, p. 51.
- Duncan, T. & Moriarty, S.E. 1998, 'A communication-based marketing model for managing relationships', *The Journal of Marketing*, pp. 1-13.
- Dunnebeil, S., Sunyaev, A., Leimeister, J.M. & Krcmar, H. 2012, 'Market engineering for electronic health services', *Hawaii International Conference on System Science (HICSS)*, pp. 952-61.
- Edvardsson, B., Tronvoll, B. & Gruber, T. 2011, 'Expanding understanding of service exchange and value co-creation: a social construction approach', *Journal of the Academy of Marketing Science*, vol. 39, no. 2, pp. 327-39.
- Eisenhardt, K.M. 1989, 'Building theories from case study research', *Academy of Management Review*, vol. 14, no. 4, pp. 532-50.
- Eissens-van der Laan, M., Broekhuis, M., van Offenbeek, M. & Ahaus, K. 2016, 'Service decomposition: a conceptual analysis of modularizing services', *International Journal of Operations & Production Management*, vol. 36, no. 3, pp. 308-31.

- Elliott, D., Hopfgartner, F., Leelanupab, T., Moshfeghi, Y. & Jose, J.M. 2009, 'An architecture for life-long user modelling', paper presented to the *User Modeling, Adaptation and Personalisation (UMAP)*, Trento, Italy.
- Etgar, M. 2008, 'A descriptive model of the consumer co-production process', *Journal of the Academy of Marketing Science*, vol. 36, no. 1, pp. 97-108.
- Ethiraj, S.K. & Levinthal, D. 2004, 'Modularity and innovation in complex systems', *Management Science*, vol. 50, no. 2, pp. 159-73.
- Evans, D.S. 2009, 'How catalysts ignite: the economics of platform-based start-ups', in A. Gawer (ed.), *Platforms, Markets and Innovation*.
- Evans, D.S. & Schmalensee, R. 2005, *The industrial organization of markets with two-sided platforms*, National Bureau of Economic Research.
- Evans, P.C. & Gawer, A. 2016, 'The rise of the platform enterprise: a global survey'.
- Facebook 2014, *Statement of Rights and Responsibilities*, viewed 12th November 2014, <<https://www.facebook.com/legal/terms>>.
- Facebook 2017, 'Annual Report'.
- Facebook Inc 2012, *Facebook Annual Report*, Facebook Inc.
- Farrelly, R. 2014, 'From Guardians to Brokers – a Taxonomy of Personal Infomediaries', *unpublished*.
- Farrelly, R. 2017, *Disrupting the Disruptors*, viewed 16th January 2018, <<https://www.youtube.com/watch?v=rnyCgbTBliQ&t=1987s>>.
- Farrelly, R. & Chew, E. 2016a, 'Designing a Personal Information Transaction Object', *System Sciences (HICSS), 2016 49th Hawaii International Conference on*, IEEE, Hawaii, pp. 4061-71.
- Farrelly, R. & Chew, E. 2016b, 'Towards a Primary Personal Information Market', *Twelfth International Conference on Technology, Knowledge, and Society*.
- Farrelly, R. & Chew, E. 2016c, 'Who's in to Win?: Participation Rate in a Primary Personal Information Market', *Proceedings of the 27th Australasian Conference on Information Systems*.
- Farrelly, R. & Chew, E. 2017, 'Designing a Primary Personal Information Market as an Industry Platform: a Service Innovation Approach', *System Sciences (HICSS), 2017 50th Hawaii International Conference on*, IEEE, Hawaii.
- Farrelly, R. & Chew, E. forthcoming, 'Transforming the worldwide market in personal information', *Journal of Enterprise Transformation*.
- Featherman, M.S. & Pavlou, P.A. 2003, 'Predicting e-services adoption: a perceived risk facets perspective', *International journal of human-computer studies*, vol. 59, no. 4, pp. 451-74.
- Fenning, R., Dogan, H. & Phalp, K. 2014, 'Applicability of SSM and UML for Designing a Search Application for the British Broadcasting Corporation (BBC)', *Enterprise, Business-Process and Information Systems Modeling*, Springer, pp. 472-86.
- Fichman, R.G. 2004, 'Real options and IT platform adoption: Implications for theory and practice', *Information Systems Research*, vol. 15, no. 2, pp. 132-54.
- Franklin, B. 2013, *Autobiography of Benjamin Franklin*, Tribeca Books.
- Frean, A. 2016, *Uber rating system driving customers to distraction*, viewed 2nd November 2016, <<http://www.theaustralian.com.au/business/uber-rating-system-driving-customers-to-distraction/news-story/191a8a907653823e0b0c2b23b56ed47b>>.
- French, M. & Hancock, J. 2017, 'What's the folk theory? Reasoning about cyber-social systems.'.
- Frow, P., Nenonen, S., Payne, A. & Storbacka, K. 2015, 'Managing Co-creation Design: A Strategic Approach to Innovation', *British Journal of Management*, vol. 26, no. 3, pp. 463-83.
- Gabisch, J.A. & Milne, G.R. 2014, 'The impact of compensation on information ownership and privacy control', *Journal of Consumer Marketing*, vol. 31, no. 1, pp. 13-26.
- Garcia, D. 2000, 'Governing electronic commerce in a global environment', *VDI-Berichte*, pp. 33-67.

- Gawer, A. 2014, 'Bridging differing perspectives on technological platforms: Toward an integrative framework', *Research Policy*, vol. 43, no. 7, pp. 1239-49.
- Gawer, A. & Cusumano, M.A. 2002, *Platform leadership: How Intel, Microsoft, and Cisco drive industry innovation*, Harvard Business School Press Boston.
- Gawer, A. & Cusumano, M.A. 2008, 'How companies become platform leaders', *MIT Sloan Management Review*, vol. 49, no. 2, p. 28.
- Gawer, A. & Cusumano, M.A. 2014, 'Industry platforms and ecosystem innovation', *Journal of Product Innovation Management*, vol. 31, no. 3, pp. 417-33.
- Gellman, R. 1996, 'Disintermediation and the Internet', *Government Information Quarterly*, vol. 13, no. 1, pp. 1-8.
- General Data Protection Regulation 2017, *GDPR Portal*, viewed 11th March 2017, <<http://www.eugdpr.org/>>.
- Ghazawneh, A. & Henfridsson, O. 2015, 'A paradigmatic analysis of digital application marketplaces', *Journal of Information Technology*, vol. 30, no. 3, pp. 198-208.
- Goes, P.B. 2014, 'Editor's comments: design science research in top information systems journals', *MIS Quarterly*, vol. 38, no. 1, pp. iii-viii.
- Goh, K.H. & Bockstedt, J.C. 2013, 'The framing effects of multipart pricing on consumer purchasing behavior of customized information good bundles', *Information Systems Research*, vol. 24, no. 2, pp. 334-51.
- Granados, N., Gupta, A. & Kauffman, R.J. 2005, 'Identifying facilitators and inhibitors of market structure change: a hybrid theory of unbiased electronic markets', *System Sciences, 2005. HICSS'05. Proceedings of the 38th Annual Hawaii International Conference on*, pp. 161c-c.
- Gregor, S. 2006, 'The Nature of Theory in Information Systems', *MIS Quarterly*, vol. 30, no. 3, pp. 611-42.
- Gregor, S. 2018, 'private correspondence'.
- Gregor, S. & Hevner, A.R. 2013, 'Positioning and presenting design science research for maximum impact', *MIS Quarterly*, vol. 37, no. 2, pp. 337-56.
- Grönroos, C. 1982, 'An applied service marketing theory', *European Journal of Marketing*, vol. 16, no. 7, pp. 30-41.
- Grönroos, C. 1984, 'A service quality model and its marketing implications', *European Journal of Marketing*, vol. 18, no. 4, pp. 36-44.
- Grönroos, C. 2011a, 'A service perspective on business relationships: The value creation, interaction and marketing interface', *Industrial Marketing Management*, vol. 40, no. 2, pp. 240-7.
- Grönroos, C. 2011b, 'Value co-creation in service logic: A critical analysis', *Marketing Theory*, vol. 11, no. 3, pp. 279-301.
- Grönroos, C. 2012, 'Conceptualising value co-creation: A journey to the 1970s and back to the future', *Journal of Marketing Management*, vol. 28, no. 13-14, pp. 1520-34.
- Grönroos, C. & Voima, P. 2013, 'Critical service logic: making sense of value creation and co-creation', *Journal of the Academy of Marketing Science*, vol. 41, no. 2, pp. 133-50.
- Grossklags, J. & Acquisti, A. 2007, 'When 25 Cents is Too Much: An Experiment on Willingness-To-Sell and Willingness-To-Protect Personal Information', *WEIS*.
- Grover, V. & Teng, J.T. 2001, 'E-commerce and the information market', *Communications of the ACM*, vol. 44, no. 4, pp. 79-86.
- Gualtieri, M. 2013, *Predictive Apps Are The Next Big Thing in Customer Engagement*, Forrester.
- Gummesson, E. 1991, 'Marketing-orientation Revisited: The Crucial Role of the Part-time Marketer', *European Journal of Marketing*, vol. 25, no. 2, pp. 60-75.
- Gummesson, E. 2002, 'Relationship marketing and a new economy: it's time for de-programming', *Journal of Services Marketing*, vol. 16, no. 7, pp. 585-9.
- Gummesson, E. 2007, 'Exit services marketing-enter service marketing', *Journal of Customer Behaviour*, vol. 6, no. 2, pp. 113-41.

- Hagel, J. & Singer, M. 1999, *Net worth: shaping markets when customers make the rules*, Harvard Business Press.
- Hagiu, A. 2009, 'Multi-sided platforms: From microfoundations to design and expansion strategies', *Harvard Business School Strategy Unit Working Paper*, no. 09-115.
- Hagiu, A. & Spulber, D. 2011, *First-party content, commitment and coordination in two-sided markets*, Harvard Business School.
- Hancock, B. 1999, 'Novell maybe causing privacy concerns with digitalme', *Computers & Security*, vol. 18, no. 3, pp. 190-1.
- Hann, I.-H., Hui, K.-L., Lee, S.-Y.T. & Png, I.P. 2007, 'Overcoming online information privacy concerns: An information-processing theory approach', *Journal of Management Information Systems*, vol. 24, no. 2, pp. 13-42.
- Harris, J. & Blair, E.A. 2006, 'Consumer preference for product bundles: The role of reduced search costs', *Journal of the Academy of Marketing Science*, vol. 34, no. 4, pp. 506-13.
- Harris, P. & Cilento, M. 2016, *Data Availability and Use*, <<http://www.pc.gov.au/inquiries/current/data-access/draft>>.
- Harrison, S. 2018, *Can you make money selling your data?*, viewed 2nd October 2018, <<http://www.bbc.com/capital/story/20180921-can-you-make-money-selling-your-data>>.
- HAT Research and Development 2017, viewed 14th July 2017, <<http://hubofallthings.com/>>.
- Hau, Y.S. & Kim, Y.-G. 2011, 'Why would online gamers share their innovation-conducive knowledge in the online game user community? Integrating individual motivations and social capital perspectives', *Computers in Human Behavior*, vol. 27, no. 2, pp. 956-70.
- Helkkula, A., Kelleher, C. & Pihlström, M. 2012, 'Characterizing value as an experience: implications for service researchers and managers', *Journal of Service Research*, vol. 15, no. 1, pp. 59-75.
- Hevner, A. 2007, 'A Three Cycle View of Design Science Research', *Scandinavian Journal of Information Systems*, vol. 19, no. 2, pp. 87-92.
- Hevner, A., March, S., Park, J. & Ram, S. 2004, 'Design Science in Information Systems Research', *MIS Quarterly*, vol. 28, no. 1, pp. 75-105.
- Hinkle, D.E., Oliver, J.D. & Hinkle, C.A. 1985, 'How large should the sample be? Part II—The one-sample case for survey research', *Educational and Psychological Measurement*, vol. 45, no. 2, pp. 271-80.
- Hoeren, T. 2014, 'Big data and the ownership in data: recent developments in Europe', *European Intellectual Property Review*, vol. 36, no. 12, pp. 751-4.
- Holbrook, M.B. 1996, 'Special Session Summary. Customer Value—A Framework for Analysis and Research', *Advances in Consumer Research*, vol. 23, no. 2, pp. 138-42.
- Holdren, J.P. & Lander, E.S. 2014, *Big Data and Privacy: A Technological Perspective*, 2016, <https://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf>.
- Holtinen, H. 2010, 'Social practices as units of value creation: theoretical underpinnings and implications', *International Journal of Quality and Service Sciences*, vol. 2, no. 1, pp. 95-112.
- Hoyer, W.D., Chandy, R., Dorotic, M., Krafft, M. & Singh, S.S. 2010, 'Consumer cocreation in new product development', *Journal of Service Research*, vol. 13, no. 3, pp. 283-96.
- Hsu, C.-L. & Lin, J.C.-C. 2008, 'Acceptance of blog usage: The roles of technology acceptance, social influence and knowledge sharing motivation', *Information & Management*, vol. 45, no. 1, pp. 65-74.
- Hu, T., Poston, R.S. & Kettinger, W.J. 2011, 'Nonadopters of online social network services: Is it easy to have fun yet', *Communications of the Association for Information Systems*, vol. 29, no. 1, pp. 441-58.
- Huber, F., Herrmann, A. & Morgan, R.E. 2001, 'Gaining competitive advantage through customer value oriented management', *Journal of Consumer Marketing*, vol. 18, no. 1, pp. 41-53.

- Iansiti, M. & Levien, R. 2004, *The keystone advantage: what the new dynamics of business ecosystems mean for strategy, innovation, and sustainability*, Harvard Business Press.
- Infact Decisions 2017, *Preparing For The Era Of Consumer Controlled Data*, viewed 11th March 2017, <<https://www.picatic.com/ConsumerControlledDataSydney>>.
- International Telecommunication Union 2017, *2017 estimates for key ICT indicators*, International Telecommunication Union, viewed 3rd June 2018, <https://www.itu.int/en/ITU-D/Statistics/Documents/statistics/2017/ITU_Key_2005-2017_ICT_data.xls>.
- Islam, H. & Miah, S.J. 2012, 'Motivation-based IS evaluation strategy: a perspective of marketing information systems', *ACIS 2012: Location, location, location: Proceedings of the 23rd Australasian Conference on Information Systems 2012*, ACIS, pp. 1-10.
- Jaakkola, E. & Alexander, M. 2014, 'The Role of Customer Engagement Behavior in Value Co-Creation A Service System Perspective', *Journal of Service Research*, vol. 17, no. 3, pp. 247-61.
- Jamrisko, M. & Miller, M. 2018, *If privacy is dead, why can't we sell our own data?*, viewed 30th September 2018, <<https://www.smh.com.au/business/consumer-affairs/if-privacy-is-dead-why-can-t-we-sell-our-own-data-20180907-p502ac.html>>.
- Jennings, H.G.N.R., Ockenfels, G.E.K.A. & Weinhardt, C. 2008, *Negotiation, Auctions, and Market Engineering*, Springer.
- Johnson, K. 2012, 'In Search of a Universal Model for Personal Data', Senior Honors Thesis, The University of Utah.
- Johnson, V. 1939, 'Aristotle's theory of value', *American Journal of Philology*, pp. 445-51.
- Jones, D. 2008, 'The Anatomy of a Design Science Paper: A Research Note', <<https://davidtjones.wordpress.com/2008/10/09/the-anatomy-of-a-design-science-paper-a-research-note/>>.
- Kaplan, A. & Haenlein, M. 2010, 'Users of the world, unite! The challenges and opportunities of Social Media', *Business Horizons*, vol. 53, p. 59—68.
- Karunakaran, A. & Purao, S. 2012, 'Designing for recombination: process design through template combination', *Design Science Research in Information Systems. Advances in Theory and Practice*, Springer, pp. 36-51.
- Katell, M.A., Mishra, S.R. & Scaff, L. 2016, 'A Fair Exchange: Exploring How Online Privacy is Valued', *2016 49th Hawaii International Conference on System Sciences (HICSS)*, IEEE, pp. 1881-90.
- Khalifa, A.S. 2004, 'Customer value: a review of recent literature and an integrative configuration', *Management Decision*, vol. 42, no. 5, pp. 645-66.
- Kickstarter 2018, *Kickstarter*, <<https://www.kickstarter.com>>.
- Kietzmann, J., Hermkens, K., McCarthy, I. & Silvestre, B. 2011, 'Social media? Get serious! Understanding the Functional Building Blocks of Social Media', *Business Horizons*, vol. 54, p. 241—51.
- Kim-Kwang Raymond, C. 2006, 'Issue report on business adoption of Microsoft Passport', *Information Management & Computer Security*, vol. 14, no. 3, pp. 218-34.
- King, N. 1994, 'Qualitative methods in organizational research: A practical guide', *The Qualitative Research Interview*.
- Kitchenham, B., Brereton, O.P., Budgen, D., Turner, M., Bailey, J. & Linkman, S. 2009, 'Systematic literature reviews in software engineering—a systematic literature review', *Information and Software Technology*, vol. 51, no. 1, pp. 7-15.
- Kleek, M.V., Moore, B., Karger, D., André, P. & Schraefel, M.C. 2010, 'Atomate It! End-user Context-Sensitive Automation using Heterogeneous Information Sources on the Web', paper presented to the *WWW 2010*, Raleigh, North Carolina, USA, April 26–30.
- Knol, A., Sol, H. & Van Wamelen, J. 2012, 'Decision enhancement for sourcing with shared service centres in the Dutch government', *Design Science Research in Information Systems. Advances in Theory and Practice*, Springer, pp. 239-55.

- Kontaxis, G., Polakis, I., Ioannidis, S. & Markatos, E.P. 2011, 'Detecting social network profile cloning', *Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 2011 IEEE International Conference on, IEEE, pp. 295-300.
- Korkman, O. 2006, *Customer value formation in practice: a practice-theoretical approach*, Svenska Handelshögskolan.
- Kraut, R.E., Resnick, P., Kiesler, S., Burke, M., Chen, Y., Kittur, N., Konstan, J., Ren, Y. & Riedl, J. 2012, *Building successful online communities: Evidence-based social design*, MIT Press.
- Krueger, M.W., Chew, E.K., Ouetani, Z.M. & Gitzel, R. 2015, 'Integrative Service Innovation: An Industrial Use Case', *Business Informatics (CBI)*, 2015 IEEE 17th Conference on, vol. 1, IEEE, pp. 217-23.
- Kumar, M.A., Arijit Bhattacharya, Jose Arturo Garza-Reyes, Vikas, Parry, G.C., Brax, S.A., Maull, R.S. & Ng, I.C. 2016, 'Operationalising IoT for reverse supply: the development of use-visibility measures', *Supply Chain Management: An International Journal*, vol. 21, no. 2, pp. 228-44.
- Kuneva, M. 2009, 'Roundtable on online data collection, targeting and profiling', vol. 156, <http://europa.eu/rapid/press-release_SPEECH-09-156_en.htm>.
- Kyriacou, D. & Davis, H. 2008, 'Moving Towards Life-Long User Modeling', *Advanced Learning Technologies, 2008. ICAALT '08. Eighth IEEE International Conference on*, pp. 647-8.
- Lam, M. 2019, *Personal Sharing Reinvented*, <<https://mediax.stanford.edu/featured-events/monica-lam-mediax2015/>>.
- Laney, D., Buytendijk, F. & Linden, A. 2013, *Predicts 2014: Innovating With Information Will Demand New Data, Organizations and Ideas*, Gartner.
- Laudon, K. 1996, 'Markets and Privacy', *Association for Computing Machinery. Communications of the ACM*, vol. 39, no. 9, pp. 92-104.
- Lee, J., Kao, H.-A. & Yang, S. 2014, 'Service innovation and smart analytics for industry 4.0 and big data environment', *Procedia CIRP*, vol. 16, pp. 3-8.
- Lee, J., Wyner, G.M. & Pentland, B.T. 2008, 'Process grammar as a tool for business process design', *MIS Quarterly*, pp. 757-78.
- Lehnerd, A.P. 1987, 'Revitalizing the manufacture and design of mature global products', National Academy Press, Washington, DC, pp. 49-64.
- Lehtinen, U. & Lehtinen, J.R. 1991, 'Two approaches to service quality dimensions', *Service Industries Journal*, vol. 11, no. 3, pp. 287-303.
- Leickly, B.L. 2004, 'Infomediaries in Information Economies', Georgetown University, Washington, DC.
- Leiner, B.M., Cerf, V.G., Clark, D.D., Kahn, R.E., Kleinrock, L., Lynch, D.C., Postel, J., Roberts, L.G. & Wolff, S. 2009, 'A brief history of the Internet', *ACM SIGCOMM Computer Communication Review*, vol. 39, no. 5, pp. 22-31.
- Lemos, R. 2001, *Security problems open Microsoft's Wallet*, viewed 25th July 2014, <http://news.cnet.com/Security-problems-open-Microsofts-Wallet/2100-1002_3-275366.html>.
- Lengnick-Hall, C.A., Claycomb, V.C. & Inks, L.W. 2000, 'From recipient to contributor: examining customer roles and experienced outcomes', *European Journal of Marketing*, vol. 34, no. 3/4, pp. 359-83.
- Lessard, L. 2014, 'Designing and Managing Value Co-Creation in KIBS Engagements', *Technology Innovation Management Review*, vol. 4, no. 7.
- Levenstein, J. 2013, *Email Statistics Report, 2013 -2017*, The Radicati Group Inc.
- Levitt, T. 1960, 'Marketing myopia', *Harvard business review*, vol. 38, no. 4, pp. 24-47.
- Li, C., Li, D.Y., Miklau, G. & Suci, D. 2013, 'A theory of pricing private data', paper presented to the *Proceedings of the 16th International Conference on Database Theory*, Genoa, Italy.

- Li, I., Dey, A. & Forlizzi, J. 2010, 'A stage-based model of personal informatics systems', *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 557-66.
- Li, I., Medynskiy, Y., Froehlich, J. & Larsen, J. 2012, 'Personal informatics in practice: improving quality of life through data', *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, pp. 2799-802.
- Li, N., Li, T. & Venkatasubramanian, S. 2010, 'Closeness: A new privacy measure for data publishing', *Knowledge and Data Engineering, IEEE Transactions on*, vol. 22, no. 7, pp. 943-56.
- Linkedin 2014, *User Agreement*, viewed 12th November 2014, <<https://www.linkedin.com/legal/user-agreement>>.
- Liu, X., Lee, C.-H. & Iyer, B. 2006, 'The Impact of Design Moves on Platform Adoption: The Case of Microsoft Windows OS', *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*, vol. 6, IEEE, pp. 141b-b.
- Lowenstein, R. 2004, *Origins of the crash: The great bubble and its undoing*, Penguin.
- Luening, E. 1998, *Microsoft to buy Firefly*, viewed 27th July 2014, <<http://news.cnet.com/2100-1001-210017.html>>.
- Lusch, R.F. & Nambisan, S. 2015, 'Service Innovation: A Service-Dominant Logic Perspective', *MIS Quarterly*, vol. 39, no. 1, pp. 155-75.
- Lusch, R.F., Vargo, S.L. & Tanniru, M. 2010, 'Service, value networks and learning', *Journal of the Academy of Marketing Science*, vol. 38, no. 1, pp. 19-31.
- Lusch, R.F. & Webster, F.E. 2011, 'A stakeholder-unifying, cocreation philosophy for marketing', *Journal of Macromarketing*, vol. 31, no. 2, pp. 129-34.
- Machanavajjhala, A., Kifer, D., Gehrke, J. & Venkatasubramanian, M. 2007, 'l-diversity: Privacy beyond k-anonymity', *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 1, no. 1, p. 3.
- Malthus, T.R. 1803, *An essay on the principle of population; a view of its past and present effects on human happiness; with an inquiry into our prospects respecting the future removal or mitigation of the evils which it occasions*.
- March, S. & Smith, G. 1995, 'Design and natural science research on information technology', *Decision Support Systems*, vol. 15, pp. 251-66.
- Marsden, P. 2011, 'eBranding and Social Commerce', *Brand Evolution*, Springer, pp. 357-72.
- Mashhadi, A., Kawsar, F. & Acer, U.G. 2014, 'Human Data Interaction in IoT: The ownership aspect', *Internet of Things (WF-IoT), 2014 IEEE World Forum on*, pp. 159-62.
- Matzler, K., Faullant, R., Renzl, B. & Leiter, V. 2005, 'The relationship between personality traits (extraversion and neuroticism), emotions and customer self-satisfaction', *Innovative Marketing*, vol. 1, no. 2, pp. 32-9.
- McCull-Kennedy, J.R., Vargo, S.L., Dagger, T.S., Sweeney, J.C. & van Kasteren, Y. 2012, 'Health care customer value cocreation practice styles', *Journal of Service Research*, pp. 370-89.
- McFedries, P. 2013, 'Tracking the quantified self', *IEEE Spectrum*, vol. 50, no. 8, p. 24.
- McGrath, R.G. 2010, 'Business models: a discovery driven approach', *Long range planning*, vol. 43, no. 2, pp. 247-61.
- McLaren, T.S., Head, M.M., Yuan, Y. & Chan, Y.E. 2011, 'A multilevel model for measuring fit between a firm's competitive strategies and information systems capabilities', *MIS Quarterly*, vol. 35, no. 4, pp. 909-29.
- Mele, C., Spena, T.R. & Colurcio, M. 2010, 'Co-creating value innovation through resource integration', *International Journal of Quality and Service Sciences*, vol. 2, no. 1, pp. 60-78.
- Merkens, H. 2004, 'Selection Procedures, Sampling, Case Construction', *A Companion to Qualitative Research*, pp. 165-71.
- Milberg, S.J., Smith, H.J. & Burke, S.J. 2000, 'Information privacy: Corporate management and national regulation', *Organization Science*, vol. 11, no. 1, pp. 35-57.

- Milgrom, P. 2011, 'Critical Issues in the Practice of Market Design', *Economic Inquiry*, vol. 49, no. 2, p. 313.
- Miller, T. & Holmes, K.R. 2010, 'Access to Information: Vital for Efficient Markets and Economic Reform', *2010 Index of Economic Freedom*, The Heritage Foundation and Dow Jones & Company, Inc., Washington, D.C.
- Milyaeva, S. & Neyland, D. Forthcoming, 'On re-devising markets, re-locating value: Online personal data and 'empowering' privacy'.
- MIT Human Dynamics Group 2017, *OPENPDS/SA*, viewed 7th March 2017, <<http://openpds.media.mit.edu/>>.
- Moeller, S., Ciuchita, R., Mahr, D., Odekerken-Schröder, G. & Fassnacht, M. 2013, 'Uncovering collaborative value creation patterns and establishing corresponding customer roles', *Journal of Service Research*, vol. 16, no. 4, pp. 471-87.
- Muffatto, M. & Roveda, M. 2002, 'Product architecture and platforms: a conceptual framework', *International Journal of Technology Management*, vol. 24, no. 1, pp. 1-16.
- Muschalle, A., Stahl, F., Löser, A. & Vossen, G. 2013, 'Pricing approaches for data markets', *Enabling Real-Time Business Intelligence*, Springer, pp. 129-44.
- Muzellec, L., Ronteau, S. & Lambkin, M. 2015, 'Two-sided Internet platforms: A business model lifecycle perspective', *Industrial Marketing Management*, vol. 45, pp. 139-50.
- Mydex 2017a, *About Mydex CIC*, <<https://mydex.org/about/>>.
- Mydex 2017b, *Mydex*, viewed 25th Feb 2017, <<https://mydex.org/>>.
- Nath, S.V. 2007, 'Champion-challenger based predictive model selection', *SoutheastCon, 2007. Proceedings*, IEEE, p. 25e.
- Negotiation, Auctions, and Market Engineering* 2006, Springer, <<http://link.springer.com.ezproxy.lib.uts.edu.au/book/10.1007%2F978-3-540-77554-6>>.
- Nononen, S. & Storbacka, K. 2010, 'Business model design: conceptualizing networked value co-creation', *International Journal of Quality and Service Sciences*, vol. 2, no. 1, pp. 43-59.
- Ng, I., Scharf, K., Pogrebna, G. & Maull, R. 2015, 'Contextual variety, Internet-of-Things and the choice of tailoring over platform: Mass customisation strategy in supply chain management', *International Journal of Production Economics*, vol. 159, pp. 76-87.
- Ng, I.C. & Wakenshaw, S.Y. 2016, 'The Internet-of-Things: Review and Research Directions', *International Journal of Research in Marketing*.
- Ngai, E.W., Moon, K.-I.K., Lam, S., Chin, E.S. & Tao, S.S. 2015, 'Social media models, technologies, and applications: an academic review and case study', *Industrial Management & Data Systems*, vol. 115, no. 5, pp. 769-802.
- Ngai, E.W., Tao, S.S. & Moon, K.K. 2015, 'Social media research: Theories, constructs, and conceptual frameworks', *International Journal of Information Management*, vol. 35, no. 1, pp. 33-44.
- Nissim, K., Vadhan, S. & Xiao, D. 2014, 'Redrawing the boundaries on purchasing data from privacy-sensitive individuals', *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science*, pp. 411-22.
- Norcie, G., De Cristofaro, E. & Bellotti, V. 2013, 'Bootstrapping trust in online dating: Social verification of online dating profiles', *International Conference on Financial Cryptography and Data Security*, Springer, pp. 149-63.
- Nordin, F. & Kowalkowski, C. 2010, 'Solutions offerings: a critical review and reconceptualisation', *Journal of Service Management*, vol. 21, no. 4, pp. 441-59.
- Oliveira, P. & von Hippel, E. 2011, 'Users as service innovators: The case of banking services', *Research Policy*, vol. 40, no. 6, pp. 806-18.
- Oxborrow, L. & Brindley, C. 2014, 'Disintermediation in the apparel supply chain', *Journal of Fashion Marketing and Management*, vol. 18, no. 3, pp. 252-68.
- Palvia, P. & D'Aubeterre, F. 2007, 'An Examination of Infomediary Roles in B2C E-Commerce', *Journal of Electronic Commerce Research*, vol. 8, no. 4, pp. 207-20.

- Parker, G. & Van Alstyne, M.W. 2014, *Platform Strategy Survey*, Boston University School of Management, Research Paper 2439323, <<http://ssrn.com/abstract=2439323>>.
- Parker, G., Van Alstyne, M.W. & Jiang, X. 2016, *Platform ecosystems: How developers invert the firm*, Boston University Questrom School of Business Research Paper No. 2861574, <<https://ssrn.com/abstract=2861574>>.
- Parker, G.G. & Van Alstyne, M.W. 2005, 'Two-sided network effects: A theory of information product design', *Management Science*, vol. 51, no. 10, pp. 1494-504.
- Peffer, K., Tuunanen, T., Gengler, C.E., Rossi, M., Hui, W., Virtanen, V. & Bragge, J. 2006, 'The design science research process: a model for producing and presenting information systems research', *Proceedings of the First International Conference on Design Science Research in Information Systems and Technology (DESRIST)*, pp. 83-106.
- Peffer, K., Tuunanen, T., Rothenberger, M. & Chatterjee, S. 2007-8, 'A Design Science Research Methodology for Information Systems Research', *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45-77.
- Perera, C., Wakenshaw, S.Y., Baarslag, T., Haddadi, H., Bandara, A.K., Mortier, R., Crabtree, A., Ng, I.C., McAuley, D. & Crowcroft, J. 2016, 'Valorising the IoT Databox: creating value for everyone', *Transactions on Emerging Telecommunications Technologies. Personal Black Box* 2016, viewed 2nd November 2016, <<http://pbb.me/>>.
- Personal Informatics 2014, *Know Thyself*, Personal Informatics, viewed 2nd April 2104, <<http://personalinformatics.org/>>.
- Piezunka, H. 2011, 'Technological platforms', *Journal für Betriebswirtschaft*, vol. 61, no. 2-3, pp. 179-226.
- Polo Peña, A.I., Frías Jamilena, D.M. & Rodríguez Molina, M.Á. 2014, 'Value co-creation via information and communications technology', *The Service Industries Journal*, vol. 34, no. 13, pp. 1043-59.
- POMI 2019, <<http://pomi.stanford.edu/index.php>>.
- Porter, M.E. & Millar, V.E. 1985, 'How information gives you competitive advantage', *Harvard Business Review*, vol. 63, no. 4, pp. 149-60.
- Prahalad, C.K. & Ramaswamy, V. 2004, 'Co-creating unique value with customers', *Strategy & Leadership*, vol. 32, no. 3, pp. 4-9.
- Prins, J. 2004, 'The Propertization of Personal Data and Identities', *Electronic Journal of Comparative Law*, vol. 8, no. 3.
- Purao, S. 2002, 'Design research in the technology of information systems: Truth or dare', *GSU Department of CIS Working Paper*.
- Quantified Self 2013, *LifeLogging*, Quantified Self, viewed 2nd April 2014, <<http://quantifiedself.com/lifelogging/>>.
- R Core Team 2015, 'R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.'
- Raasch, C. & Von Hippel, E.A. 2013, 'Innovation Effort as 'Productive Consumption': The Power of Participation Benefits to Amplify Innovation', *Available at SSRN: <https://ssrn.com/abstract=2167948> or <http://dx.doi.org/10.2139/ssrn.2167948>*.
- Radicati, S. 2013, *Mobile Statistics Report, 2013-2017*, The Radicati Group Inc.
- Ramaswamy, V. 2011, 'It's about human experiences... and beyond, to co-creation', *Industrial Marketing Management*, vol. 40, no. 2, pp. 195-6.
- Ramirez, R. 1999, 'Value co-production: intellectual origins and implications for practice and research', *Strategic Management Journal*, vol. 20, no. 1, pp. 49-65.
- Ranjan, K.R. & Read, S. 2014, 'Value co-creation: concept and measurement', *Journal of the Academy of Marketing Science*, pp. 1-26.
- Rawassizadeh, R., Tomitsch, M., Wac, K. & Tjoa, A.M. 2013, 'UbiqLog: a generic mobile phone-based life-log framework', *Personal and Ubiquitous Computing*, vol. 17, no. 4, pp. 621-37.
- Redrup, Y. 2017, *Mandatory data breach notification laws pass lower house*, *Financial Review*, viewed 7th March 2017,

- <<http://www.afr.com/technology/web/security/mandatory-data-breach-notification-laws-pass-lower-house-20170208-gu86w4>>.
- Rees, C. 2014, 'Who owns our data?', *Computer Law & Security Review*, vol. 30, no. 1, pp. 75-9.
- Reiss, S. 2004, 'Multifaceted Nature of Intrinsic Motivation: The Theory of 16 Basic Desires', *Review of General Psychology*, vol. 8, no. 3, p. 179.
- Revolutions 2018, *Revolutions*, <<http://blog.revolutionanalytics.com/2017/01/cran-10000.html>>.
- Richins, M.L. 1994, 'Valuing things: The public and private meanings of possessions', *Journal of Consumer Research*, vol. 21, pp. 504-21.
- Riederer, C., Erramilli, V., Chaintreau, A., Krishnamurthy, B. & Rodriguez, P. 2011, 'For sale: your data: by: you', *Proceedings of the 10th ACM WORKSHOP on Hot Topics in Networks*, ACM, p. 13.
- Roch, J. & Mosconi, E. 2016, 'The Use of Social Media Tools in the Product Life Cycle Phases: A Systematic Literature Review', *2016 49th Hawaii International Conference on System Sciences (HICSS)*, IEEE, pp. 1830-9.
- Rochet, J.-C. & Tirole, J. 2004, 'Two-sided markets: an overview', *Institut d'Economie Industrielle working paper*.
- Rose, F. 1999, *The Economics, Concept, and Design of Information Intermediaries*, Springer.
- Rose, J., Rehse, O. & Röber, B. 2012, *The value of our digital identity*, <<http://www.libertyglobal.com/PDF/public-policy/The-Value-of-Our-Digital-Identity.pdf>>.
- Roth, A.E. 2008, 'What Have We Learned from Market Design?', *NBER Innovation Policy & the Economy (University of Chicago Press)*, vol. 9, pp. 79-112.
- Roth, A.V. & Menor, L.J. 2003, 'Insights into service operations management: a research agenda', *Production and Operations Management*, vol. 12, no. 2, pp. 145-64.
- Russell, M. 2011, 'Adaptive mediated persuasion technologies', *Proceedings of the 6th International Conference on Persuasive Technology: Persuasive Technology and Design: Enhancing Sustainability and Health*.
- Salazar, A.J. 2015, 'Platform Competition: A Research Framework and Synthesis of Game-Theoretic Studies', *Available at SSRN 2565337*.
- Sánchez-Fernández, R., Angeles Iniesta-Bonillo, M. & Holbrook, M.B. 2009, 'The conceptualisation and measurement of consumer value in services', *International Journal of Market Research*, vol. 51, no. 1, pp. 93-113.
- Sánchez-Fernández, R. & Iniesta-Bonillo, M.Á. 2007, 'The concept of perceived value: a systematic review of the research', *Marketing Theory*, vol. 7, no. 4, pp. 427-51.
- Sarkar, M.B., Butler, B. & Steinfield, C. 1995, 'Intermediaries and cybermediaries: a continuing role for mediating players in the electronic marketplace', *Journal of Computer-Mediated Communication*, vol. 1, no. 3, pp. 1-14.
- Saunders, C. 2000, *mValue Dead in the Water*, viewed 28th June 2014, <<http://www.clickz.com/clickz/news/1700532/mvalue-dead-water>>.
- Saunders, C. 2001, *All Over at AllAdvantage*, viewed 9th June 2014, <<http://www.clickz.com/clickz/news/1691463/all-over-alladvantage>>.
- Schmidt, C. 2004, 'The analysis of semi-structured interviews', *A companion to qualitative research*, pp. 253-8.
- Schneier, B. 2015, *AT&T Charging Customers to Not Spy on Them*, viewed 29th Jan 2016, <https://www.schneier.com/blog/archives/2015/02/att_charging_cu.html>.
- Schunter, M. 2014, 'Data Security and Privacy in 2025?', *Secure Data Management*, pp. 37-41.
- Schwartz, P. 2003, 'Property, Privacy, and Personal Data', *Harvard Law Review*, vol. 115, p. 2056.
- Searby, S. 2003, 'Personalisation—an overview of its use and potential', *BT Technology Journal*, vol. 21, no. 1, pp. 13-9.
- Sellen, A.J. & Whittaker, S. 2010, 'Beyond total capture: a constructive critique of lifelogging', *Commun. ACM*, vol. 53, no. 5, pp. 70-7.

- Sevignani, S. 2013, 'The commodification of privacy on the Internet', *Science and Public Policy*, vol. 40, no. 6, pp. 733-9.
- Shadbolt, N. & Hampson, R. 2018, *Who should hold the keys to our data?*, viewed 30th September 2018, <<https://www.theguardian.com/commentisfree/2018/apr/29/in-charge-our-own-data-personal-information-facebook-scandal>>.
- Shah, M. 2015, 'Big Data and the Internet of Things', *Information Fusion*, vol. 40.
- Shibchurn, J. & Yan, X.B. 2014, 'Investigating Effects of Monetary Reward on Information Disclosure by Online Social Networks Users', *System Sciences (HICSS), 2014 47th Hawaii International Conference on*, IEEE, pp. 1725-34.
- Shilton, K., Burke, J.A., Estrin, D., Govindan, R., Hansen, M., Kang, J. & Mun, M. 2009, 'Designing the personal data stream: Enabling participatory privacy in mobile personal sensing', *Center for Embedded Network Sensing*.
- Shimojo, A., Kamada, S., Matsumoto, S. & Nakamura, M. 2010, 'On integrating heterogeneous lifelog services', *Proceedings of the 12th International Conference on Information Integration and Web-based Applications & Services*, pp. 263-72.
- Shuradze, G., Wagner, D. & Wagner, H.-T. 2015, 'Technological Platforms and Innovation: Review, Integration, and Extension', paper presented to the *Twenty-first Americas Conference on Information Systems*, Puerto Rico.
- Sitaloppi, J. & Vargo, S.L. 2014, 'Reconciling Resource Integration and Value Propositions-- The Dynamics of Value Co-creation', *System Sciences (HICSS), 2014 47th Hawaii International Conference on*, IEEE, pp. 1278-84.
- Simon, H. 1996, *The Sciences of the Artificial*, 3rd edn, The MIT Press.
- Slemko, M. 2001, *Microsoft Passport to Trouble*, viewed 25th July 2014, <<http://www.znep.com/~marcs/passport/>>.
- Smith, A. 1776, *An inquiry into the nature and causes of the Wealth of Nations*, T. Nelson and Sons.
- Smith, O. 2017, *People.io lets you cash-in by sharing your personal data*, viewed 30th September 2018, <<https://www.thememo.com/2017/07/06/people-io-lets-you-cash-in-by-sharing-your-personal-data/>>.
- Solli-Sæther, H. & Flak, L.S. 2014, 'Proposing an Entrepreneurial Process for the Co-creation of IT Value', *Electronic Government*, Springer, pp. 313-24.
- Solomon, M.R. & Buchanan, B. 1991, 'A role-theoretic approach to product symbolism: mapping a consumption constellation', *Journal of Business Research*, vol. 22, no. 2, pp. 95-109.
- Spagnoletti, P., Resca, A. & Lee, G. 2015, 'A design theory for digital platforms supporting online communities: a multiple case study', *Journal of Information Technology*, vol. 30, no. 4, pp. 364-80.
- Spiekermann, S. & Novotny, A. 2014, *The Vision: Personal Information Markets AND Privacy*, <<https://pdfs.semanticscholar.org/e1b7/3005f57c400c78b6c55a22f11be11c7cb330.pdf>>.
- Spiekermann, S. & Novotny, A. 2015, 'A vision for global privacy bridges: technical and legal measures for international data markets', *Computer Law & Security Review*, vol. 31, no. 2, pp. 181-200.
- Stahl, F., Schomm, F. & Vossen, G. 2014, *The data marketplace survey revisited*, Working Papers, ERCIS-European Research Center for Information Systems.
- Stephenson, H. 2017, *22,000 people willingly agree to community service in return for free WiFi*, viewed 5th August 2017, <<https://purple.ai/purple-community-service/>>.
- Sterman, J.D., Henderson, R., Beinhocker, E.D. & Newman, L.I. 2007, 'Getting big too fast: Strategic dynamics with increasing returns and bounded rationality', *Management Science*, vol. 53, no. 4, pp. 683-96.
- Stock, R.M., Oliveira, P. & Von Hippel, E.A. 2013, 'Impacts of Hedonic and Utilitarian User Motives on the Innovativeness of User-Developed Solutions', *Available at SSRN 2260436*.

- Sugumaran, M., Murugan, B.B. & Kamalraj, D. 2014, 'An Architecture for Data Security in Cloud Computing', *Computing and Communication Technologies (WCCCT), 2014 World Congress on*, pp. 252-5.
- Sundsted, T.E. 2002, *With Liberty and single sign-on for all, The Liberty Alliance Project seeks to solve the current online identity crisis* Java World, viewed 27th July 2014, <<http://www.javaworld.com/article/2074036/java-security/with-liberty-and-single-sign-on-for-all.html>>.
- Swan, M. 2013, 'The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery', *Big Data*, vol. 1, no. 2, pp. 85 - 99.
- Sweeney, L. 2002, 'k-anonymity: A model for protecting privacy', *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 05, pp. 557-70.
- Taitague, A. 2013, *Instant Messaging Market, 2013-2017*, The Radicati Group Inc.
- Tapscott, D. 2016, 'How the blockchain is changing money and business', <http://www.ted.com/talks/don_tapscott_how_the_blockchain_is_changing_money_and_business#t-179062>.
- Taylor, C.R. 2002, 'Private demands and demands for privacy: Dynamic pricing and the market for customer information', *Durham, NC, Duke University*, pp. 1-30.
- Tee, R. & Gawer, A. 2009, 'Industry architecture as a determinant of successful platform strategies: A case study of the i-mode mobile Internet service', *European Management Review*, vol. 6, no. 4, pp. 217-32.
- Teevan, J., Dumais, S. & Horvitz, E. 2005, 'Personalizing search via automated analysis of interests and activities', *ACM SIGIR Forum*, vol. 51, no. 3, pp. 449-56.
- Thapliyal, M., Mandoria, H.L. & Garg, N. 2014, 'Data Security Analysis in Cloud Environment: A Review', *International Journal of Innovations & Advancement in Computer Science*, vol. 2, no. 1, pp. 14-9.
- The Solid Project 2016, <<https://solid.mit.edu/>>.
- Thierer, A.D., Koopman, C., Hobson, A. & Kuiper, C. 2015, 'How the Internet, the Sharing Economy, and Reputational Feedback Mechanisms Solve the 'Lemons Problem'', Available at SSRN 2610255.
- Thomas, L.D., Autio, E. & Gann, D.M. 2014, 'Architectural leverage: putting platforms in context', *The Academy of Management Perspectives*, vol. 28, no. 2, pp. 198-219.
- Tiwana, A. 2013, *Platform ecosystems: aligning architecture, governance, and strategy*, Newnes.
- Tiwana, A. & Konsynski, B. 2010, 'Complementarities between organizational IT architecture and governance structure', *Information Systems Research*, vol. 21, no. 2, pp. 288-304.
- Tiwana, A., Konsynski, B. & Bush, A.A. 2010, 'Research commentary-Platform evolution: Coevolution of platform architecture, governance, and environmental dynamics', *Information Systems Research*, vol. 21, no. 4, pp. 675-87.
- Toppenberg, G., Henningsson, S. & Eaton, B. 2016, 'Reinventing the Platform Core Through Acquisition: A Case Study', *2016 49th Hawaii International Conference on System Sciences (HICSS)*, IEEE, pp. 4634-43.
- Tse, A.C.-b. 2003, 'Disintermediation of travel agents in the hotel industry', *International Journal of Hospitality Management*, vol. 22, no. 4, pp. 453-60.
- Tulach, J. 2008, *Practical API design: Confessions of a java framework architect*, Apress.
- Tuunanen, T., Bask, A. & Merisalo-Rantanen, H. 2012, 'Typology for modular service design: review of literature', *International Journal of Service Science, Management, Engineering, and Technology (IJSSMET)*, vol. 3, no. 3, pp. 99-112.
- Tuunanen, T., Baskerville, R. & Beck, R. 2017, 'Introduction to Advances in Design Research for Information Systems Minitrack', *Proceedings of the 50th Hawaii International Conference on System Sciences*.
- Twitter 2014, *Terms of Service*, viewed 12th November 2014, <<https://twitter.com/tos>>.
- Ulrich, K. 1995, 'The role of product architecture in the manufacturing firm', *Research Policy*, vol. 24, no. 3, pp. 419-40.

- Van Alstyne, M. & Parker, G. 2017, 'Platform Business: From Resources to Relationships', *GfK Marketing Intelligence Review*, vol. 9, no. 1, pp. 24-9.
- Van Alstyne, M.W., Parker, G.G. & Choudary, S.P. 2016, 'Pipelines, platforms, and the new rules of strategy', *Harvard Business Review*, vol. 94, no. 4, pp. 54-62.
- Van der Boor, P., Oliveira, P. & Veloso, F. 2014, 'Users as innovators in developing countries: The global sources of innovation and diffusion in mobile banking services', *Research Policy*, vol. 43, no. 9, pp. 1594-607.
- Van Kleek, M. & O'Hara, K. 2014, 'The Future of Social is Personal: The Potential of the Personal Data Store', *Social Collective Intelligence*, Springer, pp. 125-58.
- Vargo, S. & Lusch, R. 2011, 'Service-dominant logic: Looking ahead', *Presentation at the Naples Forum on Service*, June, pp. 14-7.
- Vargo, S.L. 2009, 'Toward a transcending conceptualization of relationship: a service-dominant logic perspective', *Journal of Business & Industrial Marketing*, vol. 24, no. 5/6, pp. 373-9.
- Vargo, S.L. & Akaka, M.A. 2012, 'Value cocreation and service systems (re) formation: A service ecosystems view', *Service Science*, vol. 4, no. 3, pp. 207-17.
- Vargo, S.L. & Lusch, R.F. 2004, 'Evolving to a new dominant logic for marketing', *Journal of marketing*, vol. 68, no. 1, pp. 1-17.
- Vargo, S.L. & Lusch, R.F. 2008, 'Service-dominant logic: continuing the evolution', *Journal of the Academy of Marketing Science*, vol. 36, no. 1, pp. 1-10.
- Vargo, S.L. & Lusch, R.F. 2016, 'Institutions and axioms: an extension and update of service-dominant logic', *Journal of the Academy of Marketing Science*, vol. 44, no. 1, pp. 5-23.
- Vargo, S.L., Lusch, R.F., Akaka, M.A. & He, Y. 2010, 'Service-dominant logic: a review and assessment', *Review of Marketing Research*, vol. 6, pp. 125-67.
- Vargo, S.L., Maglio, P.P. & Akaka, M.A. 2008, 'On value and value co-creation: A service systems and service logic perspective', *European Management Journal*, vol. 26, no. 3, pp. 145-52.
- Vargo, S.L. & Morgan, F.W. 2005, 'Services in society and academic thought: an historical analysis', *Journal of Macromarketing*, vol. 25, no. 1, pp. 42-53.
- Vasilescu, B., Serebrenik, A., Devanbu, P. & Filkov, V. 2014, 'How social Q&A sites are changing knowledge sharing in open source software communities', *Proceedings of the 17th ACM conference on Computer Supported Cooperative Work & Social Computing*, ACM, pp. 342-54.
- Vázquez, X.H. 2004, 'Allocating decision rights on the shop floor: A perspective from transaction cost economics and organization theory', *Organization Science*, vol. 15, no. 4, pp. 463-80.
- Venable, J., Pries-Heje, J. & Baskerville, R. 2012, 'A comprehensive framework for evaluation in design science research', *Design Science Research in Information Systems. Advances in Theory and Practice*, Springer, pp. 423-38.
- Viljakainen, A., Toivonen, M. & Aikala, M. 2013, 'Industry transformation towards service logic: A business model approach', *The Cambridge Service Alliance working paper series*, University of Cambridge.
- Voima, P., Heinonen, K. & Strandvik, T. 2010, *Exploring customer value formation: a customer dominant logic perspective*, <<http://urn.fi/URN:ISBN:978-952-232-088-9>>.
- von Hippel, E.A., Ogawa, S. & PJ de Jong, J. 2011, 'The age of the consumer-innovator'.
- Voss, C.A. & Hsuan, J. 2009, 'Service architecture and modularity', *Decision Sciences*, vol. 40, no. 3, pp. 541-69.
- Waldfoegel, J. & Reimers, I. 2015, 'Storming the gatekeepers: Digital disintermediation in the market for books', *Information economics and policy*, vol. 31, pp. 47-58.
- Wang, S., Zheng, S., Xu, L., Li, D. & Meng, H. 2008, 'A literature review of electronic marketplace research: Themes, theories and an integrative framework', *Information Systems Frontiers*, vol. 10, no. 5, pp. 555-71.

- Wang, Y., Min, Q. & Han, S. 2016, 'Understanding the effects of trust and risk on individual behavior toward social media platforms: A meta-analysis of the empirical evidence', *Computers in Human Behavior*, vol. 56, pp. 34-44.
- Weinhardt, C. & Gimpel, H. 2006, 'Market engineering: An interdisciplinary research challenge', *Negotiation and Market Engineering*, no. 06461.
- Weinhardt, C., Holtmann, C. & Neumann, D. 2003, 'Market-engineering', *Wirtschaftsinformatik*, vol. 45, no. 6, pp. 635-40.
- Weintraub, K. 2013, 'Quantified self: The tech-based route to a better life?'
- Weld, L.D. 1917, 'Marketing functions and mercantile organization', *The American Economic Review*, vol. 7, no. 2, pp. 306-18.
- Wetter-Edman, K., Sangiorgi, D., Edvardsson, B., Holmlid, S., Grönroos, C. & Mattelmäki, T. 2014, 'Design for Value Co-Creation: Exploring Synergies Between Design for Service and Service Logic', *Service Science*, vol. 6, no. 2, pp. 106-21.
- Weyl, E.G. 2010, 'A price theory of multi-sided platforms', *The American Economic Review*, vol. 100, no. 4, pp. 1642-72.
- Wheelwright, S.C. & Clark, K.B. 1992, *Creating project plans to focus product development*, Harvard Business School Pub.
- Whittaker, S., Kalnikaitė, V., Petrelli, D., Sellen, A., Villar, N., Bergman, O., Clough, P. & Brockmeier, J. 2012, 'Socio-Technical Lifelogging: Deriving Design Principles for a Future Proof Digital Past', *Human-Computer Interaction*, vol. 27, no. 1/2, pp. 37-62.
- Wiegand, N., Witt, S., Steiner, M. & Backhaus, K. 2015, 'Platform Adoption In Network Markets: Selecting Beneficial Partners To Achieve Market Dominance', *International Journal of Innovation Management*, vol. 19, no. 02, p. 1550028.
- Wieland, H., Polese, F., Vargo, S. & Lusch, R. 2012, 'Toward a service (eco) systems perspective on value creation', *International Journal of Service Science, Management, Engineering, and Technology (IJSSMET)*, vol. 3, no. 3, pp. 12-25.
- Wiese, J. 2013, 'Enabling an ecosystem of personal behavioral data', *Proceedings of the Adjunct Publication of the 26th Annual ACM Symposium on User Interface Software and Technology*, pp. 41-4.
- Williams, L.Y. & Neal, D.M. 2012, 'The Digital Aggregated Self: A Literature Review', *Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), 2012 International Conference on*, IEEE, pp. 170-7.
- Williamson, O.E. 1981, 'The economics of organization: The transaction cost approach', *American journal of sociology*, pp. 548-77.
- Wolf, G. 2010, 'The Data Driven Life', *The New York Times*, 26 April.
- Wolfson College Cambridge & HAT Community 2017, *Wolfson Hat Annual Symposium*, viewed 11th March 2017, <<https://hatcommunity.org/wolfson-hat-annual-symposium/>>.
- Woodall, T. 2003, 'Conceptualising 'value for the customer': an attributional, structural and dispositional analysis', *Academy of Marketing Science Review*, vol. 12, no. 1, pp. 1-42.
- World Economic Forum and Bain & Company, I. 2011, 'Personal data: The emergence of a new asset class', *An Initiative of the World Economic Forum*.
- Wright, K.B. 2005, 'Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services', *Journal of Computer-Mediated Communication*, vol. 10, no. 3.
- Xie, C., Bagozzi, R.P. & Troye, S.V. 2008, 'Trying to prosume: toward a theory of consumers as co-creators of value', *Journal of the Academy of Marketing Science*, vol. 36, no. 1, pp. 109-22.
- Yassine, A., Nazari Shirehjini, A., Shirmohammadi, S. & Tran, T.T. 2011, 'Online information privacy: Agent-mediated payoff', *Privacy, Security and Trust (PST), 2011 Ninth Annual International Conference on*, IEEE, pp. 260-3.
- Yin, R.K. 2013, *Case study research: Design and methods*, Sage publications.
- Yoo, Y. 2010, 'Computing in everyday life: A call for research on experiential computing', *MIS Quarterly*, pp. 213-31.

- Yoo, Y. 2013, 'The tables have turned: How can the information systems field contribute to technology and innovation management research?', *Journal of the Association for Information Systems*, vol. 14, no. 5, p. 227.
- Zannier, F. 2013, *A bite of Me*, viewed 20th April 2015, <<https://www.kickstarter.com/projects/1461902402/a-bit-e-of-me>>.
- Zhu, F. & Iansiti, M. 2012, 'Entry into platform-based markets', *Strategic Management Journal*, vol. 33, no. 1, pp. 88-106.
- Zimmerman, G. 2018, *New Company Hu-manity.co Uses Blockchain to Declare a 31st Human Right, Empowering All Humans to Claim Legal Ownership of Inherent Human Data*, viewed 30th September 2018, <<https://www.businesswire.com/news/home/20180605005581/en/New-Company-Hu-manity.co-Blockchain-Declare-31st-Human>>.
- Zomorodi, M. 2017, 'Day 5: Your Personal Terms of Service', in M. Zomorodi (ed.), *Note To Self*, NPR, <<http://www.npr.org/podcasts/452538677/note-to-self>>.
- Zwass, V. 2010, 'Co-creation: Toward a taxonomy and an integrated research perspective', *International Journal of Electronic Commerce*, vol. 15, no. 1, pp. 11-48.