12-10-2024

StandFram – A Method to Design and Build a Standards and Frameworks Knowledge Graph

Rosetta Romano
University of Canberra, rosetta.romano@canberra.edu.au

Blooma John
University of Canberra, blooma.john@canberra.edu.au

Marcus Jowsey
None, marcus.jowsey@surroundaustralia.com

Simon Thompson
Australian National University, Simon.Thompson@anu.edu.au

Jayan Kurian
University of Technology Sydney, jayanchirayathkurian@uts.ed.au

StandFram – A Method to Design and Build a Standards and Frameworks Knowledge Graph

Full research paper

Rosetta Romano
Information Systems Capability
Faculty of Science and Technology
University of Canberra
Email: rosetta.romano@canberra.edu.au

Blooma John
Information Systems Capability
Faculty of Science and Technology
University of Canberra
Email: blooma.john@canberra.edu.au

Marcus Jowsey
Surround Australia
Email: marcus.jowsey@surroundaustralia.com

Simon Thompson
Australian National University
Email: simon.thompson@anu.edu.au

Jayan Chirayath Kurian
School of Computer Science
University of Technology Sydney
Email: JayanChirayathKurian@uts.edu.au

Abstract

There are many cybersecurity standards and frameworks, each representing the interests of different communities. While these standards and frameworks are rich sources of information, they are often complex, overlapping, scattered across different websites, hard to compare, subscription-based, and unstructured. For smaller businesses with limited resources, navigating these complexities in a rapidly evolving cyber threat landscape is particularly challenging. To address these challenges, industry and academia partners collaborated to research the collocation of cybersecurity standards and frameworks relevant to Australian businesses into a knowledge graph. This graph structures the information and describes both explicit and implicit semantic connections. This paper introduces the STANDFRAM, a novel method for designing, building, and evaluating a standards and frameworks knowledge graph. The goal is to extend its use beyond the originating communities. The STANDFRAM method was evaluated by experts, and the Knowledge Graph was assessed by users to determine the utility of the developed artifacts.

Keywords Cybersecurity, Standards, Frameworks, Knowledge Graph, Design Science Research (DSR), StandFram, Smaller Businesses.

1 Introduction

All organizations and businesses today strive to protect their cyberspace. Cyber threats represent Information Technology (IT) leaders' biggest concern (Kappelman et al. 2019). While more big companies are being hacked, small businesses are attacked more frequently, with one in five small to medium businesses being hacked yearly (Segal, 2022). Smaller businesses report affordability barriers preventing them from accessing technologies, trained cybersecurity staff, and external security services that can keep their organizations safe from cyber-attacks (Cynet, 2022). Limited resources concerning the budget, lack of highly skilled cyber experts to keep up with the rapidly evolving cyber threat landscape, lack of awareness and understanding of cybersecurity risks as well as the importance of implementing a framework, relying on third-party vendors and suppliers for various services and products, and also trusting these entities with sensitive data and systems are all different levels of challenges experienced by small and micro businesses today (Tam et al. 2021; Cartwright et al. 2023).

Global regulation increasingly relies on alternatives to legal rules called standards (Kerwer, 2005). In principle, following a standard is voluntary (Brunsson and Jacobsson, 2002) but compliance with them may be required by certain certifiers compliance to demonstrate that a particular situation is being managed (Brunsson and Jacobsson, 2002). A standard is a rule that others have provided about how to organize, what policies to pursue, what kind of services to offer, or how to design their products (Brunsson and Jacobsson, 2002; Kerwer, 2005). A framework refers to the overall structure to support a system (Romano et al. 2024). A variety of cybersecurity tools and techniques, standards and frameworks, processes and procedures have been developed over the years to identify and assess security vulnerabilities, i.e., the flaws in a system or its design that allow an attacker to execute malicious commands, access data in an unauthorized way and conduct various denial-of-service attacks (Humayun, 2020). Cybersecurity standards explain and provide methods one by one, specifying what to do to complete a process and clarifying the methods provided by the standard. In contrast, a cybersecurity framework is a general guideline that covers many components or domains without specifying the steps to follow (Baskerville et al. 2014; Seeburn, 2014). Some examples of popular cybersecurity frameworks, standards, and guidelines are ISO 27001 and the NIST Cybersecurity Framework 2.0. Although rich information exists in cybersecurity standards and frameworks (CSF), they are complex, overlapping, and electronically scattered, and some attract annual fees to access them. Different organizations have developed the range of CSFs available today, and choosing between them is challenging for businesses. Selecting the most appropriate standard or framework is a serious decision for an organization that must match its suitability with its business demands (Taherdoost, 2022), particularly for smaller businesses. Relying on third-party vendors and suppliers for various services and products and trusting these entities with sensitive data and systems is a challenge experienced by smaller businesses today (Tam et al. 2021).

The problem is that standards and frameworks are generally community or problem specific, have complex overlapping information, and are not easy to find and hence posits a significant research gap. For example, understanding applicable cyber security standards and frameworks applicable in Australia requires a comparison of state, national, international, and globally developed resources (see Figure 1). Standards may provide hundreds of controls to manage cyber security[1] without a map of where to start. Different standards offer protection for similar data assets such as sensitive information without describing the new information provided, or a comparison of information covered by existing standards rather there are mappings produced to show the connections[2]. Some of the most known standards are published behind Standards organisations paywalls, published as pdfs on various government websites[3].

[1] ISO/IEC 27002 – Information security, cybersecurity and privacy protection – Information security controls offers over 1000 cyber security controls, and the Information Systems Manual (ISM) offers 819, NIST 800-171- Protecting Controlled Unclassified Information in Nonfederal Systems and Organizations has 97 controls, and NIST 800-53 (Security and Privacy Controls for Information Systems and Organizations) has 1000 controls.

[2] NIST 800-53 Appendix H-2 maps its controls to ISO 27001 controls.

[3] Standards paywall – AS/NZS ISO/IEC 27002:2022 and AS/NZS ISO/IEC 27001:2023; PDFs on websites – NIST 800-171 and 800-53, Various government websites – ISM, Australian Cyber Security Centre (ACSC) Mitigation Strategies Maturity Model (including the Essential Eight and Defence Industry Security Program.

While this may present a problem for all organisations, this research attempts to remove the barriers for smaller Australian businesses to access rich resources in CSF to improve the cybersecurity maturity of smaller businesses. To address this research gap, the aim of this research is to explore the use of a knowledge graph (KG). A KG is advanced data structures that integrate information from multiple sources, capturing entities, tasks, and formally describing semantic connections between them (Nickel et al. 2016). KGs can mine, organize, and effectively manage knowledge from multiple large-scale data sources to improve information retrieval quality (Pujara et al. 2013). They are particularly useful for mining, organizing, and managing large-scale data to improve the quality of information retrieval (Paulheim, 2017). KGs enhance threat detection, situational awareness, and automated reasoning capabilities, thereby improving the accuracy and efficiency of cyber threat intelligence and response (Almgren et al. 2019). They aid in visualizing complex relationships, understanding the cyber threat landscape, and mapping threat actions to countermeasures (Pan et al. 2009). A KG can effectively organize, link, and represent knowledge to be efficiently utilized in advanced applications (Chen et al. 2020).

State-of-the-art research highlights that no common knowledge graph captures the domain knowledge for using CSF (Cartwright et al. 2023), particularly for smaller businesses. Approaches to improving smaller businesses' security posture should provide simple and practical advice (Bada et al. 2019). This process of constructing such a knowledge graph is not a trivial task, hence this research. This paper addresses the following research question: ***How can smaller Australian businesses use a Knowledge Graph to improve access to applicable Cyber Security Standards and Frameworks?***

To address the research question, this research paper describes the design and build of a cybersecurity standards and frameworks knowledge graph (CSFKG). It will be a knowledge base that will serve as a backbone for collecting, consolidating, sorting, and sharing cybersecurity-related rules and regulations for access in a single repository for smaller businesses in Australia. This paper contributes the development of a method for collating standards and frameworks into a knowledge graph and demonstrates it using cyber security standards and frameworks for smaller businesses. In this paper, smaller businesses include Indigenous, micro, small, or medium businesses employing less than 200 people (ABS, 2024). In Australia, an Indigenous business is also known as an eligible Aboriginal and Torres Strait Islander owned business that is at least 50% owned by an Aboriginal person(s) and/or a Torres Strait Islander person(s). It is either a sole trader, partnership, incorporated entity, or trading through a trust and has a current Australian Business Number (Queensland Government, 2023). A microbusiness employs 0-4 employees and is more likely to be a sole proprietor and/or partnership and include many non-employing businesses; a small business employs 0-19 people, while a medium business employs between 20-199 people (Australian Bureau of Statistics, 2016). The remainder of this paper presents a literature review, methodology, discussion and conclusion. The steps involved to design and develop the CSFKG are detailed throughout the sections.

2 Literature Review

The National Institute of Standards and Technology (NIST) defines cybersecurity as the prevention of damage to, unauthorized use of, exploitation of, and, if needed, the restoration of electronic information and communications systems and the information they contain to strengthen the confidentiality, integrity, and availability of these systems. Cybersecurity is a global contemporary issue concerning the management and utilization of information technology (IT) (Kelley, 2008). Cybersecurity supports social sustainability goals as it is a frequently used tool for data management to secure data and protect privacy (Piccarozzi et al. 2023).

The practical and cohesive application of cybersecurity practices in industry is accomplished by adopting cybersecurity frameworks that provide structure and methodology (Carias et al. 2020). Despite this, the unstructured and semi-structured format of the current CSF restricts the ability to navigate and link them. Table 1 lists some, not all, of the CSF applicable to all Australian businesses.

Standard		Framework	
S1	ISO/IEC 27001 Information Security Manual	F1	NIST Cybersecurity Framework
S2	ISO/IEC Information Security, Cybersecurity and Privacy Protection – Information Security Controls	F2	CIS (Critical Infrastructure Controls)
S3	NIST 800-53 Security & Privacy Controls for Information Systems and Organizations	F3	COBIT (Control Objectives for Information Technology)
S4	VIC (Victorian State) Protection data Security Standard	F4	CCM (Cloud Control Matrix)
S5	ISM (Information Security Manual)	F5	PSPF (Protective Security Policy Framework)
S6	PCI DSS (Payment Card Industry Data Security Standard)	F6	AESCSF (Australian Energy Sector Cyber Security Framework)
S7	ACSC (Australian Cyber Security Centre) Mitigation Strategies Maturity (including the Essential Eight)	F7	Tasmania's Protective Security Policy Framework
S8	NIST 800-171 Safeguarding sensitive information on federal contractors' IT systems and networks	F8	South Australian Cyber Security Framework
		F9	DSPF (Defence Security Principles Framework)
		F10	DISP (Defence Industry Security Program) 16.1 Defence Industry Security Control

Table 1. Cybersecurity Standards and Frameworks Applicable to Australian Businesses

There are many CSFs applying to Australian businesses. Figure 1 depicts many of these applying to particular Australian States (Standard (S)4, and Framework (F)8), all of Australia (S5,&7 and F 5,6,9 & 10), Australia and internationally (S3,8 and F1&2), Australia and globally (S1,2,6 and F3&4). CSFs the use of various standards and frameworks around Australia where this study is situated.

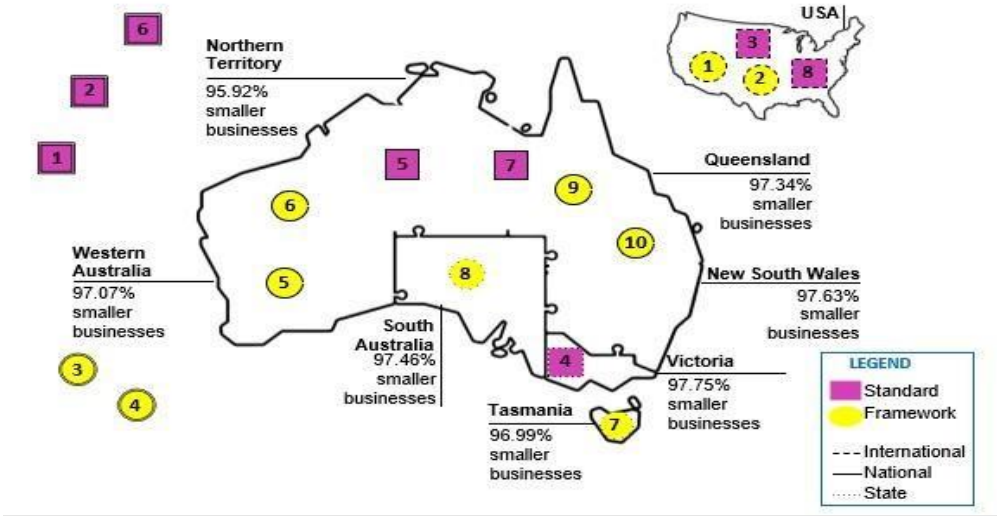

Figure 1. State/National/International/Global Cyber Security Standards and Frameworks Applying to All Australian Businesses

The complexity of standards and frameworks, such as NIST Cybersecurity Framework or ISO 27001, can be daunting for small business owners. Understanding and mapping the requirements to their specific business needs can be challenging without specialized knowledge (Tam et al. 2021). Smaller businesses lack cybersecurity expertise and financial resources compared with that found in large enterprises and lack of expertise is a significant challenge (Alahmari and Duncan, 2020). IT maturity of organizations is also an important challenge for small businesses; many will be new entrants. There is a need to recognize the challenges experienced by small businesses.

In cybersecurity, KGs have been utilized to represent vulnerabilities, weaknesses, attack patterns, attack strategies, and threat detection (Jia et al. 2018). For example, the SEPSES CKG integrates an abundance of cybersecurity information into a regularly updated public CKG for improved vulnerability assessment (Almgren et al. 2019). Additionally, ontologies and KGs are frequently mentioned together, providing formal representations of entities, and maintaining distinct definitions, which aids in the interoperability and standardization of data (Pan et al. 2009). Studies have also demonstrated the use of machine learning techniques with CKGs to enhance the detection and analysis of cyber threats, showcasing innovative methods to process and interpret vast amounts of security data (Zhu et al. 2023).

Despite the various applications and advancements in CKGs, there is limited focus on how these frameworks can be specifically tailored to address the cybersecurity needs of smaller businesses (Cartwright et al. 2023). There is a significant gap in creating a common CKG that captures domain knowledge for utilizing cybersecurity frameworks tailored for smaller businesses (Bada et al. 2019). While there are several methods and metrics proposed for evaluating KGs, such as the K Score, I Score, and C Score derived from the science of science, information theory, and causality perspectives, there is still a need for standardized and objective approaches to ensure their effectiveness across various applications (Zhu et al. 2023). The complexities involved in human-centered evaluation and the practical implications of these metrics require further research (Paulheim, 2017). The dynamic nature of cyber threats and the continuous evolution of attack strategies pose a challenge in keeping KGs updated and relevant (Chen et al. 2020). The need for continuous updates and the integration of new data sources into KGs remain areas that require ongoing attention. Additionally, the development of sector-specific KGs, such as those for critical infrastructure or healthcare, and their integration into existing cybersecurity frameworks is still an area with potential for further exploration and research (Mayer et al. 2017).

This research proposes the CSFKG to serve as a comprehensive resource, facilitating a better understanding of existing and applicable cybersecurity standards and frameworks (Romano et al. 2024 (2) and make the knowledge accessible in a single repository to support the development of a Questions and Answering (Q&A) system for smaller Australian businesses. This Q&A system should improve the accessibility to the rich information about cybersecurity for these businesses and help them to improve their cybersecurity maturity.

The majority of Cybersecurity Knowledge Graphs (CKG) reported in the literature are used to represent cybersecurity vulnerabilities (Jia et al. 2018), weaknesses and attack patterns (Kiesling et al. 2019), attack strategies (Chen et al. 2021), and threat detection (Kurniawan et al, 2022). Other researchers use CKGs for security analysts, to assist the development of cyber-situational awareness and acquire cyber threat intelligence (Sikos, 2023). In addition, a few studies have created CKGs to address the challenges posed by social engineering attacks (Wang et al. 2021). Existing work incorporates the many varied perspectives on cybersecurity frameworks in a concise view, integrating heterogeneous data and knowledge schemas from different cybersecurity systems and the most used cybersecurity standards for information sharing and exchange (Syed et al. 2016). Yet other work incorporates the many varied perspectives on cybersecurity frameworks in a concise view, orating the many varied perspectives on cybersecurity frameworks in a concise view, allowing contrasting different intentions and distilling shared concepts (Azmi et al. 2018). Another Security, Privacy, and Safety Enhanced Smart Environments (SEPSES) CKG, integrates an abundance of cybersecurity information into a regularly updated and public CKG for improved vulnerability assessment (Kiesling et al. 2019). Although the application of CKG is diverse, to the best of our knowledge, none of the studies have examined CSF to address the needs of smaller businesses. This paper describes the design and build of a CKG based on prominent CSF to address this gap.

The researchers applied a semantic graph using the Resource Description Framework (RDF) standard W3.org using the structure of a triple in the form of <subject> <predicate> <object> (Figure 2) synonymous with <noun> <verb> <noun> (W3.org, 1999). RDF is a W3C standard to represent knowledge as a semantic graph in which the nodes represent entities, concepts or literal values, and the edges represent relations between nodes in a Triple (Syed et al. 2016). The RDF triple includes a logical statement that is both human and machine readable.

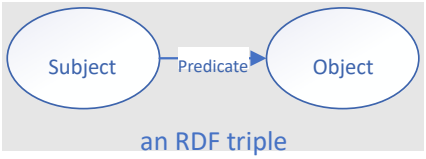

Figure 2. An RDF Triple

The researchers used the enterprise knowledge graph platform, Stardog to design and develop the CSFKG. The CSKFG is a semantic knowledge base that can organize and manage the complex data available in CSF from diverse sources (Sikos, 2023). RDF specifications allow linking. Due to the importance of knowledge graphs in processing heterogeneous information within a machine-readable context, a considerable amount of research has been conducted continuously on these solutions in recent years (Dai et al. 2020; Kong et al. 2022). This paper describes the design and evaluation of the CSFKG to answer the research question.

3 Methodology

A Design Science Research (DSR) methodology is used in the research reported in this paper. DSR involves designing and evaluating artifacts, such as new software, processes, algorithms, or systems, to improve or solve an identified problem. It can also include constructs, models, methods, or instantiations (Hevner et al. 2004). The researchers developed a CSFKG artifact. The development of the artifact follows a DSR design pattern: build, evaluation and if necessary, iterate. IT artifacts are intended to solve identified organizational problems. It involves a rigorous process to design artefacts to solve observed problems, make research contributions, evaluate the designs, and communicate the results to appropriate audiences (Dresch et al. 2015).

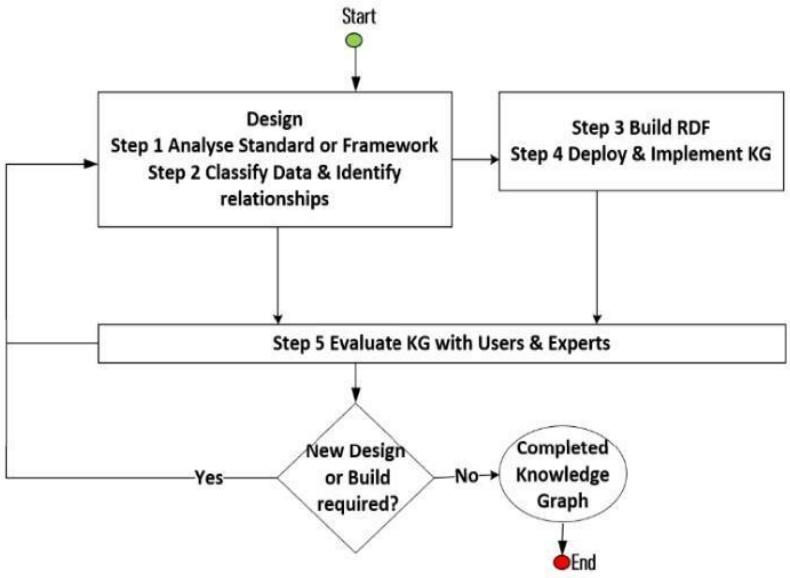

Figure 3. The conceptual framework and architecture (StandFram) architecture

A conceptual framework depicts how the theory related to the DSR is connected and how it will answer the research question. The steps are depicted in Figure 3. In **step one**, the researchers analyzed the CSF content (Saunders, 2011) listed in Table 1 to determine which standards and frameworks would be included. In Australia, the Department Industry Security Program (DISP) is a framework established to ensure that all suppliers meet specific security requirements. The DISP provides four standards that the Department will accept, and these were used as the first selection from the many available standards for the researchers[4]. After the initial selection, the researchers considered the ISO/IEC 27002 would provide a logical extension. Extensions thereafter were made in agreement with all researchers. In **step two**, the researchers detected and classified the entities and the inter- and intrarelationships from the standards and frameworks content to transform metadata/data into structured data from the classification captured in sets of enumerated values (including identifiers and definitions), declared meaning for terms used in controls (including terms and associated descriptions, and including acronyms), cyber security control sets (including: associated identifiers, metadata including: Links between classifiers and controls), and different types of sub-classifications (e.g., maturity levels, security classifiers, etc.). Metadata describing cyber security frameworks and standards was modelled in the ontology to reflect existing structure and constraints. Metadata modelling included support from various upper-level ontology (W3C based) and the Data Privacy Vocabulary (DPV) (W3C, 2022). The DPV enables the expression of machine-readable metadata about the use and processing of personal data based on legislative requirements such as the General Data Protection Regulation (GDPR) (ibid.). The vocabulary provides classifications for activities critical to cyber security to enable record keeping with links to controls. For example, the concept of 'consent' can be linked to controls in policies and other documents and information pertaining to 'consent provided' can then be logged. The DPV is comprehensive and can be adapted and applied to activities such as planning, execution, status capture and other standards-based activities within a business. The DPV has been included to support communities of software developers who are building applications that source cyber security meta data from the CSFKG. In this research, common contexts provided in the cyber security frameworks/standards were ignored, for instance the scheduled publishing events for new versions. The output of this step is a spreadsheet with structured data coded from the CSF content.

The **third step** required translating the structured data recorded in the two-dimensional spreadsheets and any related data from the standards and frameworks to RDF-based triples. The RDF triples were derived using the individual ontologies in the spreadsheets and the connections between the individual spreadsheets. The researchers used spreadsheets to delineate between the different but similar predicates or rules (Bersagol et al. 1996) and to represent RDF (Han et al. 2008). See Figure 4 for the basic triple in each spreadsheet (intra-standard or framework) or between the spreadsheets (inter-standard or framework). **Step four** provides a backend RDF-based knowledge graph that enables Application Programming Interfaces (API) based on saved queries in the KG. The last and **fifth step** focus on the CSFKG user. In this paper, the sample users are the smaller businesses. Context as the basis of the ontology design process, was specified in terms of a smaller business undertaking planning and implementation of cyber security strategy. Design consideration was also given to questions asked about cyber security by smaller business stakeholders. The CSFKG design is intended as an application backend that is extensible and that minimises the need to develop data logic in applications whilst supporting integration with existing/emerging business applications, and to provide support services for cyber security standards planning/implementation/compliance activities (e.g. business record keeping for security audits, implementation planning and execution, people and procedure).

[4] . The Defence Security Principles (DISP) Framework Annex A provides a requirement that entities seeking certification meet or exceed the following information standards across all of the Entity's ICT corporate systems used to correspond with Defence, including, but not limited to the Australian Signals Directorate Essential 8, ISO/IEC 27001 Information Security management; NIST SP 800-171 Protecting Controlled Unclassified Information in Non-Federal Systems and Organisations, and Def Stand 5-138 Cyber security for Defence.

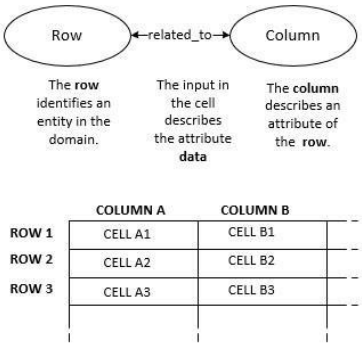

Figure 4. Intra (single) or Inter (between two or more) Standards and Frameworks triple

By representing the CSF as spreadsheets, the researchers have enabled a semantic formalization of descriptions of concepts, classification of the data, and mapping. Formal descriptions of the rows and columns (concepts) support constructing complex relationships between the concepts developed as Description Logic (DL), a formal knowledge representation language. Formal data classification, such as the cybersecurity controls and evaluation frameworks supporting constraint-based deductive and automated reasoning. Formal mapping across concepts and the ability to reliably employ linking semantics (including across concepts that are similar but not equal) to enable concept mapping where not formally declared. The formal descriptions, classification and mapping motivated the researchers to consider adopting (RDF-based) KG technology as the preferred approach to navigate across CSF.

Information systems (IS) research investigates the phenomena of the development and use of information systems and technology for human enterprise, covering IS use by individuals, teams, organizational units, and organizations as well as IS use by communities, markets, industries and societies (Grover et al. 2015). Research in the IS field examines more than just the technological system, or just the social system, or even the two side by side; in addition, it investigates the phenomena that emerges when the two interact (Gregor and Jones, 2007). The study reported in this paper applies a form of design theory called utility theory to evaluate that the artifact being StandFram, the method, has utility, i.e. the method is useful for solving or improving a problematic situation (Venable, 2006), and by applying StandFram, the problem will be improved (Iivari, 2020). StandFram is an exaptation to extend a known solution (KG) to a new problem (Gregor and Hevner, 2014) of collocating disparate community-based standards and frameworks for a broader user base. Knowledge about KGs is high however, the use of KGs to improve cybersecurity maturity of users is low. Exaptation projects should demonstrate that the extension of known design knowledge into a new field is non-trivial and interesting in order to claim a knowledge contribution (ibid). The design and build of StandFram is non-trivial and emerged as an original idea from interconnections conceptualized by the academics and fleshed out by industry practitioners in partnership. This exaptation research is distinguished from adoption because it is the first attempt create such a repository as a basis for a Q&A system. Subsequent extensions will constitute adoption.

Gregor and Jones (2007) provide six mandatory and testable propositions concerning the design product to understand the effects of the design artefact (Iivari, 2020) to which the researchers apply the following:

1. **Purpose and scope:** What the system is for? StandFram provides a method to analyse standards and frameworks developed for, and by other communities.
2. **Constructs of theory:** StandFram establishes a baseline for incremental development. Noting the instantiation of StandFram in this instance is applied to cybersecurity standards and frameworks applicable to the Australian context.
3. **Principle of form and function:** The architecture of the IS artifact IS development method is provided in Figure 5.

4. **Artifact mutability** (i.e., the changes in the state of the artifact anticipated in theory): The development of StandFram supports similar form of KG and its mutability in any other domain. For example, standards and frameworks for responsible AI.
5. **Testable propositions**, i.e., the truth statements about the design theory: The StandFram can be used to build or extend a KG, and this has been tested by incrementally introducing one standard or framework at a time.

3.1 StandFram Design

This study highlights the findings in various dimensions based on the methodology presented. First, application developers focusing on smaller businesses can use the steps to meet their needs. For example, to plan for implementing threat mitigation, increase the smaller business maturity, and identify and secure critical business information. The researchers acknowledge the diversity of smaller businesses, the diversity of applications that they can use, and the complexity of CSF. For these reasons, the researchers developed the KG design specifications in Table 2.

Provides a repository with definitions of concepts that can be applied and extended to many (yet to be determined) application domains.
Provides a tool for the education of communities about available standards and frameworks.
It supports both a design-level backend and a user-level front-end.
Practical and flexible.
Supports top-down and bottom-up ontology development.
Is extensible to other standards and frameworks.
Supports both inter (singular) and intra (many) standards and frameworks.
Acknowledges that the KG is one conceptualization, but there may be many.

Table 2. StandFram Knowledge Graph Design Specification

A representation of the CSFKG appears in Figure 5. The ontology for the CSFKG is captured as spreadsheets. These ontologies represent the information architecture that enables linking of information from across frameworks and standards, the ability to safely link between frameworks and cyber security controls through inferencing (including where linking is not formalized by the associated Standards bodies), search and discovery, linking to natural language capability to enable natural language querying of the CSFKG. While not yet tested, the information architecture could support a business with the development and capture of cyber security strategies through selecting from standards-based controls supporting both new and pre-existing strategies, and the capture of provenance of changes to the frameworks/standards in the CSFKG.

Notwithstanding that KG is greater than the sum of the individual data parts, the coloured rectangles represent the different components within the CSFKG individually. The **blue (upper middle)** rectangle includes all CSF provided by the spreadsheets—for example, ISO 27001 and ACSC Essential 8. The **red (upper far-right)** rectangle encapsulates graphs for the cybersecurity reference datasets that classify controls and associated frameworks—for example, personnel security and technical security. The **orange (upper left)** rectangle encapsulates the graphs for glossary terms and acronyms used and provided by the standards and frameworks—for example, acronyms such as CISO (Chief Information Security Officer), any definitions of terms, and metadata such as the description of the spreadsheet column headings. The **black (lower right)** rectangle encapsulates the stepwise strategy structure for developing applications that assist with strategy and planning for standards implementation. Based on the example strategies offered by the Australian Signals Directorate (ASD) to help small businesses prioritize mitigation strategies for incidents caused by cyber threats. The **green (lower left)** rectangle

captures the provenance data for changes to the CSFKG in a Provenance Graph (W3C-PROV-O, 2013) and references logged changes in an instance of GitHub.

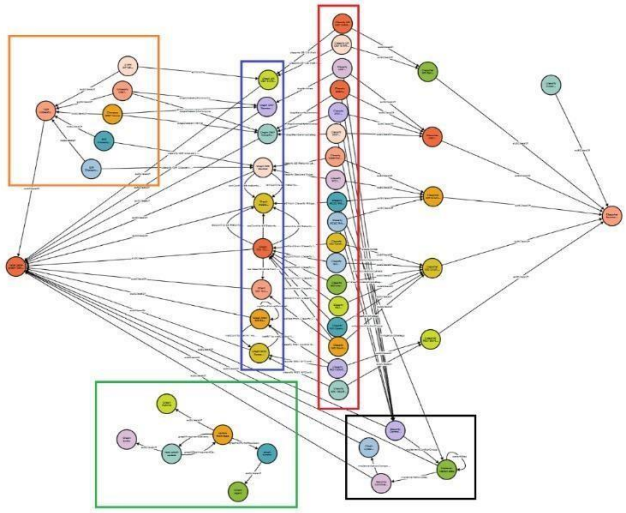

Figure 5. The Cybersecurity Knowledge Graph

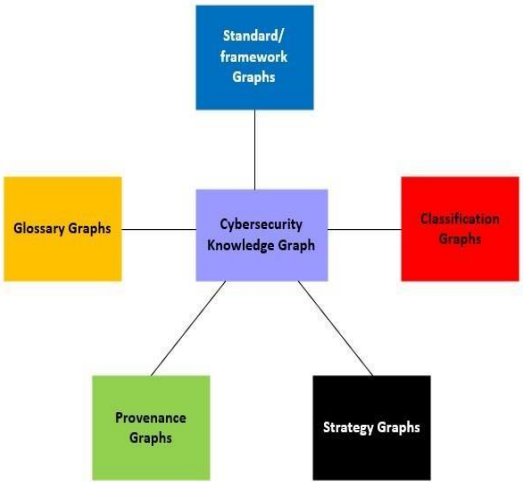

Figure 6. An overview of the Cybersecurity Standards & Frameworks Graph

A summarized model of the graphs applied (appearing as coloured rectangles in Figure 5) in the CSFKG architecture is provided in Figure 6. The lines depict the links between the graphs that are declared in frameworks/standards (lines joining circles) but not the inferred links that are inferred in graph query result sets. A limitation of Figure 5 is that it does not represent the DPV which supports the application integration with the framework or standard metadata and data.

In its entirety, the CSFKG enables a semantic understanding of the connected four types of graphs as depicted in the same-coloured and modelled rectangles in Figure 6. Glossary graphs with descriptions of concepts and acronyms appear in the orange rectangle. Classification graphs provide the standards and framework categories in the **red** rectangle. Strategy graphs supporting the ACSC stepwise planning structure in **black** rectangle. Standards/Frameworks graphs include the **blue** rectangle's controls, threats, and assessment frameworks. Finally, the Provenance graphs, the information about the agents, activities, and entity changes to the CSFKG in **green**. The greater than its parts emerge from the semantics that these connections enable, including description logic, constraint-based reasoning, inferencing, and linked concepts that appear as the different graph types connected in the CSFKG architecture.

3.2 StandFram Evaluation

The researchers presented the CSFKG to experts for their evaluation. As the final steps, the StandFram was evaluated. An ethics approval to seek both expert and user feedback enabled the evaluation of the StandFram. Experts reviewed the inclusions in the CSFKG and commented on extensions to it. Users, i.e., representatives from smaller businesses, were asked about their maturity and knowledge about CSF and were also asked what questions they had about cybersecurity. Findings related to StandFram evaluation are detailed in the following section.

Several methods can be used to evaluate DSR, including ex-ante (before artifact construction) and expost (after artifact construction), naturalistic settings such as case studies, and artificial settings for example, lab experiments (Pries-Heje, et al. 2008). Where the design artefact is a method, evaluation can be done against some criteria or opinions of the method from the users. The overall efficacy that satisfies its users and efficiencies can also be evaluated (Pries-Heje, 2008). The StandFram evaluation phase provides a set of steps or processes to build a KG from the standards and/or frameworks supporting a community of interest. A specification for the StandFram evaluation method is summarized in Table 3.

The specification of the StandFram evaluation method found in Table 3 can be used to conduct an expost evaluation in real settings with human subjects who are being asked about its application for another community or automatically using experimental designs.

The data to test the architecture of the CSFKG is drawn from interviews with experts in cybersecurity or in data/information management. During the interview, the CSFKG was explained and demonstrated to experts. The following insights are drawn from the interviews held so far:

- They were excited about the focus on smaller businesses. For extensions, they were interested in covering critical infrastructure. For technology, they were enthusiastic about a chatbot for smaller businesses and indicated that they would watch to see what artificial intelligence could offer in this space. The 3-minute video is clear, and they would happily share it.
- Respondents all commented on their support for the research.

1. A method that supports the collection of standards and frameworks applicable to a community in a single repository (in this research it is a collection of cybersecurity standards and frameworks that apply to all Australian businesses).
2. A method that is documented for technical experts (in this research it is for implementation of a KG using Stardog).
3. A method that is described for non-technical users (in this research it is for smaller business users).
4. A method that can be followed for and by other communities.
5. A method that detects both common and specific community questions (in this research it is cybersecurity answers for Australian smaller businesses from the cybersecurity standards and frameworks that apply to them).

Table 3. Specification requirements of the StandFram

The data used to evaluate the CSFKG is drawn from interviews with representatives from smaller businesses. The following insights have been drawn from the interviews held so far:

- Smaller businesses are excited to be the focus of this research (all respondents).
- There are common questions that sit outside of the standards and frameworks like: what is cybersecurity; what customer information should I prioritize; and my business does not collect any private information from customers, only their name, phone number, and credit card details when they pay for their goods/services does it still apply to me?
- Frequent questions apply to organizations working with the Department of Defence: how are the metrics communicated to members? Is there specific easy-to-locate guidance on implementing the non-Australian Signals Directorate frameworks, including NIST? If not, then how can industry comprehensively and consistently implement these frameworks? Is there one out of the four recommended frameworks that can be implemented for an organization with little time, expertise, and financial resources (such as a start-up or a micro business in the non-IT Government sector?), Does compliance with the suggested frameworks ensure compliance with AS 4811-2006 Employment screening?
- Some questions apply to organizations that must show they have an appropriate cybersecurity program: Which framework and standards are the easiest, cheapest, quickest, and with the least administration to adopt? The ISO/IEC 27001 does not offer implementation guidance, which may lead to inconsistency and misinterpretation in adopting standards across a specific industry. How can this inconsistency in adoption be mitigated? Are small businesses expected to build upon these frameworks?
- The 3-minute YouTube video (https://www.youtube.com/watch?v=RE59wRX69Sc&t=105s) was clear, and they would be happy to share it with others.

The interviews with experts and smaller businesses identified questions that the CSF did not answer. In future, these answers need to be considered but for now, knowing which questions are being asked and which questions can be answered is the focus of this research. The observations reported so far are limited by the number of interviews held. As more interviews are conducted, more questions from smaller businesses will emerge. The target is to expand the evaluation to 10 experts and 50 users. This is relevant because of the diverse sectors that smaller businesses are operating in. However, it does not detract from the StandFram evaluation method, which is the focus of this paper.

4 Discussion

The study was driven by an interest in understanding how a KG could be used to improve cybersecurity for smaller businesses, given that there is already so much information held in standards and frameworks. As industry-supported research, the industry partners were keen to investigate how smaller businesses can ingest voluminous information about cybersecurity.

Cybersecurity is a problem for all businesses. While smaller businesses are a large fraction of the economies of many countries, research is rarely focused on this sector (Chidukwani et al. 2022). While knowledge graphs have been used to manage cybersecurity, the CSFKG developed in this research uses the information existing in the many standards and frameworks to help smaller businesses improve their cybersecurity maturity. Recalling the research problem is that standards and frameworks are generally community or problem specific, have complex overlapping information, and are not easy to find, navigate and use. Collocating standards and frameworks in a system enabling complex querying to be undertaken with a simple user interface will improve the accessibility to this rich source of information. The research reported in this paper presents the StandFram evaluation method to build a KG that collocates and improves accessibility to standards and frameworks for users outside the intended communities.

In this paper, StandFram is used to develop a CSFKG backend application. For the research, the application for cybersecurity is narrowed by the focus on smaller businesses to understand the questions they ask about cybersecurity. Can the answers to these questions be found in the CSFKG, if a front end of the CSFKG is a Q&A system that is provided for smaller Australian businesses? This industry-partnered research uses DSR methodology encompassing a cycle of design, build, evaluation, and iteration if necessary. The three artifact types are the ontology spreadsheets, the CSFKG, and the future Q&A system. DSR enabled design, build, and evaluation iterations of solutions to occur without impeding the creative problem-solving between the partners. The StandFram specifications were also evaluated. StandFram provides a repository of definitions of concepts that can be applied and extended to many (yet to be determined) application domains. The glossary graphs provide a repository of definitions that can be extended. Other applications for these graphs will emerge in time, but one example could provide the harmonization of terms across standards and frameworks.

As communities build more standards and frameworks, StandFram can be used as a tool to explain these to others. StandFram is practical and flexible, allows for understanding standards and frameworks in different domains, and is extensible. StandFram can support a single standard or framework, or the connections between multiple standards or frameworks. Both top-down using a standard or framework and bottom-up using the information in a standard or framework are ways to develop the ontologies that are recorded in the spreadsheets and the choice is made by the builder of the KG. StandFram offers the implementation of a conceptualization as provided in Figure 3, but the researchers acknowledge that there are always other valid ways to conceptualize.

In the future, the StandFram evaluation method will be used and improved by others. As other standards and frameworks are added to the CSFKG, new search capabilities across several interrelated KGs will emerge (Sarrafzadeh et al, 2014). As versions of standards and frameworks are published, the Provenance Graphs will provide new research opportunities. The software development community may find other opportunities to use the information in community KGs developed using this method. Evaluation and validation by more experts will improve the StandFram over time.

5 Conclusion

The StandFram method and the application of the CSFKG were introduced in this paper. A conceptual framework is presented for use by other researchers to describe other domains of interest. Each new KG

described using the framework can be shared with other researchers and industry to produce further applications. When a KG is linked to another KG, a larger network of standards and frameworks results. The research reported in this paper presents the StandFram method to build a KG that collocates and improves accessibility to standards and frameworks for users outside the intended communities. This design approach facilitates applications to automate procedures and integrate with existing business recordkeeping systems to support ongoing cybersecurity controls in business, as well as the ability to leverage standardized controls to develop novel approaches to advanced cybersecurity challenges. In the future, the knowledge in the KG can be shared with the software development community, which can then apply generative and predictive machine learning applications and simple form-based applications that a community of software developers can build for smaller businesses. This will result in further empirical validation of the method The StandFram and conceptual framework were used to develop the CSFKG for smaller Australian businesses and offer practitioners a single KG repository that improves accessibility for this sector. The research reported in this paper seeks to identify questions that smaller businesses are asking about cybersecurity. Knowing which questions cannot be answered for smaller Australian businesses may be the impetus to self-organise and develop cybersecurity standards and frameworks that answer their questions about cybersecurity. This paper has presented only one application of the CSFKG developed in this research for smaller businesses. There are many other applications that future researchers may imagine, and application developers implement that may improve the cybersecurity of more than 97.3% of Australia's businesses.

6 References

Australian Bureau of Statistics. 2024, Business Indicators, Australia, ABS Website, accessed 20 October 2024.

Australian Bureau of Statistics. 2016, accessible at https://www.abs.gov.au/AUSSTATS/abs@.nsf/mf/1321.0#:~:text=micro%20businesses%20-%20businesses%20employing%20less%20than%205,or%20more%20people%2C%20but%20less%20than%2020%20people%3B

Alahmari A, and Duncan B., 2020. “Cybersecurity risk management in small and medium-sized enterprises: A systematic review of recent evidence”. In *2020 international conference on cyber situational awareness, data analytics and assessment* (CyberSA) 2020 Jun 15 (pp. 1-5). IEEE.

Almgren, M., Lantz, P., and Lagerström, R. 2019. “The SEPSES Knowledge Graph for Cybersecurity,” *International Journal of Information Security*, (18:3), 323-337.

Azmi, R., Tibben, W., and Win, K. 2018. "Review of cybersecurity frameworks: context and shared concepts," *Journal of Cyber Policy* (3:2), pp. 258-283.

Bada, M., and Nurse, J. R. 2019. "Developing cybersecurity education and awareness programmes for small-and medium-sized enterprises (SMEs)," *Information & Computer Security*, 27(3), pp. 393410.

Bada, M., Sasse, M. A., and Nurse, J. R. C. 2019. “Cyber security awareness campaigns: Why do they fail to change behaviour?,” arXiv preprint arXiv:1901.02672.

Baskerville, R., Spagnoletti, P., and Kim, J. 2014. "Incident-centered information security: Managing a strategic balance between prevention and response," *Information & management*, (51:1), pp. 138151.

Bersagol, V., Dessalles, J.L., Kaplan, F., Marze, J.C. and Picault, S., 1996. “XMOISE: A logical spreadsheet to elicit didactic knowledge.” In *Computer Aided Learning and Instruction in Science and Engineering: Third International Conference*, CALISCE'96 San Sebastian, Spain, July 29–31, 1996 Proceedings (pp. 430-432). Springer Berlin Heidelberg.

Brunsson, N., and Jocobsson, B. 2002. “'Following Standards”, *A World of Standards* (Oxford, 2002; online edition, Oxford Academic, 1 Jan. 2010).

Carias, J.F., Borges., M.R.S., Labaka, L., Arrizabalaga, S., and Harnantes., J. 2020. "Systematic approach to cyber resilience operationalization in SMEs," *IEEE Access*, 8, pp. 174200-174221, DOI: 10.1109/ACCESS.2020.3026063.

Cartwright, A., Cartwright, E., and Edun, E. 2023. "Cascading information on best practice: Cyber security risk management in UK micro and small businesses and the role of IT companies," *Computers and Security*, 103288.

Cartwright, E., Post, G. V., & Sommer, P. 2023. "Towards a Small Business Cybersecurity Knowledge Graph. Journal of Cybersecurity Education," *Research and Practice*, 2023(1), 2.

Chen, X., Shen, W., and Yang, G. 2021. "Automatic generation of attack strategy for multiple vulnerabilities based on domain knowledge graph," *In the proceedings of the IECON 2021–47th Annual Conference of the IEEE Industrial Electronics Society*, pp. 1-6.

Chen, H., Zhu, D., and Zhang, T. 2020. "Change Management in Knowledge Graphs", *Journal of Knowledge Management*, 24(5), 1234-1248.

Chidukwani, A., Zander, S., and Koutsakis, P. 2022. "A Survey on the Cyber Security of Small-to Medium Businesses: Challenges, Research Focus and Recommendations," *IEEE Access* 10, pp. 8570185719. DOI: 10.1109/ACCESS.2022.3197899/.

Cynet, 2022 Survey of CISOs with Small Cyber Security Teams, https://go.cynet.com/2022_ciso_survey (Accessed 1 September 2023)

Department of Home Affairs, 2023. Cyber Security Strategy 2023-2030. https://www.homeaffairs.gov.au/about-us/our-portfolios/cyber-security/strategy/20232030australian-cyber-security-strategy

Dai, Y., Wang, S., Xiong, N. N., & Guo, W. 2020. "A survey on knowledge graph embedding: Approaches, applications and benchmarks," Electronics, (9:5), 750.

Dresch, A., Lacerda, D. P., & Antunes, J. A. V. 2015. "Proposal for the conduct of design science research," *Design Science Research: A Method for Science and Technology Advancement*, 117-127.

Gregor, S., & Hevner, A. R. 2014. "The Knowledge Innovation Matrix (KIM): A clarifying lens for innovation.*," Informing Science: the International Journal of an Emerging Transdiscipline,* 17, 217239.

Gregor, S. and Jones, D., 2007. "The anatomy of a design theory," *Journal of the Association for Information System*, (8:5), 312-335.

Grover, V., and Lyytinen, K. 2015. "New state of play in information systems research: The push to the edges," *MIS Quarterly*, (39:2), 271-296.

Han, L., Finin, T., Parr, C., Sachs, J., and Joshi., A. 2008. "RDF123 from Spreadsheets to RDF," *In the proceedings of the international workshop on the Sematic Web, Lecture Notes in Computer Science.*

Hevner, A. R., March, S.T., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *MIS Quarterly* (28:1), pp.75-105.

Humayun, M., Niazi, M., Jhanjhi, N. Z., Alshayeb, M., and Mahmood, S. 2020. "Cyber security threats and vulnerabilities: a systematic mapping study," *Arabian Journal for Science and Engineering* 45, pp. 3171-3189.

ISO/IEC 27001:2022 "Information security, cybersecurity and privacy protection," *Information security management systems Requirements* (2024) Retrieved April 2024 from International Standards Organization (2022), https://www.iso.org/standard/27001

Iivari J. 2020. "A critical look at theories in design science research," *Journal of the Association for Information Systems.* 2020;21(3):10.

Jia, Y., Qi, Y., Shang, H., Jiang, R., and Li, A. 2018. "A practical approach to constructing a knowledge graph for cybersecurity," *Engineering* (4:1), pp. 53-60.

Jia, Y., Xiang, G., & Li, L. 2018. “Constructing a cybersecurity knowledge graph for threat detection,” *Proceedings of the 27th International Conference on World Wide Web Companion*, 1791-1797.

Kappelman, L., Torres, R., McLean, E., Maurer, C., Johnson, V., and Kim, K. 2019. "The 2018 SIM IT issues and trends study," *MIS Quarterly Executive* (18:1), pp. 51-84.

Kelley, G. (Ed.). 2008. "Selected Readings on Information Technology Management: Contemporary Issues," IGI Global.

Kerwer, D., 2005. "Rules that many use: standards and global regulation," *Governance*, (18:4), pp.611632.

Kiesling, E., Ekelhart, A., Kurniawan, K., and Ekaputra, F. 2019. "The SEPSES knowledge graph: an integrated resource for cybersecurity," In *the proceedings of the International Semantic Web Conference, Cham: Springer International Publishing*, pp. 198-214.

Kong, Y., Liu, X., Zhao, Z. 2022. "Bolt defect classification algorithm based on knowledge graph and feature fusion," *Energy Rep* (8), pp. 856–863

Kurniawan, K., Ekelhart, A., Kiesling, E., and Quirchmayr, G. 2022. "KRYSTAL: Knowledge graph-based framework for tactical attack discovery in audit data," *Computers and Security* (121), pp. 102828.

Mayer, R., Lünendonk, A., Jeschke, S., & Sadeghi, A. R. 2017. “Security and Privacy in Cyber-Physical Systems: Foundations, Principles, and Applications”. John Wiley & Sons.

Nickel, M., Murphy, K., Tresp, V., & Gabrilovich, E. 2016. “A review of relational machine learning for knowledge graphs,” *Proceedings of the IEEE*, 104(1), 11-33.

Pan, J. Z., Stamou, G. B., Taylor, S., & Horrocks, I. 2009. “Semantic Web Technologies: Trends and Research in Ontology-based Systems,” John Wiley & Sons.

Paulheim, H. 2017. “Knowledge Graph Refinement: A Survey of Approaches and Evaluation Methods,” *Semantic Web*, 8(3), 489-508.

Piccarozzi, M., Stefanoni, A., Silvestri, C., and Ioppolo, G. 2023. "Industry 4.0 technologies as a lever for sustainability in the communication of large companies to stakeholders," *European Journal of Innovation Management*.

Pries-Heje, J., Baskerville, R., and Venable, J.R. 2008. "Strategies for Design Science Evaluation," in *proceedings of the European Conference on Information Systems*.

Pujara, J., Miao, H., Getoor, L., and Cohen, W. 2013a. "Knowledge graph identification," In the *proceedings of the International semantic web conference*, pp. 542–557.

Queensland Government. 2023 accessible at https://www.dsdsatsip.qld.gov.au/resources/dsdsatsip/work/atsip/business-economic-development/qipp/indigenous-business-procurement-guide.pdf

Romano, R., and John, B. 2024a. “Cybersecurity Education – Answering the Questions that Smaller Businesses are Asking”, Association for Information Systems. *Proceedings of the 2024 AIS SIGED European Conference on Information Systems Research,* ECISER 2024.

Romano, R., and John, B. 2024b. “A Cybersecurity Standards and Frameworks Knowledge Graph for the Education of Sustainable Australian Smaller Businesses”, *Thirty-Second European Conference on Information Systems* (ECIS 2024, 2024, Paphos, Cyprus. Avital, M. Karahanna, E., Themistocleous, M., Constantiou, I., Fitzgerald, B. & Seidel, S. (eds.). Association for Information Systems, p. 1-3 3p.

Sarrafzadeh, B., Vechtomova, O. and Jokic, V. 2014. "Exploring knowledge graphs for exploratory search," in *Proceedings of the 5th Information Interaction in Context Symposium*, pp. 135-144.

Saunders, M. N. K. 2011. "Content Analysis In: The SAGE Dictionary of Qualitative Management Research," in R. Thorpe, and R. Holt (eds.), *The sage dictionary of qualitative management research,* SAGE Publications Ltd, London.

Seeburn, K. 2014. “Basic Foundational Concepts Student Book: Using COBIT 5,” ISACA: Schaumburg, IL, USA, 2014.

Segal, E. 2022. “Why Small and Medium-Sized Companies Face More Cyber Challenges Thank Large Ones: Survey”. Forbes, www.forbes.com/sites/edwardsegal (13 July 2022, 12:56 PM).

Sikos, L. F. 2023. "Cybersecurity knowledge graphs," *Knowledge and Information Systems* (65:9), pp. 3511-3531.

Syed, Z., Padia, A., Finin, T., Mathews, L., and Joshi, A., 2016. "UCO: A Unified Cybersecurity Ontology." in the *Workshops of the Thirteenth AAAI Conference on Artificial Intelligence for Cyber Security: Technical Report* WS-16-03.

Taherdoost, H. 2022. "Understanding Cybersecurity Standards and Information Security Standards – A Review and Comprehensive Overview," *Electronics* (Switzerland) (11: 14).

Tam, T., Rao, A., and Hall, J. 2021. "The good, the bad and the missing: A Narrative review of cybersecurity implications for Australian small businesses," *Computers & Security* (109), pp. 102385.

Venable J. 2006. “The role of theory and theorising in design science research,” *In proceedings of the 1st international conference on design science in information systems and technology* (DESRIST 2006) Feb 24 (pp. 1-18).

W3C 1999, "Resource Description Framework (RDF) Model and Syntax Specification". 22 Feb 1999.

W3C 2013, "PROV-O: The PROV Ontology". 30 Apr 2013.

W3C 2022, “Data Privacy Vocabulary (DPV) version 2.0,” Data Privacy Vocabulary (DPV) (w3c.github.io) accessed 16 July 2024.

W3c 2024, “RDF 1.2 Concepts and Abstract Syntax,” W3C Working Draft 04 July 2024. https://www.w3.org/TR/rdf12-concepts/#resources-and-statements accessed 16 July 2024.

Wang, Z., Zhu, H., Liu, P., and Sun, L. 2021. "Social engineering in cybersecurity: a domain ontology and knowledge graph application examples," *Cybersecurity* (4), pp. 1-21.

Zhu, D., Chen, H., & Zhang, T. 2023. “The Measurement of Knowledge in Knowledge Graphs,” *Journal of Artificial Intelligence Research,* 60(4), 123-145.

Acknowledgements

The researchers appreciate the funding and in-kind contributions received from Surround Australia, as well as the funding from Pathfinders, and in-kind support from Procure Spot which made this research possible.

Copyright