Curriculum profile: modelling the gaps between curriculum and the job market

Aleksandr Gromov
University of Technology Sydney
PO Box 123, Broadway, Sydney, NSW, 2007
aleksandr.gromov@uts.edu.au

Andrei Maslennikov
University of Technology Sydney
PO Box 123, Broadway, Sydney, NSW, 2007
andrei.maslennikov@uts.edu.au

Nikolas Dawson
University of Technology Sydney
PO Box 123, Broadway, Sydney, NSW, 2007
nikolas.dawson@uts.edu.au

Katarzyna Musial
University of Technology Sydney
PO Box 123, Broadway, Sydney, NSW, 2007
katarzyna.musial-gabrys@uts.edu.au

Kirsty Kitto
University of Technology Sydney
PO Box 123, Broadway, Sydney, NSW, 2007
kirsty.kitto@uts.edu.au

ABSTRACT

This study uses skill-based curriculum analytics to mine the curriculum of an entire university. A curriculum profile is constructed, providing insights about university curriculum design and the match between one institution's curriculum and the job market for a cluster of data-intensive fields. Automating the delivery of diagnostic information like this would enable institutions to ensure that their professionally-oriented degrees meet the needs of industry, so helping to improve learner outcomes and graduate employability.

Keywords

curriculum mapping, curriculum profile, skill-based curriculum analytics, ontologies, skills profile, job market

1. INTRODUCTION

People around the world see universities as an important step in building a successful career [3]. They invest time and finances in undergraduate and postgraduate courses, with a goal to gain new competencies that will help them to find a job [2]. Over the decades, a number of institutions in the tertiary sector have worked hard to adapt their curriculum to market requirements, seeking to prepare more work-ready graduates. However, there is an ongoing debate about whether university efforts to develop students' skills have a noticeable influence upon graduate employability [18]. In particular, employers continue to express doubts that university education does indeed lead to professional competence, claiming that it fails to provide students with the skills they actually require in the workforce [17]. At the same time, undergraduate employment rates are slumping [7], which often leads to further delays in the time it takes students to find work upon graduation, in turn leading to requirements for further professional training [10].

Until recently, much of this debate has been poorly supported by evidence and data. Claim and counterclaim prevail, but a large amount of the data supplied has been *ad hoc*, or cherry picked to support vested interests [20, 9]. However, with the rise of online job advertisements it has become possible to collect data about what potential employers demand in the workplace. A number of datasets can now be created, using data collected from web platforms such as LinkedIn[1], SEEK[2] and Monster[3]. Indeed, vendors such as Burning Glass (BG) technologies[4] now market aggregation services and data that can be used to understand changing trends in the workforce. The next sections provide a brief overview of the ways in which this data can be used.

1.1 Curriculum analytics

Many attempts have been made to understand what gaps there might be in the curriculum offerings of educational institutions. For example, Knight and Yorke [13] describe the Skill plus project as an attempt to manually audit the university curricula for four universities and 17 departments. Trying to find curriculum gaps, Davis et al. [5] conducted a survey of graduates, Lang et al. [15] surveyed industry representatives, Lempp and Seale [16] conducted a study among health students. However, the manual and resultingly not sufficient scalability of this work, has limited the use of this work in linking to workforce needs [11].

Aleksandr Gromov, Andrei Maslennikov, Nikolas Dawson, Katarzyna Musial and Kirsty Kitto "Curriculum profile: modelling the gaps between curriculum and the job market" In: *Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020)*, Anna N. Rafferty, Jacob Whitehill, Violetta Cavalli-Sforza, and Cristobal Romero (eds.) 2020, pp. 610 - 614

[1]https://www.linkedin.com/jobs/
[2]https://www.seek.com.au/
[3]https://www.monster.com/
[4]https://www.burning-glass.com

More recently we start to see data science techniques used to overcome these limitations. For example, attempts have been made to infer curriculum information from student performance and interactions with the curriculum [20, 21], but not the skills offered in the curriculum itself. *Skills based curriculum analytics* [12] makes use of Natural Language Processing (NLP) to map curriculum documents to a defined skills taxonomy, in this case for the purpose of recognising prior learning. A number of other works have made use of NLP in analysing curriculum [1, 8, 24, 19, 23], but only for specific subsets of curriculum (normally computer science). Thus, opportunities are emerging for automating the analysis of curriculum, but how might this be linked with labour market demand?

1.2 Connecting market demand and curriculum

Educational institutions have recognised their role in supplying market demand by offering various courses that prepare graduates to enter the workforce [4]. This transforms the problem of mapping curriculum into one of alignment with labour market requirements [14]. One of the most common methods applied to this task adopts industry skill frameworks and then formulate graduate attributes which match market expectations [14, 22]. This method allows curriculum managers and developers to see gaps, and work to align existing and new courses towards market expectations. It also helps students to plan their study course according to their career desires [25]. However, both frameworks and curriculum are living documents that adapt to the environment, and there is a time lag between acknowledging and implementing new skills and technologies into frameworks and curriculum. Furthermore, this curriculum mapping task is usually completed manually (see section 1.1, which makes it time-consuming, tedious, and prone to mistakes. Worse still, some industries may rarely update their industry frameworks, while other industries might not even have a formal listing of their skill requirements [6].

Overall, while many institutions make use of industry advisory bodies, market reports, etc. to map their curriculum by hand, we are yet to find examples of curriculum mappings to workforce requirements that are supported by the emerging large scale employment datasets that are becoming common in the field. This is the gap that we seek to address here.

1.3 Research questions and contribution

This work aims to use automated methods to map gaps between the subset of the curriculum taught at one institution and labour market demands. We will do this by asking the following three research questions:

RQ1: What is the skills profile for a complete institution?
RQ2: How can we explore the gap between university curriculum and market demand?
RQ3: How can we differentiate the match between contrasting curriculum pathways and labour market demand?

Our contributions include: (i) a preliminary method for automatically constructing a curriculum profile for an institution (ii) a way to compare subsets of curriculum within that institution (iii) a method for finding gaps between closely related set of degree programs and the job market.

2. CURRICULUM PROFILE

This paper introduces the concept of a university curriculum profile. We make use of the BG ontology, which provides a static set of skills that can be consistently mapped into a range of different higher level clusters to extract information about what mix of skills is being taught across an entire institution. We chose the University of Technology Sydney (UTS) curriculum as a data source, which consists of 486 degree programs (termed courses) offered across 9 faculties at the undergraduate and postgraduate levels. In total, there are 3,739 subjects offered across these degrees. Information about the curriculum can be obtained using the UTS handbook (https://www.handbook.uts.edu.au/).

2.1 Method

This section discusses three experiments that have been performed, each designed to extract information about the skills taught at UTS. We start with a course profile across the entire curriculum of UTS in a bid to respond to RQ1 (Section 2.1.1), before performing a deeper analysis of the data analysis curriculum offerings at the same institution to respond to RQ2 (Section 2.1.2). Finally, in Section 2.1.3 we determine how well aligned these data analysis offerings are with local labour market demand, so responding to RQ3.

2.1.1 The UTS curriculum profile

For the first part of our analysis, we performed a skills analysis of the entire UTS curriculum, with a view towards developing a skills profile across the university. This then enables us to drill into sub-profiles for specific degree programs, demonstrating their similarities and differences. Our analysis implemented the following steps:

STEP 1: We scraped the UTS curriculum handbook.
STEP 2: Subject names and descriptions were mapped to lists of skills using the BG content tagger.
STEP 3: Skills were mapped to higher level skill clusters and families according to the BG Skill Ontology.

Drilling into the skill cluster families tagged by BG makes it possible for us to start exploring the distribution of skills taught by each Faculty at UTS. Encouragingly, this method reveals that the majority of the skills developed by each Faculty are in sensible skill cluster family domains, with some spread into other families, in explainable patterns. Thus, the Faculty of Engineering and Information Technology (FEIT) teaches 100% of all *Engineering* and 97% of all *Information Technology* skills covered at UTS; the Faculty of Health (Health) teaches the largest proportion of *Health Care* skills, followed by the Graduate School of Health (GSH); and the Faculty of Business covers the *Business* and *Finance* skills.

2.1.2 Within the data analysis curriculum

For our second analysis, we decided to perform a deep dive into a subset of the UTS curriculum. We chose to explore the Data Analytics related degrees available across UTS. This decision was based upon the existence of three potentially competing degrees that are currently offered at UTS

MDSI: Master of Data Science and Innovation (MDSI)

MIT: Master of Information Technology (Data Analytics)
MBA: Master of Business Administration (Data Analytics)

We sought to explore how much overlap existed in the curriculum associated with these three degrees, and whether there was a possibility that inefficiencies could be identified where the same skills were being taught across multiple subject offerings. We implemented this analysis by following a similar sequence to that presented in Section 2.1.1. Each of the three selected courses consists of a set of compulsory core subjects and an optional choices. We performed two separate analyses, extracting a skills profile for two different course structures: one using just the core subjects required for each of the three courses (core), and a second analysis that added the Data Analytics subject selections for that degree (data). We chose the IT, Business and Analytics skill cluster families for a further investigation of possible reasons of skills changes, because these are the claimed focus of the chosen degrees.

After skill profiling a number of patterns can be noted, for example, the MIT has the most complete coverage in the *Information Technology* skill cluster family, but has almost no coverage of the *Analytics* skills cluster family. However, adding the Data Analytics subject choices to the Core leads to an increase of skills in the *Analytics* skill cluster family (almost to the point where it has the same number of skills as the entire data science-oriented MDSI degree). Similarly, the Core MBA subjects cover all three skill families, but selecting more specific Data subjects leads to a growth in the number of skills in *IT* and *Analytics* skill cluster families. Finally, as expected given the exclusion of the full set of optional subjects no change is observed for the MDSI when expanding with Data subjects.

2.1.3 Finding a gap between market demand and curriculum offerings

For our final study, we combined UTS curriculum data with data about skills sought in the Australian job market:

STEP 1: We selected the top 10 Data Science and Analytics (DSA) skills, found by Dawson et al. [6] and required by the market, and checked if they exist in UTS curriculum and MIT, MDSI and MBA courses.
STEP 2: Then, we selected the top 10 DSA skills that showed the highest growth in the market and cross-checked to see if the UTS curriculum adapts to these rapid market changes.
STEP 3: After that, we selected three DSA Occupations from the BG ontology, retrieved the skills linked with these occupations in the BG ontology and compared them with the skills covered by the UTS curriculum.

Encouragingly, all of the top-10 DSA skills exist in the curriculum. However, some skills are missing from the selected Data Analytics courses. Overall, selected courses cover most of the demanded DSA skills.

At the same time, only four skills with the highest compound annual growth rate (CARG) in 2019 [6] exist in the resulting skills profile. Other technologies and tools, such as Blockchain, TensorFlow, Internet of Things, are missing in the curriculum we analysed. However, some of these skills are yet to be integrated into the skills clusters and families of the BG ontology, which points to their very recent emergence. This gap points to an opportunity for UTS to identify rapidly growing skills that it considers beneficial to deliver: a curriculum gap that could be rectified.

Finally, we retrieved three DSA occupations (Data Analyst, Data Scientist and Business Intelligence Analyst) and their skills from the BG occupation ontology comparing them with the skills profiles for the MIT, MDSI and MBA (see Table 1 for the core subject selections). The results mirror those obtained in the previous section.

Overall, none of the UTS courses has more than 50% of the skills taught that required in our three selected occupations which potentially demonstrates a gap between university curriculum and market demand.

Occupation	Course	Both exist	Only in Occupation	Only in Course
Data Scientist	MDSI	27	73	97
	MIT	42	58	287
	MBA	37	63	153
Data Analyst	MDSI	20	80	104
	MIT	42	58	287
	MBA	34	66	156
Business Intelligence Analyst	MDSI	17	83	107
	MIT	35	65	294
	MBA	32	68	158

Table 1: Three DSA occupations from BG ontology with the number of skills existing and not existing in three selected UTS courses.

3. TOWARDS A CURRICULUM PROFILE

A number of findings about the curriculum taught at UTS emerge from the preliminary curriculum profile presented in the previous section. Firstly, there is a gap between BG skill ontology and the UTS curriculum profile. However, UTS develops knowledge not just software skills. The analysis shows only seven families are more than 50% covered by the UTS curriculum (Business, Economics, Engineering, Environment, Legal, Media and Science). At the same time, the most well-presented in the curriculum IT and Health families cover only 25% and 38% respectively. The majority of skills in these families are missing. However, the reasons for the gap are interesting in themselves, including:

1. **Novelty skills:** there is a lag between new skills emerging (e.g. TensorFLow, WebAssembly) in the market and their incorporation into curriculum offerings.
2. **Legacy skills:** in contrast, some skills (e.g. COBOL, ALGOL and the early versions of Microsoft Server) are present in the ontology but not taught at UTS.
3. **Generalised knowledge:** the role of universities is larger that of simple skill development. We see evidence that UTS is developing generalisable knowledge, rather than the more specific skills.

Another finding from our approach is that the faculties at UTS do appear to have specialisations which largely match the subject materials we would expect to see taught in the faculty. Thus, for example, the Business Faculty has a focus on "Finance" and "Business" (87% and 73% of all UTS skills in this families), Faculty of Engineering and IT includes "IT" and "Engineering" (97% and 100% of all UTS skills in this families) and Science Faculty prepares students in "Science and Research" and "Environment" (100% and 77% of all UTS skills in this families). Overall, the UTS curriculum profile demonstrates that the university tends to develop knowledge, not just skills and an ability to use specific tools.

3.1 Data Analysis projection on different courses

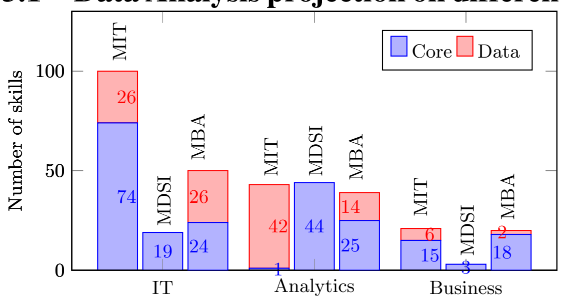

Figure 1: Number of skills in three courses (Core and Data selections) for three BG skill families.

All three selected DSA courses have unique skill profile and prepare different types of Data Experts. By selecting three Data related courses from three different faculties, we wanted to find out if the courses, offered across UTS, are repetitive and faculties lack collaboration. The results in Figure 1 show that each course has a different flavour which aims to develop different fundamental skills. The MDSI focuses on Analytical skills, the MIT focuses upon IT skills and the MBA is giving fundamental knowledge in IT, Business and Analytics. However, expanding the Core subject selection to DSA major, we see that: (i) MBA almost stopped developing business skills and focused on Analytics, which also resulted in the number of IT skills increasing. (ii) MIT becomes far more aligned with the Data Analytics skill cluster, with an accompanying increase in IT and Business skill development (iii) MDSI did not change the curriculum profile at all, which is related more to the extremely open course structure of the degree and the fact that this could not be captured by our method. Overall, each course prepared almost the equal number of skills in Analytics (39 for MBA, 43 for MIT, 44 for MDSI). However, the content and the student pathway are not the same. Students from MDSI are focusing on Data Analytics, as expected from a Data Science course, but MIT students have a strong background in IT, supported by Analytics. Finally, MBA students are well-rounded specialists in IT, Business and Analytics.

3.2 Finding a curriculum gap

The first gap revealed in section 2.1.3 illustrates a disconnection between the understanding of graduate capabilities possessed by universities and the market. While some skill sets are too narrow to be useful to a graduate (e.g. Atlas.ti and Alteryx), others are popular on the job skills for data analysts and data scientists (e.g. SQL and Git). A domain expert is required to make the distinction between these skills, but our method of building curriculum profiles could help with decision support, showing up skills that are essential but still largely unrepresented in the curriculum.

The second gap we found is a lag between the appearance of a new area of knowledge in the labour market and its introduction to the curriculum. This gap should be distinguished from the emergence of new tools alone, although there can be some overlap (e.g. with TensorFlow and Apache Spark). More general skills like Deep Learning, Data Lakes and Random Forests also feature in this list. Similarly, there is a need for Internet of Things and Blockchain specialists which the UTS curriculum is yet to respond to. This second gap is potentially more dangerous because it shows incapability to cover new areas of knowledge in time. However, more analysis is required to investigate the actual demand for these emerging skills. For example, while the CAGR can be very high, this can be achieved by doubling the demand for a skill that was previously only advertised for twice in a time period. Care must be taken to disentangle growth from absolute demand, a task that we reserve for future work.

4. CONCLUSIONS AND FUTURE WORK

In this article, we introduced the university curriculum profile that allowed us to explore the skills taught across an entire university, and to establish that the faculties at UTS do indeed teach the skills we expect. It also enabled us to demonstrate the existence of a potential gap in what was taught, but which was explained by unearthing the too specific and technology dependent nature of many skills in the BG ontology. This gap was explained with the observation that universities should be developing graduates who can generalise knowledge from their skillsets, not just make use of a highly specific tool, and was therefore identified as not critical.

This work can be extended in several ways. Firstly, it will be important to find ways of representing the complexity of a curriculum structure using more than counts. Many of the potential gaps our analysis identified turned out to be understandable once we looked deeper into the skills that were not being taught (Section 2.1.3). Secondly, the analysis can and should be extended beyond the DSA degrees considered here to explore what differences may result from using different skill sets. Third, finding the curriculum gaps will benefit from the development of automated tools for finding changes in labour market demand. Such instruments could be tuned to track changes in real-time and provide historical data for more in-depth analysis of the university curriculum and its development. Finally, the current method of extracting skills from the curriculum cannot identify differences between novice and advanced skills. It is essential that we develop additional tools for evaluating these different levels of skill proficiency.

We believe that the method of profiling curriculum developed here will help institutions to adjust existing courses or initialise new ones as required by the market. It will also help students to choose more effective learning paths according to the market demand. As such, it has the potential to help institutions improve outcomes for all learners.

5. ACKNOWLEDGMENTS

We acknowledge the support of Burning Glass in provisioning the API tools that were used in this study.

References

[1] ACM/IEEE-CS Joint Task Force on Computing Curricula. Computer science curricula 2013. Technical report, ACM Press and IEEE Computer Society Press, December 2013. URL `http://dx.doi.org/10.1145/2534860`.

[2] T. Brock. Young adults and higher education: Barriers and breakthroughs to success. *The future of children*, pages 109–132, 2010.

[3] N. Chhinzer and A. M. Russo. An exploration of employer perceptions of graduate student employability. *Education+ Training*, 2018.

[4] M. Clarke. Rethinking graduate employability: The role of capital, individual attributes and context. *Studies in Higher Education*, 43(11):1923–1937, 2018.

[5] R. Davis, S. Misra, and S. Van Auken. A gap analysis approach to marketing curriculum assessment: A study of skills and knowledge. *Journal of Marketing Education*, 24(3):218–224, 2002.

[6] N. Dawson, M.-A. Rizoiu, B. Johnston, and M.-A. Williams. Adaptively selecting occupations to detect skill shortages from online job ads. In *2019 IEEE International Conference on Big Data (Big Data)*, pages 1637–1643, Dec 2019.

[7] B. De La Harpe, A. Radloff, and J. Wyber. Quality and generic (professional) skills. *Quality in Higher Education*, 6(3):231–243, 2000.

[8] E. Durant, J. Impagliazzo, S. Conry, R. Reese, H. Lam, V. Nelson, J. Hughes, W. Liu, J. Lu, and A. McGettrick. Ce2016: Updated computer engineering curriculum guidelines. In *2015 IEEE Frontiers in Education Conference (FIE)*, pages 1–2. IEEE, 2015.

[9] B. G. T. (Firm)(US). Moving the goalposts: How demand for a bachelor's degree is reshaping the workforce. 2014. URL `https://www.burning-glass.com/wp-content/uploads/Moving_the_Goalposts.pdf`.

[10] J. Keating et al. *Current vocational education and training strategies and responsiveness to emerging skills shortages and surpluses.* National Centre for Vocational Education Research, 2008.

[11] J. Keevy and B. Chakroun. Level-setting and recognition of learning outcomes: The use of level descriptors in the twenty-first century. Technical report, 2015. URL `http://unesdoc.unesco.org/images/0024/002428/242887e.pdf`.

[12] K. Kitto, N. Sarathy, A. Gromov, M. Liu, K. Musial, and S. Buckingham Shum. Towards skills-based curriculum analytics: Can we automate the recognition of prior learning? In *Proceedings of the 10th International Conference on Learning Analytics and Knowledge*. ACM, 2020.

[13] P. T. Knight and M. Yorke. Employability through the curriculum. *Tertiary education and management*, 8(4): 261–276, 2002.

[14] B. R. Konsky, A. Jones, and C. Miller. Embedding professional skills in the ict curriculum. In *ASCILITE-Australian Society for Computers in Learning in Tertiary Education Annual Conference*, pages 883–887. Australasian Society for Computers in Learning in Tertiary Education, 2013.

[15] J. D. Lang, S. Cruse, F. D. McVey, and J. McMasters. Industry expectations of new engineers: A survey to assist curriculum designers. *Journal of Engineering Education*, 88(1):43–51, 1999.

[16] H. Lempp and C. Seale. The hidden curriculum in undergraduate medical education: qualitative study of medical students' perceptions of teaching. *Bmj*, 329 (7469):770–773, 2004.

[17] K. Lowden, S. Hall, D. Elliot, and J. Lewin. Employers' perceptions of the employability skills of new graduates. *London: Edge Foundation*, 2011.

[18] G. Mason, G. Williams, and S. Cranmer. Employability skills initiatives in higher education: what effects do they have on graduate labour market outcomes? *Education Economics*, 17(1):1–30, 2009.

[19] Y. Matsuda, T. Sekiya, and K. Yamaguchi. Curriculum analysis of computer science departments by simplified, supervised lda. *Journal of Information Processing*, 26: 497–508, 2018.

[20] X. Ochoa. Simple metrics for curricular analytics. In *Proceedings of the 1 st Learning Analytics for Curriculum and Program Quality Improvement Workshop*, 2016.

[21] Z. A. Pardos, Z. Fan, and W. Jiang. Connectionist recommendation in the wild: on the utility and scrutability of neural networks for personalized course guidance. *User Modeling and User-Adapted Interaction*, 29 (2):487–525, 2019.

[22] R. S. Pillutla and M. Narayana. Framework integrating multiple dimensions of competency and related pedagogies: A case for it industry. In *2013 12th International Conference on Information Technology Based Higher Education and Training (ITHET)*, pages 1–8. IEEE, 2013.

[23] D. Ramage, D. Hall, R. Nallapati, and C. D. Manning. Labeled lda: A supervised topic model for credit attribution in multi-labeled corpora. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1*, pages 248–256. Association for Computational Linguistics, 2009.

[24] T. Sekiya, Y. Matsuda, and K. Yamaguchi. Curriculum analysis of cs departments based on cs2013 by simplified, supervised lda. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, pages 330–339. ACM, 2015.

[25] L. S. Xun, S. Gottipati, and V. Shankararaman. Text-mining approach for verifying alignment of information systems curriculum with industry skills. In *Information Technology Based Higher Education and Training (ITHET), 2015 International Conference on*, pages 1–6. IEEE, 2015.