UNIVERSITÉ DU QUÉBEC EN OUTAOUAIS

IMPLEMENTING INFORMATION TECHNOLOGY MANAGEMENT WITH AN ONTOLOGY OF CYBERSECURITY PROFESSIONAL SKILLS

DOCTORAL THESIS PRESENTED TO

DÉPARTEMENT D'INFORMATIQUE ET D'INGÉNIERIE

IN PARTIAL FULFILLEMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

MARC-ANDRÉ LÉGER

GATINEAU, 15 DECEMBER 2021

ABSTRACT

Standards of professional cybersecurity skills are multifarious, complex, and difficult to integrate and exploit for talent management in organizations. This study proposes using an ontology for facilitating the integration of skill repositories, proprietary standards, and open standards. Using an action design research methodology, this study aims to develop and test an innovative ontology of cybersecurity professional skills for talent management of large organizations, specifically in the financial services industry in Canada. The open collaborative development lifecycle of this study involves a community of experts.

SOMMAIRE

Les normes de compétences professionnelles en cybersécurité sont nombreuses et complexes. Elles sont difficiles à intégrer et à exploiter pour la gestion des talents dans les organisations. Nous proposons d'utiliser une ontologie afin de faciliter l'intégration de référentiels de compétences, des normes propriétaires et ouvertes. En utilisant une méthodologie de recherche Action Design, nous proposons de développer et de tester une ontologie innovante des compétences professionnelles en cybersécurité pour la gestion des talents de grandes organisations, plus précisément dans l'industrie des services financiers au Canada. Notre cycle de vie de développement ouvert collaboratif implique une communauté d'experts.

BOARD OF EXAMINERS

THIS THESIS HAS BEEN EVALUATED BY THE FOLLOWING BOARD OF EXAMINERS:

Pre Anna Margulis, UQO, Chair, anna.margulis@uqo.ca, https://uqo.ca/profil/margan01

Pre Romilla Syed, University of Massachusetts, Boston, External Evaluator, Romilla.Syed@umb.edu, https://www.umb.edu/faculty_staff/bio/romilla_syed

Pr Péricles Sobreira, UQO, Internal Evaluator, pericles.sobreira@gmail.com
http://w4.uqo.ca/dii/dyn/profs/pericles.delimasobreira.php

Pr Raul Valverde, UQO, Thesis Codirector, raul.valverde@concordia.ca
https://www.concordia.ca/jmsb/faculty/raul-valverde.html

Pr Stéphane Gagnon, UQO, Thesis Director, stephane.gagnon@uqo.ca
https://gagnontech.org/

Table of contents

Table of acronyms

Acronym	Definition
ADR	Active design research
BIE	Building, intervention, and evaluation
CIA	Confidentiality, integrity, and availability
CCM	Cybersecurity competency model
DSR	Design science research
DSS	Decision support system
EU	European Union
FinTech	Financial technologies
GDPR	General Data Protection Regulation
HRM	Human resource management
IRI	Internationalized Resource Identifiers
IS	Information system
IT	Information technology
KSA	Knowledge, skills, and abilities
MCC	Matthews correlation coefficient
NCWF	National Cybersecurity Workforce Framework
NICE	National Initiative for Cybersecurity Education
NIST	National Institute of Science and Technology
OWL	Ontology Web Language
PCI-DSS	Payment Card Industry Data Security Standard
RDF	Resource Description Format
SPARQL	Semantic Query Language for data stored in RDF format
SWRL	Semantic Web Rule Language
UML	Unified Modeling Language
UQO	Université du Québec en Outaouais
URI	Uniform Resource Identifier
USDOLETA	United States Department of Labor, Employment and Training Administration

Table of figures

1 Introduction

Financial institutions require information to provide financial products and services to their customers. This makes information technology (IT) their key strategic tool, and these institutions must manage the security of the information while using the same. This study aims to contribute to improving how financial organizations manage information security.

In today's complex world, financial institutions are confronted with a wide array of challenges, as society is transitioning from financial transactions based on physical currency to a purely electronic financial world. New financial technologies, fintech, are emerging as a driving force of innovation in this industry benefiting from continuous advances in high-bandwidth networking, mobility, software development, and artificial intelligence. Business technologies combining information systems, IT, and fintech provide strategic tools for organizations. With business technologies, the financial sector, as it exists today, is very different from what it was 10 years ago and is anticipated to be extremely distinct in another 10 years.

To profit from this innovation, not only the financial industry but also the society requires multiple levels of trust between the actors to exist to maintain the integrity of the economic system. A vital component of creating trust is risk management. By determining and managing unacceptable risks, financial institutions can create a safe marketplace for financial institutions, their customers, and industry stakeholders. There are various risk categories that concern the financial industry (Chornous & Ursulenko, 2013; Hunton, Wright, & Wright, 2004; Isaca, 2009), such as market risk, credit risk, operational risk, liquidity risk, network and database security risks, and overall internal control risk (Ochuko, 2013). This study investigates one of these categories, the reduction of risks that are caused by cybersecurity competency gaps by introducing the use of a cybersecurity competency ontology. Insufficient or deficient cybersecurity competencies give rise to risks related to the use of business technologies in organizations.

The financial sector deals with money and how it circulates in the economy. Hence, in addition to being a critical issue for institutions themselves, cybersecurity is of prime

importance to the world economy and the national interest. Many people get so caught up in making money, and some are even willing to adopt a felonious conduct to acquire the same. Financial institutions must adequately protect money so that it cannot be misappropriated. Given that money is the data circulating on cyber networks, protecting the monetary system and the flow of money necessitates cybersecurity. Financial institutions are primary targets of cybercriminals, as they offer a multitude and concentration of financial services, which render them an attractive target for cybercriminals.

In response to the rapidly growing threats, regulatory requirements to protect financial sector organizations and their customers have risen up at both national and international levels (Leung, 2018). Regulatory, legal, and contractual obligations, in addition to the obligation to conform to norms, standards, and international treaties on information security, put additional pressure on financial organizations. These include regulatory directives such as the US Payment Service Directive 2, statutes such as the Sarbanes–Oxley Act (USC, 2002), and standards such as the Payment Card Industry Data Security Standard (PCI Council, 2021), and privacy regulations such as the European Union (EU) General Data Protection Regulation ("GDPR," 2021) and various other country-specific or international laws and regulations (Benaroch, Chernobai, & Goldstein, 2012). The presence of this multitude of different regulations and requirements makes information security compliance a complex task.

Cybersecurity is also required, as managers have a fiduciary duty to act diligently in the best interest of their organizations. This brings to forefront issues, such as ethics and governance. Financial institutions must not only embrace measures that ensure maximum information security but also abide by the numerous laws and regulations governing the security and privacy of data (Ula et al., 2011). Nowadays, good governance entails risk management. More specifically, good IT governance requires cybersecurity.

However, there are many problems that make it difficult for organizations to fulfill the expectations of their customers. A few problems are enumerated here. The complex nature of interconnected global networks, such as the Internet of telecommunication

networks, creates many access points that can become sources of vulnerabilities (Elahi, Yu, & Zannone, 2010; MITRE, 2021; Partida & Andina, 2010a). Moreover, these networks are highly heterogeneous and evolve so quickly that new vulnerabilities are continuously being discovered (Allodi & Massacci, 2017; "CVE," 2020; Partida & Andina, 2010b; Taubenberger, 2014). At the same time, human, technical, and systemic threats continue to emerge, proliferate, and evolve (Allodi & Massacci, 2017; Armstrong, Jones, Namin, & Newton, 2018; Bevilacqua & Ciarapica, 2018).

There are a number of opportunities to explore in the search for solutions to help financial organizations. Several technical, human, procedural, or strategic avenues offer interesting challenges for research and innovation, which are too many to explore in a single study. Because of the researchers' interest and of opportunities that have arisen, this study is centered on what was regarded as the most critical strategic cybersecurity issue currently faced by organizations and financial institutions in Canada, namely the shortage of competent cybersecurity human resources. The findings of a recent survey revealed the global shortage of workers at 3.12 million in 2020, with 56% of respondents claiming that their organizations are at risk due to the shortage (PricewaterhouseCoopers, 2020). Other studies estimated a global shortage of 4.07 million cybersecurity staff in 2019, a 26% increase from 2018 (ISC2, 2020a) and 512,000 unfilled cybersecurity positions in the USA in 2021 (CyberSeek, 2021).

Given the combination of social, cultural, and historical reasons, there has been a reduction in the rate of the growth of the labor force in Canada (Government of Canada, 2019). This has become more acute in the past few years and should reach a critical stage very soon as the last of the Baby Boomer generation reaches retirement age. This problem is compounded when national economies are rated on the basis of continued economic growth. Furthermore, stakeholders and markets generally expect increasing financial returns and sustained growth. While an increasing number of questions arise as to the sustainable nature of limitless continued growth in a finite ecosystem, it is still the current reality that organizations face. The general shortage of workers becomes more acute when specialists are needed. The more specialized, the worse the problem. Thus, as

financial organizations must find ways to improve their effectiveness and efficiency with a shortage of IT experts, it is even more problematic with cybersecurity experts.

At present, there is ample pertinent evidence of the already conspicuous shortage of talented cybersecurity professionals who possess the requisite skills. This shortage is forecasted to grow that is being felt in the field (Furnell & Bishop, 2020; Herjavec, 2020):

- Thirty percent of organizations were struggling with a skill shortage (van Kessel, 2018).
- Fifty-three percent of the IT professionals surveyed by the Enterprise Strategy Group considered the problematic shortage of cybersecurity skills as their foremost issue (Oltsik, 2019).
- Sixty-five percent of the organizations surveyed by the ISC2 reported a cybersecurity skill shortage, with 51% of them considering that their organization was at moderate or extreme risk as a result (ISC2, 2020b).

To compound the problem, there is enormous confusion between talent and competency (Draganidis & Mentzas, 2006; Man et al., 2002; Subramaniam et al., 2019). Who is a competent professional? A competent professional has the capability and directed intent of an individual within the requirements that a job demands in the context of a particular organization (Boyatzis, 2008). Organizations do not just need resource but also talented, competent individuals who can help them maintain and improve cybersecurity. A competent individual for a position must have more than the skills or degrees required, and the individual must be capable of adequately doing the required job. As this is a complex matter, research such as the current study is crucial to propose solutions.

This is a critical issue as organizations need solutions. Plausible avenues include better identification of competency, identification of individuals with nontraditional backgrounds who can be trained or cross-trained, identification of students who exhibit potential for competency, and identification of new immigrants who have difficulty integrating into the Canadian job market and others. There are numerous possibilities to identify these nontraditional individuals or individuals with transferable skills who could

become cybersecurity workers. Notwithstanding, organizations must be able to scale up the current, mostly manual, recruitment process to do this efficiently. Tools and methodologies are required to industrialize how this process takes place to make it more successful and efficient. This is where this study aims to provide solutions for the financial industry in Canada.

1.1 Purpose of this study

This study aims to improve information system-related risk management activities in financial institutions, chiefly in Canada, as this is the available geographic area for research. As there is a vast scope for improvement in information security, the contribution of this study materialized with a better alignment between the competencies of individual actors in the cybersecurity work roles of the financial institutions and the business requirements for information security that the organizations have. Misalignment of competencies and work roles increases risks, which can result in additional expenses, damages, losses, or negative financial impacts. The proposed ontology contributes to the said alignment by providing strong bases for a unified framework that managers with information security responsibilities in strategic business units, information security departments, and human management capacities can use for producing reliable performance indicators about the organizational fit of competencies and roles with the needs of the organizations.

1.2 Significance of this study

Much has been written about cybersecurity as a business issue and from a technical standpoint. This study is focused on another aspect related to the cybersecurity work competencies required to operationalize technical and business solutions. The growing talent shortage that was mentioned previously is also a popular subject in mainstream media. There are scientific journal publications on cybersecurity curriculum design and on the cybersecurity workforce, but there appears to be a gap in how competencies and talent are linked. In addition, there is a dire need to provide financial organizations with tools, which are validated by scientific research, to help them become more efficient in this area. However, to get to a point where tools can be developed and provided,

knowledge and skills gaps need to be understood and subsequently bridged. This is where this study can have a considerable impact, bridging theses gaps and laying the foundations for financial organizations to mitigate the risks associated with cybersecurity competencies and talent with tools that rely on emerging technologies, such as machine learning. The plausible dearth of similar studies (ontology based) for other cybersecurity or IT professions, further substantiates the need for this study. More specifically, the ontology that was developed and which is presented in this dissertation can be used in a multiplicity of ways to assist managers and organizations. For example, as shown in the validation tests, the proposed ontology can be used for processing large volumes of documents to identify the most likely matches. This could allow an organization to extract the best candidates to fill available cybersecurity positions from large databases of candidates or by using large external data sources, such as LinkedIn or other sources that can be purchased. Another potential application could be to scan existing employees' databases to identify potential candidates for hard-to-fill specialized positions or identify who could take the position with some basic training to fill gaps in competencies. Another use that we have discovered during this study is conformity. In fact, the data collected, as along with the formal process of competency management put in place, can be used to answer auditor questions during conformity audits. As such, the ontology and the queries that were developed and are presented in this dissertation are the first bricks toward building a solution.

1.3 Organizations of Sections

Section 2 presents the research question for this study. This is followed by the literature review, in Section 3, which includes the definition of the primary concepts that are covered in this dissertation. Section 4 presents the research methodology and the steps that were used in the execution of this study. The participants are described, in Section 5, the data collection strategies in Section 6, the validation in Section 7, and the research calendar in Section 8. Section 9 delineates the ontology design approach that was developed as a part of this study. The development of the ontology is then presented in Section 10, which describes how the cybersecurity competency ontology was built. Section 11 addresses the internal validity, or how the ontology was designed and then

represents the reality in the organization where this study was conducted. Sections 12 and 13 discuss issues associated with external validity, or how the ontology can describe the reality of other Canadian financial institutions as it regards cybersecurity competencies. Finally, Section 14 presents a discussion, followed by a conclusion and a bibliography.

2 Research Questions

This study was conducted as one of the requirements for a Ph.D. in Sciences and Information Technology program at Université du Québec en Outaouais (UQO). The designed ontology presents the conditions for decision-makers to have a better understanding of the potential gaps and, most importantly, the vulnerabilities caused by the gaps in cybersecurity competencies. Ultimately, cybersecurity competency gaps can result in vulnerabilities that negatively impact the confidentiality, integrity, and availability of information. Implementing this ontology, increasing the awareness of the competency gaps, and helping managers fill the same contribute to reducing unacceptable risks. This brings us to this study's principal research question:

> **RQ-1: Can a cybersecurity competency ontology provide an effective tool for financial institutions to manage cybersecurity talents?**

This research question is structured into three sub-questions, identified as RQ-1.1, 1.2, and 1.3. The answers to these questions are examined in further details in this dissertation in Sections 9 through 13. Section 9 presents how the ontology was built (RQ-1.1); Section 10 exhibits how it was mapped to the subject domain (RQ-1.2); lastly, Sections 11, 12, and 13 explain how it was formally validated and tested using semantic queries and rule-based inferences given cybersecurity scenarios (RQ-1.3).

The first sub-question addresses the ontology design approach:

> **RQ-1.1: What is the most effective approach for developing a new cybersecurity ontology that represents the competencies, skills, and abilities of effective practices in this field?**

Answering this sub-question led to elaborating and justifying how ontologies and semantic reasoning strategies can be employed to develop, validate, and test the innovative ontology for cybersecurity talent management.

The second sub-question utilizes the results of the first one and expands from there:

RQ-1.2: What should be the structure and contents of an Ontology Web Language (OWL) ontology representing the core competencies of the cybersecurity domain?

Following the approach presented in Section 9, Section 10 presents the new ontology results using OWL in the application Stanford Protégé. This artifact, which uses data coming from the fieldwork gathered by applying the research methodology, is based on the cybersecurity requirements of a large Canadian financial institution, aligned with the National Cybersecurity Workforce Framework (NCWF) (NIST, 2021).

Answering this second sub-question essentially led us to mapping the cybersecurity ontology with cybersecurity competency reference models and the requirements of financial institutions to give a practical use of the ontology for human resource management (HRM).

Once the ontology was completed, it was validated and tested for the usability of the cybersecurity competency in the specific context of the target organization, answering the following question:

RQ-1.3: What is the level of validity of the ontology in accurately representing the cybersecurity domain, and to what extent is it effective as a talent management decision tool?

This third sub-question explored the opportunity of using the ontology to help a large Canadian financial institution in the management of cybersecurity talents. Sections 11, 12, and 13 present how it was validated and tested as a management tool that financial institutions can utilize to find, identify, and retain cybersecurity talents that are well targeted to the needs of the organizations, given specific cybersecurity scenarios and requirements.

3 Literature Review

To adequately frame the main concepts of this exploratory study, a thorough literature review was conducted. This study encompasses several areas of knowledge, chiefly cybersecurity and cybersecurity competencies, the current state of knowledge of these concepts had to be identified to identify the existing gaps. As research does not take place in a vacuum, this study started with materials from previously conducted research in risk management that informed the current study and formed a concrete academic basis for the researchers interested in this domain (Léger, 2001, 2003).

A literature review was conducted using academic databases available through the library system research tools of the UQO, online academic tools, such as Google Scholar and other credible Internet sources. The search for new material was focused on articles published after 2010 on the topics such as cybersecurity, competency, cybersecurity competency, and related concepts that were identified at the onset of this study. For example, cybersecurity has various related concepts, such as information security, computer security, and cybercrime, which were included in the search. Appropriate and pertinent non-scholarly resources were also utilized for information about national frameworks, standards, certifications, and educational programs for example. The research also included methodological aspects and the chosen research methodology. The articles were sorted, analyzed, and perused once their relevance was ascertained. Furthermore, various tools were used in this study. Zotero (https://www.zotero.org/) was used to help collect, organize, cite, and manage the references and bibliography. MindManager (https://www.mindmanager.com/) helped manage the research themes into conceptual maps that could help structure the researcher's initial understanding of the subject. These tools were very useful for becoming organized from the beginning of this study and structuring the collected data. As this study evolved, these tools were also used during the analysis to help manage the references and have more structure in the data collection. A conceptual map, presented in Appendix J, was created using MindManager and maintained over the duration of the study to assist in creating and maintaining a thorough understanding of the research subject.

In this section of the dissertation, the current state of knowledge in main concepts of the study is presented. This section starts with cybersecurity, the principal problem domain, with a focus on risk and how it is managed in financial institutions. Dynamic capabilities theory is then explained in detail. From there, competency is defined and then presented in the context of cybersecurity. Finally, the concept of ontology, which is being developed as one of the results of the study, is defined.

3.1 Cybersecurity

Organizations require information security because they need to protect the information they use as a strategic asset and technology by nature creates risks (Bahli & Rivard, 2003; Beucher, Veyret, & Reghezza, 2004; Crichton, 2002; Léger, 2001, 2003). Organizations wish to take acceptable risks and foster a security culture as they endeavor to maintain an equilibrium between operational performance and the costs of mitigating unacceptable risks (Ahmed & Abraham, 2013; Alter & Sherer, 2004; Bahli & Rivard, 2003; Bannerman, 2008; Cooper, 2000; Cox S & Flin R, 1998; De Haes & Van Grembergen, 2009; Douglas & Wildavsky, 1983; Furnell & Bishop, 2020; Guldenmund, 2000; Hiller & Russell, 2013; Pidgeon, 1998; Richter & Koch, 2004). To achieve this equilibrium, they develop guidelines, policies, and safeguards (Bonollo & Massimiliano, 2012; Camillo, 2017; Elshahat, Parhizgari, & Hong, 2012; Hiller & Russell, 2013; Joshi et al.,2013; Leung, 2018; Ochuko, 2013; Ula et al., 2011). Keeping information systems free from unacceptable risks to maintain information security is at the heart of cybersecurity.

Cybersecurity is a security risk management process followed by organizations to protect the confidentiality, integrity, and availability of data and access and assets that are utilized in cyberspace (Edgar & Manz, 2017; Schatz, Bashroush, & Wall, 2017). Building on the definitions of security and information security, cybersecurity can be perceived as information security in an interconnected world (Alter & Sherer, 2004; Banham, 2017; Callen-Naviglia & James, 2018; Cleveland & Spangler, 2018; Cleveland & Cleveland, 2018; Hiller & Russell, 2013; Schatz et al., 2017; van Kessel, 2018).

3.2 Risk management

As diligent managers are required to manage cyber risks, they assess cybersecurity risk in a continuous process of identification and prioritization to determine appropriate countermeasures that could mitigate unacceptable risks (Dawson, Crespo, & Brewster, 2013; Elahi et al., 2010; Hiller & Russell, 2013; Léger, 2001, 2003; Maisey, 2014; Partida & Andina, 2010b; Yoe, 2011). Mitigation measures must be carefully selected, as they have a financial and human cost and may impact the ability to fulfill business goals (Abawajy, 2014; Agrawal, Finnie, & Krishnan, 2010; Ioannidis, Pym, & Williams, 2012). One approach to assist organizations in accomplishing this goal is the use of scenarios to avoid assuming that the future will look like the present (Bradfield, Wright, Burt, Cairns, & Van Der Heijden, 2005; Ergashev, 2012; Mcube, 2017; Rippel & Teply, 2010; Thomsen, Sørensen, Fauser, & Porragas, 2006; Wilkinson & Kupers, 2013). This strategic foresight confers several benefits, such as allowing organizations to consider inconceivable or imperceptible futures, increasing their ability to perceive change, and helping them interpret and respond to change and their capacity for organizational learning (Gilbert, 2000; Marcelo, Rodríguez, & Trucharte, 2008; Rigby & Bilodeau, 2007; Rohrbeck & Schwarz, 2013).

There exists a correlation between cybersecurity and other information security strategies in organizations (Jirasek, 2012). Nonetheless, cybersecurity strategies support broader organizational goals, ideally integrated into enterprise risk management strategies (Babb, 2014; Bakshi, 2012; De Haes, Van Grembergen, & Debreceny, 2013a; Frelinger, 2012). Once cybersecurity strategies are defined, technical, procedural, and internal controls are implemented using organizational policies, standards, optimal practices, and frameworks (Amarachi, Okolie, & Ajaegbu, 2013; Asosheh, Hajinazari, & Khodkari, 2013; De Haes, Van Grembergen, & Debreceny, 2013b; Disterer, 2013, 2013; Fenz, Goluch, Ekelhar, Riedl, & Weippl, 2007a; Humphreys, 2006; Leitner & Schaumuller-Bichl, 2009). Controls include media protection, risk assessment, contingency planning, and configuration management, to name only a few (Sheikhpour & Modiri, 2012; Twum & Ahenkora, 2012). Conversely, implementing technical controls is the responsibility of competent cybersecurity professionals, and their overall target is to attain a balance

business functionality and security (Asosheh et al., 2013; Galliano, 2017; Jirasek, 2012; Scarfone, Jansen, & Tracy, 2008).

3.3 Dynamic capabilities

An organization's ability to create and sustain a competitive advantage depends on multiple factors, which include strategic tools and business processes (Fernandes et al., 2017; Jarzabkowski & Paul Spee, 2009; Teece, 2018). Decision-makers in an organization need to identify the strategic tools that are most likely to help them cope with the uncertainties of their business environment (Iszatt-White, 2010; Jarzabkowski & Kaplan, 2015). Developing dynamic capabilities is one of the strategic tools (Sanchez, 2004; Teece, 2007; Teece, Pisano, & Shuen, 1997; Westley & Mintzberg, 1989). Accelerating changes mean that once a competitive advantage is created, organizations need to continuously evaluate, select, and implement emerging technological innovation to maintain the same (Beck & Wiersema, 2013). This is where dynamic capabilities can be useful. In other words, dynamic capabilities are about creating a culture where a dynamic adaptation of its resources, both human and material, is what an organization can do successfully and continuously to create and maintain a competitive advantage (Teece, 2007; Teece et al., 1997).

Dynamic capabilities can be further defined as the plans and business processes that organizations devise to allow them to transform and evolve more readily (Beck & Wiersema, 2013). What sets dynamic capabilities apart from other strategies is the intent to create an agile environment that could foster adaptability and a strong long-term competitive advantage (Mirabeau & Maguire, 2014; Pfeffer & Sutton, 2006). It sets to develop an inherent capability that allows an organization to use all of its resources purposefully and optimally (Eisenhardt & Martin, 2000). Findings of different studies have revealed that dynamic capabilities model is a tool used by organizations to adapt to emerging threats and identify new opportunities (Beck & Wiersema, 2013). Dynamic capabilities can help an organization ensure it has adequate coverage and is prepared for diversified situations (Ericson, 2014) and deal with change (Lê & Jarzabkowski, 2015). It is vital for sensing change, seizing opportunities, and reconfiguring the organization to

adapt to a new situation (Helfat & Peteraf, 2015; Fernades et al., 2017; Beck & Wiersema, 2013). Dynamic capabilities leading toward a form of organizational agility are a key strategic tool for survival in the current corporate landscape. As Porter (1996) explained, not only businesses need to be operationally effective, but managers must also plan for unforeseen threats and opportunities (Nag, Hambrick, & Chen, 2007). In this regard, dynamic capabilities can contribute to effective cybersecurity by creating an organizational capability to rapidly adapt to unforeseen and emerging threats and vulnerabilities to an essential component of modern organizations – its information technologies.

3.4 Competency

Central to this study is the concept of competency, which can be defined as a characteristic of a successful performer in a work role (Boyatzis, 2008; Prescott, 2012; Subramaniam et al., 2019). Competency can be demonstrated through behavior and actions (Man, Lau, & Chan, 2002b). It is generally regarded as more vital for the success or failure of an individual in a work role than formal education (Draksler & Širec, 2018a). Individuals are not born with competencies but acquire and develop them over time. Competency is, therefore, variable and learnable, which allows intervention in terms of choice and teaching (Draksler & Širec, 2018a; Man et al., 2002b; McClelland, 1973; Mitchelmore & Rowley, 2010). It can further be defined in relation to know-how, know-what, and know-how-to-be but also in some models in relation to knowledge, skills, and abilities. As the literature shows, these are closely related and, in some instances exactly the same.

Know-how is knowledge – it knows how to do something. On its own, it does not guarantee to accomplish a task successfully, which lies in the definition of competency; it is rather a step in the direction of competency. Know-how can be acquired through formal and informal education as a blend of tacit and explicit knowledge (Draganidis & Mentzas, 2006). It requires not only high-level problem-solving abilities applied while adhering to best practices and recognized standards but also technical and business knowledge. Additionally, at high levels, individuals have the ability to combine elements

of know-how to propose alternative or innovative solutions (Bacigalupo, Kampylis, McCallum, & Punie, 2016; Man et al., 2002b). Formal education is a source of know-how, which contributes to competency, but the ability to perform successfully in a role extends beyond theory, as know-what, the ability to successfully apply it in a real-world scenario is of paramount importance (Bacigalupo et al., 2016; Boyatzis, 2008, 2008; Man et al., 2002a).

Know-what refers to the ability of competent individuals to skillfully demonstrate practical knowledge of the work, tasks, methods, business, and ecosystem of an organization applying know-how. It is linked to skills and abilities. When individuals have know-what, they also have in-depth mastery of how their domain of competency operates as a coherent system. Likewise, know-how and know-what, effective performance and career efficiency can also be linked to the emotional, social, and cognitive intelligence of individuals, i.e., know-how-to-be (Bacigalupo et al., 2016; Boyatzis, 2008, 2008; Man et al., 2002a).

Know-how-to-be is a characteristic of competent individuals who demonstrate high emotional and human relations abilities, mental and physical capacities, basic sense attitudes, strong value systems, and behaviors compatible with the organization's culture and the dominant socio-cultural values of various internal and external stakeholders, including the ability to interact with colleagues (Bacigalupo et al., 2016; Boyatzis, 2008, 2008; Draksler & Širec, 2018b; Man et al., 2002a; McClelland, 1973; Mitchelmore & Rowley, 2010). Figure 1 presents a Unified Modeling Language (UML) model that we developed to represent how different competency elements are related in the construction of competency. This UML model was constructed to help the researchers better conceptualize how the elements of competencies are connected. This is useful not only while presenting the concept of competency but also for assisting in early designs of the ontology.

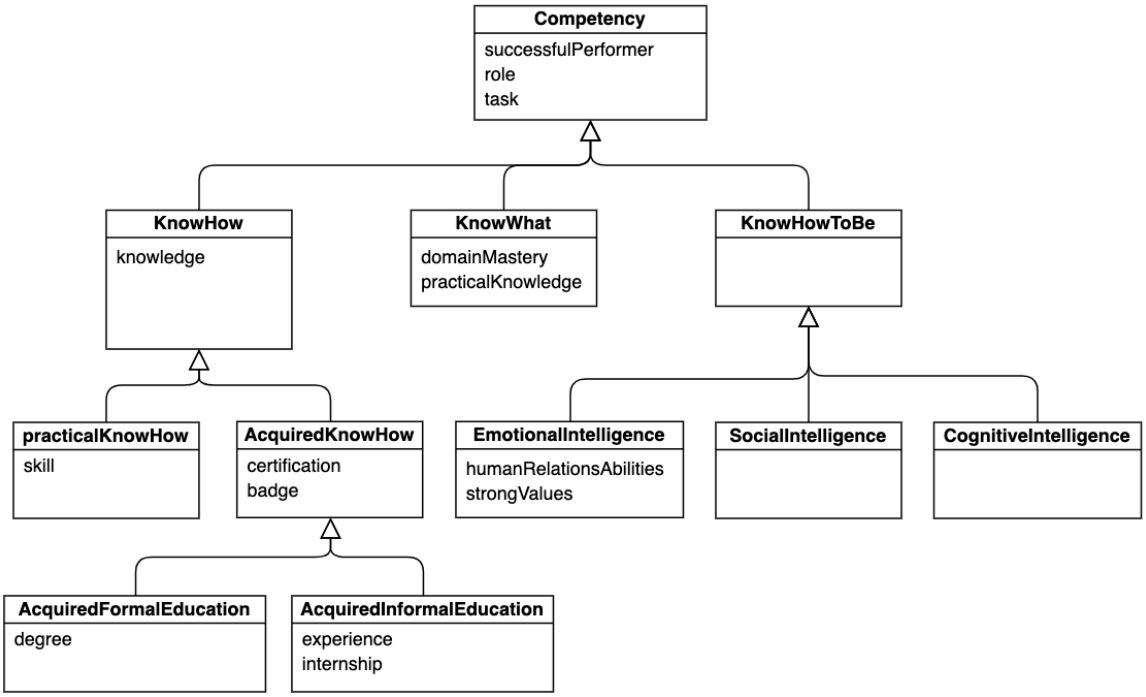


Figure 1: UML Knowledge Model of Competency

An important aspect of competency in the context of this study is how competency can be measured and appraised. Even if this is not part of this study itself, the organizations that will use the results of this study and further research projects will be able to complement this study by developing measurement tools. Having a metric to quantify competency is an essential aspect of evaluating the competency levels of individual actors in work roles. Levels of competency and the three components of competency can be evaluated in relation to the six levels of Bloom's revised taxonomy (Krathwohl, 2002).

1. **Remember:** At this first level, individuals exhibit a memory of previously learned materials. At this level, there is no competency.
2. **Understand:** This is the minimum level at which an embryonic competency can be considered. At this level, individuals demonstrate an understanding of facts and ideas. They begin to develop know-how but not necessarily know-what.
3. **Apply:** When individuals can successfully use their know-how to solve problems and apply the same – know-what – competencies and abilities are demonstrated to find and apply solutions to problems in the real world. They

demonstrate the ability to function with other stakeholders within organizational constraints, demonstrating know-how-to-be.

4. **Analyze:** At this more advanced level, competent individuals can also identify motives or causes, make inferences, and find evidence to support generalizations.
5. **Evaluate:** At a high level of competency, individuals can further present and defend opinions and make judgments about information, validity of ideas, or quality of work based on a predetermined criterion.
6. **Create:** At the highest level of competency in this model, highly competent individuals can compile information, combine elements in new ways, or propose alternative solutions to existing or new problems they confront.

3.5 Cybersecurity competencies

As defined earlier, the central focus of this study is cybersecurity, which is why the literature review started by first defining cybersecurity and then competency. However, there is a need to define how these two concepts are connected in the literature as they would apply to this study. When a review of the extant literature was performed, the best model of cybersecurity competency identified was the cybersecurity competency model (CCM) (Administration) proposed by the United States Department of Labor, Employment and Training Administration (US DOLETA) and the National Institute of Science and Technology (NIST). This model is best known as the National Cybersecurity Workforce Framework (NCWF) (NIST) from the National Initiative for Cybersecurity Education (NICE) (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017). The NCWF had already been selected by the participating organization prior to the commencement of this study. This model, as was later known, was also selected by the Canadian Bankers Association's working group on cybersecurity as their reference model.

While other cybersecurity competency models are found in the scientific literature and presented by industry or specialist associations, they do not offer the depth of the NCWF, with all the different components available, as described in the next few paragraphs. The

2021 InfoSec Cybersecurity Role & Career Path Clarity Study, which surveyed over 370 IT and security team managers from the US- and Canada-based organizations, indicated that 81% of organizations are using, or plan to use the NCWF. An example of available frameworks from an industry association is TechNation's recently published Canadian Cybersecurity Skills Framework, which is an implementation of the NCWF https://technationcanada.ca/en/future-workforce-development/cybersecurity/cybersecurity-skills-framework/. However, a model such as the TechNation model is not widely used yet (Infosec, 2021). In a similar manner, academic models are dependent on research funding and researcher's interest. The CCM received initial funding from the Cybersecurity Workforce Strategy that supports the Cybersecurity National Action Plan initiatives that US President Barack Obama presented in 2016 (White House, 2016). It proposed investing $62 million in fiscal year (FY) 2017. With long-term funding and resources, NCWF seemed the best choice for this study. Furthermore, the participating organization had already been using this model and has committed to using the NIST cybersecurity framework to support their activities and conformity requirements. Finally, because cybersecurity certifications, training programs and tools, academic training programs, risk management frameworks and other material that will be used in this study, all seemed to have been mapped to the CCM and NCWF. For example, Immersive Labs (https://www.immersivelabs.com/), Secure Code Warrior (https://www.securecodewarrior.com/), and Point 3 (https://ittakesahuman.com/) cyber-range cybersecurity training environment have been mapped to the NCWF. Initially, when a decision was required on the direction to select as a starting point in this study, the CCM made it possible to connect many different pieces of the puzzle together. It looked like the most optimal choice for this study.

The CCM is designed to represent cybersecurity skills in organizations aligned with the NIST framework. It applies to experienced cybersecurity employees of organizations who use networks and professionals who are new to the field of cybersecurity. As illustrated in Figure 2, it is not only focused on computer skills. It also describes the knowledge, skills, and abilities (KSAs) required to have the know-how and know-what required in an organizational context. This model is structured in six layers, called tiers, divided into three sections (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017).

The first section, **Foundational Competencies**, presents know-how-to-be, which is also referred to as soft skills and work-readiness skills. These include **Personal Effectiveness Competencies** (Tier 1) essential to all roles in life to help individuals navigate in social settings; **Academic Competencies** (Tier 2) based on formal education as it is acquired in schools, colleges, and universities; and **Workplace Competencies** (Tier 3) applicable to many professions and industries and learned from a job or a workplace setting that can be specific to a particular role in a particular environment. The second section presents industry-specific Skills, including **Industry-wide Technical Competencies** (Tier 4) that cover the transversal KSAs of workers in an industry, regardless of the sector in which they operate. The third section presents **Industry-Sector Functional Areas** (Tier 5) that correspond to the categories of the work of the NCWF (Newhouse et al., 2017b; NIST, 2021; Petrella, 2017).

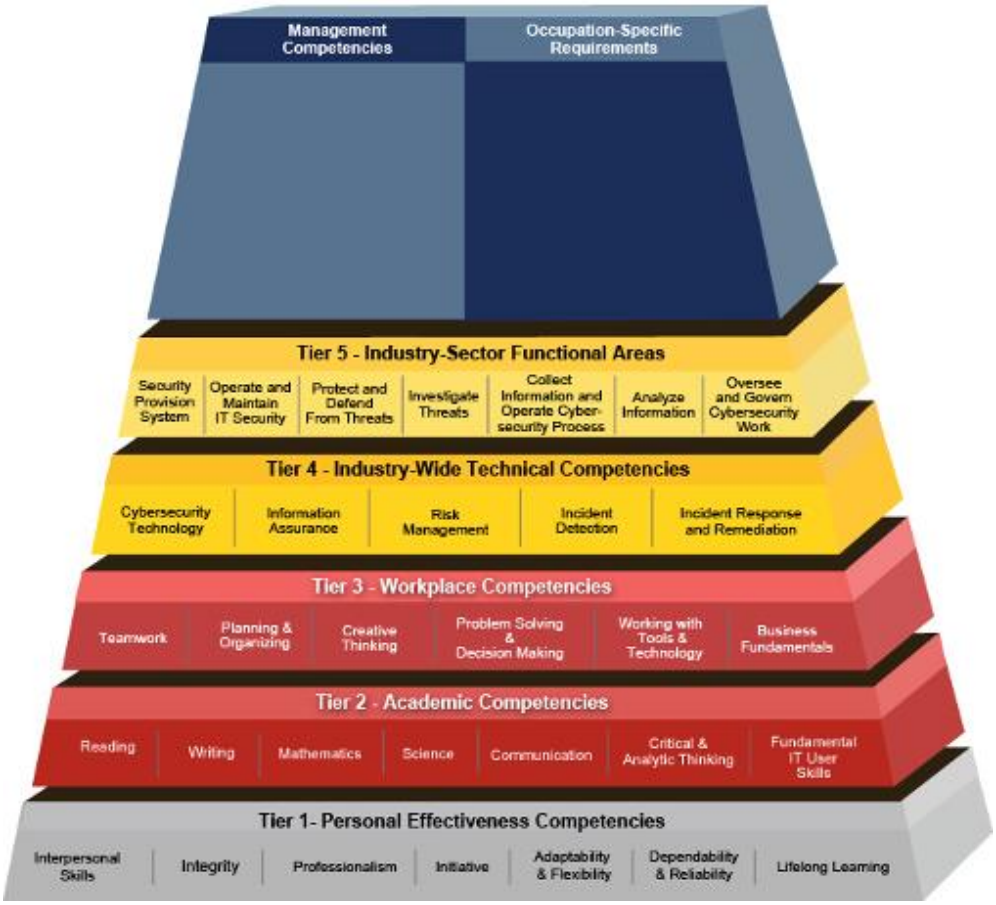


Figure 2: Cybersecurity Competency Model (Newhouse et al., 2017).

3.5.1 Functional and technical competencies

The NCWF presents an extensive description of cybersecurity roles, tasks, and associated KSA in Tiers 4 and 5 (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017). This part of the CCM is commonly referred to as the NICE framework (Newhouse et al., 2017b; NIST, 2021; Petrella, 2017). It includes seven categories, 33 specialties, 52 work roles, 1,007 tasks, 630 knowledge elements, 374 skills, and 176 abilities. It also shows the relationship between the KSAs and the tasks with work roles. The framework is useful as it provides a common lexicon and taxonomy to describe cybersecurity roles that are highly beneficial for the cybersecurity ontology because it serves as a foundation for defining cybersecurity roles. It also proposes an understanding of the KSAs necessary to successfully complete these roles in an organizational setting.

In the context of this study, the NICE framework provides a concrete basis to understand the core knowledge domains for an ideal cybersecurity professional in a normalized role as a theoretical model for us to compare (Newhouse et al., 2017b; NIST, 2021; Petrella, 2017). It describes the required skill sets in different knowledge domains, with some being more strategic business skills and others more technical. However, competency is more than just knowledge and is often developed from the combination of knowledge and several years of experience. A more practical goal for organizations should be employing individuals with a high level of competency in one domain and a fair understanding of other domains; however, the identification of other domains is a complex task. Organizations cannot expect individuals to be true experts in multiple domains and determining the most essential domain required for a particular individual in a specific role is challenging. Organizations require cybersecurity workers to understand their businesses when making decisions rather than only technological aspects. Some of the future cybersecurity workers are expected to develop from the current workforce; thus, they would master essential security principles and concepts. However, the NCWF also informs us that workers should have other technical competencies, such as understanding system design, software development, software verification, validation, firmware, malware, and hardware (Ross, McEvilley, & Oren, 2018). In addition, while functional and technical competencies are essential to successfully fulfill their work roles,

cybersecurity workers in an organizational setting are required to have cybersecurity business competencies.

3.5.2 Cybersecurity business competencies

Cybersecurity workers need analytical and business skills to be competent in their roles in the organization they are part of. This was confirmed by most of the participants in this study as extremely crucial. These are essential skills to ensure that proposed cybersecurity solutions are practical, and they are described in Tier 3 of the NCWF (Newhouse et al., 2017b; NIST, 2021; Petrella, 2017). Competent cybersecurity workers also need industry-wide technical competencies, including experience in areas such as IT project management and system development that are found in Tier 3 and risk management or incident response competencies that are found in Tier 4. Developing resilient and secure systems requires the entire lifecycle of the system to be considered, from the cradle to the grave, including its secure operation. This entails additional skills that include the analysis of increasing costly upgrades, patches, identification, and remediation of vulnerabilities and many other issues. Accordingly, skills in engineering processes, design reviews, decision criteria, and project milestones are essential (Musman, 2016).

More than ever, cybersecurity is being managed as a part of the wider organizational security, not as a separate IT function as it was often done in the past. A correlation can be observed between cybersecurity and other strategies of business security (Jirasek, 2012). The cybersecurity risk assessment is increasingly viewed as an ongoing process of assessing IT security and an organization's posture to determine optimal measures to keep the risks acceptable (Dawson et al., 2013). The risk assessment considers the impacts of measures on the organizations' ability to fulfill its goals (Abawajy, 2014). Cybersecurity strategies, which are often are defined by security managers, should support broader organizational goals. Once strategies are defined, the next step is the implementation of the technical and administrative security control methodologies. This implies that cybersecurity workers must understand the relationship between the business, risk, and security to do their job well. Cybersecurity workers must comprehend how

controls, including topics such as media protections, risk assessments, contingency planning, and configuration management, are connected to the business goals. Conversely, cybersecurity workers must also understand how implementing technical controls can be done in ways that maintain a balance between business functionality and security (Jirasek, 2012).

3.5.3 Foundational competencies

As mentioned earlier, technical and business knowledge are not the only elements of the cybersecurity competency that are required to develop an ontology that represents an effective cybersecurity workforce. Failing to understand the role of organizational culture in cybersecurity opens the door to security and knowledge gaps. Convincing users to follow best practices and respect policies requires persuasion and social skills (Shillair et al., 2015). Likewise, crimes, such as phishing attacks, exploit known social and human behavioral flaws and predict human behavior. Hence, social skills are indispensable competency elements that should be included in the cybersecurity competency ontology. The social factors combined with technical KSAs in Tiers 1 and 2 of the NCWF would enable the cybersecurity ontology to provide a complete view of key attributes (Newhouse et al., 2017b; NIST, 2021; Petrella, 2017).

With new vulnerabilities and risks continually emerging, cybersecurity workers must have a commitment to life-long learning to stay abreast of novel concepts and new attack vectors to stay proficient (Champion, Rajivan, Cooke, & Jariwala, 2012). In particular, workplace mentoring and training provide the highest increase in performance benefits (Champion et al., 2012).

Cybersecurity professionals should have pattern matching and good mental flexibility abilities and situational awareness. These professionals working as a team are likely to solve complex tasks as compared to individual analysts because their expertise and talents are distributed across analysts (Rajivan & Cooke, 2018). Strong expertise facilitates excellence and creativity while performing different tasks (Huang & Zhu, 2019).

Hence, future cybersecurity professionals must have a strong ethical code, like other professionals working in complex environments. The lack of value systems creates an avenue for potential exploitation by bad actors, dissatisfied employees, or professionals with good intentions (Hannah, Jennings, Bluhm, Peng, & Schaubroeck, 2014). Future cybersecurity professionals should be trustworthy and reliable, and employees should be selected by organizations on the basis of the expectation that they will match their skills, knowledge, interests, and values (Cable & Parsons, 2001).

3.5.4 Academic competencies

Academic competencies (Tier 2) are based on formal education. Today, in the field of cybersecurity, this is principally acquired in vocational schools, colleges, and universities. Some vocational schools have technical IT programs that have cybersecurity competencies in the curriculum, and vocational education would often be intended for more junior positions in IT support roles that have a cybersecurity component. Most cybersecurity academic programs would be offered in colleges or universities. There is an inventory of these programs in Canada that was used as a starting point (SERENE-RISC, 2020), which was complemented by additional data from a search of the websites of Eastern-Canada colleges and universities for courses and programs in cybersecurity or that included a cybersecurity component.

In the interviews and workshops, the participants indicated that academic competencies are useful when cybersecurity workers possess the ability to use them in a work setting. For instance, in a penetration testing role, the target organization's managers are looking for individuals who can do the job, regardless of their academic degrees or other variables, such as age or gender.

The upper levels of the CCM, present the specialization, profession-specific requirements, and management skills in specific occupations within an industry (Newhouse et al., 2017b; NIST, 2021; Petrella, 2017).

With regard to technical and social skills, Dawson and Thomson (2018) proposed additional trait requirements for the future cyber workforce:

- **Systemic thinkers**, demonstrating that they can make sense of complex, interconnected environments with multiple physical, logical, and virtual layers.
- **Team players**, being comfortable working with others in effective teams exhibiting cohesion and trust that involve a shared sense of identity.
- **Civic duty**, being loyal to the ideals of an organization.
- **Life-long learners**, seeking out the latest information about security, vulnerabilities, and capabilities and a passion for learning and problem-solving.
- **Communicators**, having the ability to expressly communicate technical information to a non-technical audience.

3.6 Ontology

One of the aspects of this study requires us to represent knowledge, specifically related to cybersecurity competencies in a structured manner. We used an ontology to achieve the same.

The concept of ontology originated from ancient Greece to designate the study of existence, categorization, and relationship of objects or things (Eloumri, 2019; Grimm, Abecker, Völker, & Studer, 2011). It has since evolved into the knowledge representation models of a domain (Grimm et al., 2011). An ontology can be described as a simplified, specific, and abstract view of a reality of a particular subject area (Grimm et al., 2011; Gruber, 1993, 1995; Velasco & Rodriguez, 2017). This representation can be of concepts, objects, and other components of interest in a specified subject area and their relationship with one another, expressed formally. Ontologies contribute to constructing a shared understanding of a subject domain (Grimm et al., 2011; Gruber, 1993, 1995), and the literature informs us of some of the important aspects of ontologies (Grimm et al., 2011; Gruber, 1993, 1995; Velasco & Rodriguez, 2017).

1. Ontologies must provide a knowledge representation language based on formal semantics and logic to ensure that the specifications of domain knowledge are interpreted semantically and logically correct.
2. Concepts must be explicitly stated or defined so that they can be processed using software.

3. There must exist a shared agreement of distinct stakeholders.
4. Ontologies must be captured as a general abstract conceptual model, identifying the concepts and their relationships in a particular context.
5. Ontologies must specify the knowledge of a particular domain.

For this study, ontologies can be understood and used as a tool to create a formal representation of explicit knowledge and semantic vocabulary, which allows the sharing of data that can be performed with the help of IT (Grimm et al., 2011; Gruber, 1993, 1995). Providing this formal and explicit domain knowledge in this form allows software systems, such as Stanford Protégé OWL to be used to structure the information. For example, the information in the ontology that was developed in this study can be employed in combination with artificial intelligence (AI) or business process modeling (Grimm et al., 2011; Gruber, 1995; Velasco & Rodriguez, 2017). Ontologies are becoming an increasingly popular tool in research for many applications, such as knowledge representation, decision support tools, system engineering, and the semantic web (Grimm et al., 2011; Gruber, 1993, 1995).

Ontologies can be categorized into domain, mid-level, and upper ontologies on the basis of their abstraction levels (Obrst, 2010). The highest level includes the upper or universal ontologies, which are domain-independent and provide the basis for more general knowledge representation. Mid-level ontologies are less abstract while extending to multiple domain ontologies but still providing specific representations of theoretical concepts contained in upper ontologies. They are in the middle, between the very general upper ontology and the very specific domain ontology. As such, the distinction between mid-level and upper ontologies is epistemological. At the other end, a domain ontology contains specific concepts for a defined and limited domain and presents how the divergent concepts that define a domain are related to one another. Recently, many studies have examined specific technical aspects within the cybersecurity field. Existing cybersecurity ontologies cover certain aspects of the domain, such as network securities, vulnerabilities, attack vectors, defense strategies, assets, security protocols, and integration tools (Obrst, 2010).

3.7 Ontology design

As indicated by Keet(2020), there are several specific approaches for ontology development. Notwithstanding, after reviewing the available documentation for these approaches, it remained difficult to identify a specific set of instructions that could be followed to proceed with the design phase. The level of details that could be found made it difficult to implement them at the time this study was ready to proceed. At this point, early in the project, the strategy was to develop an ontology design approach.

While looking at existing ontologies, a few ontologies in the field of human resources provided some help to guide the study (Gomez-Perez, 1998; Radevski & Trichet, 2006). From an architectural perspective, the cybersecurity competency ontology relies on three major building blocks: **cybersecurity work roles** (what the organization requires individuals to do to help it with its mission), **cybersecurity competencies** (what individuals must know or be able to do well in the work roles), and **cybersecurity education** (how they can acquire the requisite knowledge, skills, or abilities). These building blocks are further detailed as the data from the field instructs us regarding the organizational requirements. For instance, cybersecurity education must also consider how acquired competencies can be formally recognized so that cybersecurity education can be used to fill the gap between the competencies recognized by a formal degree-granting program and the competencies an individual has acquired through previous experiences or by autonomous or informal education. Considering the aim to provide a solution that can help organizations, real-world situations and the existing workforce are used.

3.8 Ontology of cybersecurity competencies

The ontology representing the cybersecurity workforce competencies considers technical skills and social behavior on the network (Fontenele & Sun, 2016). For instance, developing cybersecurity talents involves understanding the fact that interested individuals have unique social–psychological tendencies and traits, which make them more likely to excel in the field. While the NCWF covers the technical KSAs well, it has limited the coverage of social and organizational aspects and soft skills (Newhouse et al.,

2017a; NIST, 2021; Petrella, 2017; Seong, Kristof-Brown, Park, Hong, & Shin, 2015). Notwithstanding, it has the basics, which can be enhanced by the data from the field, collected in the study. Other competency elements should be included in the ontology as human behavior introduces vulnerabilities that go beyond the technical elements (Bell et al., 2014).

4 Research methodology

The approach used for developing the cybersecurity competency ontology employs bottom-up and top-down analyses, referred to as a middle-out strategy (Ahmed-Kristensen et al., 2007; Keet, 2020). The bottom-up analysis involves considering the preexisting data sources on cybersecurity competencies, such as the CCM and NCWF described in this document that should be integrated (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017). The top-down analysis involves considering the needs of the users of the ontology (Ahmed-Kristensen et al., 2007). The top-down analysis utilized information from interviews and other data sources collected in the research study. Combining these from the middle out, enables the formulation of questions that the ontology should answer to provide the expected value. To this end, such questions may be regarded as queries, which would be executed using Protege OWL. In turn, these queries can be utilized as a validation tool, indicating when the ontology is complete for a specific development stage, when the queries provide usable results (Gruber, 1995; Uschold & Gruninger, 1996). Incorporating the competency questions is an essential requirement for the analysis stage of constructing an ontology. The analysis essentially aids the identification of scenarios and use cases. Integrating the use cases, competency questions, and scenarios permits the fleshing out of the requirements. For a successful construction of the ontology, analyzing the cybersecurity domain, including its entities, frameworks, relationships, properties, and rules needed in guiding the model, is crucial.

In essence, the approach proposed for constructing the cybersecurity ontology is founded on three principles, namely emphasizing, reuse, and parsimony (Ahmed-Kristensen et al., 2007). The existing ontologies in the cybersecurity domain are reused where applicable. The reuse approach includes the following steps:

1. Consider the applicability of existing ontologies in the cybersecurity domain. This includes utility, foundational, and reference ontologies.
2. When constructing the cybersecurity ontology, include properties, definitions, and classes that are identified in the first step.

3. In situations where the properties, definitions, and classes adopted from step 1 into the current model grow large, the given ontology should be imported directly into the cybersecurity ontology. Furthermore, equivalence relations should be established between the classes in step 1 and the cybersecurity ontology classes.

Existing standards from national frameworks, such as the NIST framework, and data dictionaries, schemas, and glossaries are harvested. Other definitional and structured resources from the existing literature are included when applicable as a source of the core acquisition of domain knowledge. Such resources are evaluated on the basis of the kind of relationship, entity, attribute, property, and value. Where applicable and correlated with the cybersecurity domain, they are included in the cybersecurity ontology upon improvement in accordance with the principles of ontological engineering. However, the conceptualization should be kept simple for easy understanding.

This study adopted the action design research (ADR) approach to implement the design science research (DSR) methodology taking advantage of activities from ADR (Mullarkey & Hevner, 2019). ADR differs from other DSR approaches as the former starts from the practical stakeholder setting rather than the theory (Keijzer-Broers & de Reuver, 2016). ADR was chosen because it has been successfully deployed in an organizational setting to generate knowledge through an iterative process of construction and evaluation of artifacts (Keijzer-Broers & de Reuver, 2016; McCurdy et al., 2016; Sein et al., 2011). Another reason for this choice was that the researcher involved in the study had used it before successfully in other studies. Selecting a research methodology for this qualitative study was different from the ontology design approach developed as a deliverable in this study, addressing the research question RQ-1.1, looking to determine the most fruitful approach for developing a new cybersecurity ontology that could represent the competencies, skills, and abilities of effective practices in this field.

Approaches for ontology development were investigated, such as those proposed by several researchers (Keet, 2020), (Sure-Vetter, Staab, & Studer, 2009) or reviewed by others (Stadlhofer, Salhofer, & Durlacher, 2013). The level of details provided in the approaches made it difficult to estimate how they could be used in this study.

Furthermore, with the lack of sufficient knowledge on the problem domain, there were no apparent benefits to using these. It was decided to proceed with ADR to develop and document the process in more detail in a formal process to facilitate future reuse and provide useful evidence of the rigor of the process.

DSR is focused on a problem-solving paradigm, generating artifacts for engineering and applied sciences. Implementing the ADR approach to DSR provides us with a process to achieve the construction of the artifact, constructs, models, and methods applied in the development, and the use of the information system (Hevner et al., 2004; Kukulies et al., 2016; Mullarkey & Hevner, 2019; Niederman & March, 2012). In this study, this consists of a tool for organizations, rooted in the needs of the participating organization. It aims to propose solutions to a problem through interventions and their implementation. Using a theoretical framework, ADR focuses on multiple cycles of the building, implementing, and evaluating (BIE) artifacts.

ADR considers the intention of researchers, needs of users, and continuous use in a context. Moreover, ADR, as used in this study, offers a four-step iterative research process based on seven principles, as presented in Figure 3, which emerged from the use in the field of action and design research, the methodologies that influenced ADR (Sein et al., 2011; van Aken, 2004). The steps and principles of the ADR approach to DSR are presented in the next sections.

4.1 Step 1: Problem formulation

Using ADR in a research study starts with identification of a problem grounded in the reality of an organization. This problem can come from different stakeholders or by following initial empirical research (Sein et al., 2011). At this step, researchers and the participating organization determine the initial scope, roles, and research questions. The problem formulation stage makes the identification and conceptualization of the research opportunity on the basis of existing knowledge. In this study, the problem formulation was completed at the onset and resulted in a research proposal and the identification of the initial research material and references. These were then incorporated into initial

documents that formed the basis of what was used in the next step to execute, as described in Section 4.2.

As in any academic study, obtaining a commitment from organizations and defining how it fits into all organizational problems are essential. This guides research and allows the creation of scientific knowledge. Through collaboration between researchers and participants, like what is done in action research, a mutual understanding of the purview and objective of the survey is co-constructed. This is linked to **Principle 1:** "Research is inspired by practice" and **Principle 2:** "Artifacts are rooted in theory."

4.2 Step 2: Building, intervention, and evaluation

The second step of ADR uses the problem and the theoretical foundations from the first step, described in Section 4.1, as a starting point for the design of the artifact using an iterative approach in this second step of the ADR methodology. In this study, the initial artifact was in the form of a concept map created using MindManager (www.mindmanager.com), a popular conceptual mapping software, for which licenses were available to the members of the research team. The concept map was the initial model developed in this study from the literature review and became the starting point for the second step. The artifact was shaped by its use in the organization, and by subsequent cycles, it evolved during the study to form the basis of the ontology using UML in a subsequent cycle, as UML provides a good representation of the ontology and the concepts that are being defined (Opdahl & Henderson-Sellers, 2002; Tilakaratna & Rajapakse, 2017). The UML model is the second model that was developed in this study. Eventually, the UML model evolved through additional cycles and was used to create the OWL database, making the ontology the final model that was created in multiple cycles. At the end of this study, a graphical database model was realized, which could be used for subsequent research projects or to help develop a management tool. This graphical database model is presented in Appendix O. Executed as an iterative process in a restricted environment, the development cycles completed in the second step of the ADR research methodology involved combining the construction of the artifact (building), intervention in the organization, and evaluation, – BIE. The output from the many

iterations is the final artifact, the main deliverable of this research. During the BIE, the problem and the artifact were continuously assessed, and the principles of design were developed. These principles can then be generalized to a broader problem in subsequent BIE cycles. This step also aims to gradually clarify the targeted innovation space, which can come from the design of the artifact or organizational intervention through the multiple iterations. In this study, three BIE cycles were used to develop the ontology, which was then further refined through the testing phases, as possibilities for improvement were identified.

This step is influenced by the following three principles: reciprocity in formatting, mutual influence in roles, and simultaneous evaluation providing a gradual and continuous improvement feedback loop. Together, these principles underline the mutual influence of the process steps of the ADR methodology.

BIE is suitable for ADR efforts that aim to create an innovative technological solution from the start, as was the situation in this study. The first designs fueled interventions in a limited organizational context (Sein et al., 2011). The emerging artifact and the theories that are anchored in it were constantly instantiated and tested several times through an organizational intervention and subject to the assumptions, expectations, and knowledge of the participants. This participatory process strengthens the organizational commitment and guides the eventual design of the artifact. Building on these initial interactions, the ADR team then integrates a more mature artifact into a broader organizational framework. This step allows the evaluation of the artifact in use, and this allows the continuous improvement of the artifact as it is shaped and reshaped by the context of use. This intervention stage can lead to the end of the study or generate a new BIE cycle. In a way, it is the artifact that guides innovation.

The second step of ADR is based on **Principle 3:** “Reciprocal shaping,” **Principle 4:** “Mutually influential roles,” and **Principle 5:** “Authentic and simultaneous assessment.”

During the evaluation stage of the BIE cycles, the validity of the ontology is evaluated. To validate the ontology at each BIE cycle, the utility it provides is evaluated. This needs to be integrated into a development process prior to use (Verdonck & Gailly, 2016).

OWL ontology is based on the quality of the utility, which must be incorporated through executed methods. An OWL ontology has a structure that permits its assessment with a systematic validation process. This assessment benefits from reflection and learning, described in Section 4.3, which is done in the third step. The cyclical nature of ADR, with multiple BIE cycles and feedback-based improvement, contributes to generating scientific knowledge, as presented in the results section of this dissertation.

4.3 Step 3: Reflection and learning

The reflection and learning stage, the third step of ADR, take a solution for a particular case to apply the learned knowledge to a more extensive problem. It is continuous and done in parallel to the first two steps, and it recognizes that the research process involves more than just solving a problem. Reflection on the framing of problem, theories, and emerging knowledge is used to ensure that contributions to knowledge are identified and adjust the research process in accordance with the first results of the assessment to reflect the growing understanding of the artifact. Reflection and learning occur at all stages of this research project and are presented throughout this dissertation. This step is based on **Principle 6:** "Guided emergence."

4.4 Step 4: Formalization of learning

The objective of the fourth step of ADR is to formalize learning. The lessons learned from the study are developed into a solution that can be generalized to resolve similar problems in other organizations, within the limits of internal and external validity. At this step, the researchers describe the achievements made in the computer artifact and describe the organizational results to formalize learning. The results can be characterized as design principles and with further reflection, as improvements to the theories that contributed to the initial design. In this study, this became the OWL ontology of cybersecurity competencies. However, earlier ADR cycles started with simpler artifacts, such as the descriptions of the various components, conceptual maps, and UML models. The discussions and results, outlines in Section 9, contribute to this step of the ADR methodology, as applied to this research project. This step is based on **Principle 7:** "Generalization of results."

The steps and principles of ADR are presented in Figure 3, adapted from (Sein et al., 2011).

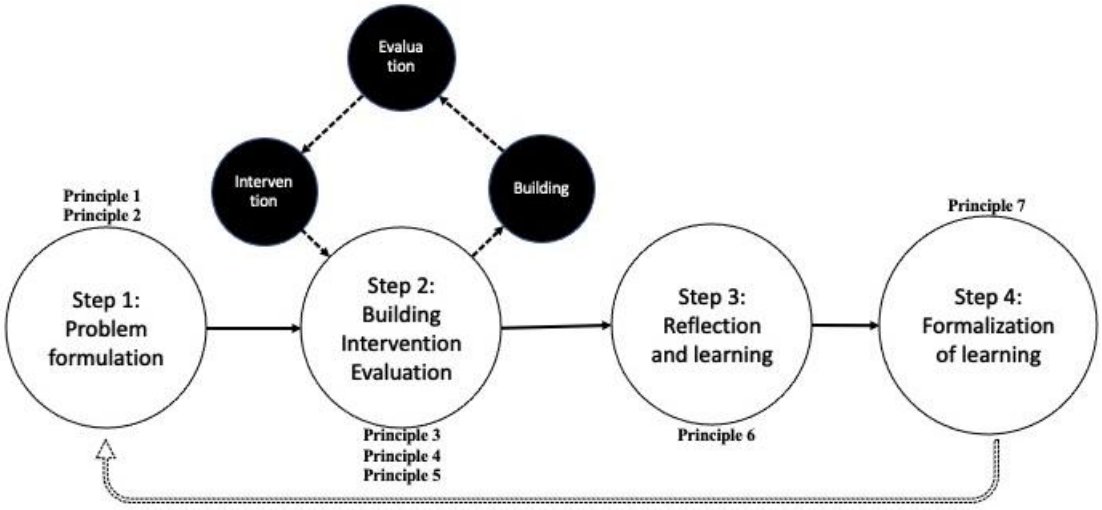


Figure 3: Action Design Research

4.5 Hypothesis

Throughout this study, based on ADR methodology, several hypotheses were proposed and investigated. In this section, we explicate these hypotheses, which will be discussed more in depth later in this dissertation as they are relevant in different stages of the research project.

The first hypothesis that was tested is the predictive ability of the cybersecurity competency ontology, as presented here:

> **H1:** The prediction results obtained using the cybersecurity competency ontology are better than what could be expected in a random choice of 50%.

This was done using an F1-score, as seen in Section 8.4. To achieve the same, the ontology was needed to be designed, constructed, and populated, which is described in Section 5. Thus, there is an underlying hypothesis to H1, identified as H2, our second hypothesis:

H2: The cybersecurity competency ontology representing work roles, tasks, and competency elements in Canadian financial institutions can be designed, constructed, and populated.

The ability of the ontology to represent the work roles was tested, and the findings are presented in Sections 6 and 7. We are proposing a solution based on the data collected and the process put in place in early stages of the research. As an incremental strategy, with ADR, is used, this involved a database version of the ontology in Stardog, as will be explained later in this dissertation. While testing H2, queries were performed and evaluated, to check if they corroborate H1, if H2 was supported.

Should H1 and H2 be supported, a third hypothesis can be tested – whether the ontology can assist the organization matching work roles in risking scenarios by inferring the connection, via the association of work roles with competency elements and the link between risk mitigation measures associated with risk scenarios in MITRE ATT&CK, competencies associated to risk mitigations and, finally, competencies and work roles. Thus, the third hypothesis was formulated as:

H3: The ontology allows organizations to match work roles to risk scenarios.

In Section 8, risk scenarios are developed and used to identify the work roles best suited to assist the organization in dealing with risk mitigation activities pertinent to address MITRE ATT&CK risk scenarios.

Subsequently, the fourth hypothesis was then formulated, asserting that improving cybersecurity competency management with the proposed ontology will contribute to creating dynamic capabilities, which can, in turn, contribute to effective cybersecurity by creating an organizational capability to rapidly adapt to new and emerging threats and vulnerabilities to an essential component of modern organizations, its information technologies.

H4: improving cybersecurity competency management will contribute to create dynamic capabilities, which contribute to effective cybersecurity.

4.6 Study Participants

This study was conducted in a large Canadian financial institution. Nevertheless, to secure the agreement of the organization and because the topic is cybersecurity and considering the current worldwide situation as it relates to cybercriminals, it was agreed not to name the organization in publications, as not to introduce any vulnerability. The research supervisors and UQO do have this information secured in a file for evaluation purposes.

The study participants were initially estimated (n = 32) from the cybersecurity group under the authority of the chief information security officer (CISO). This number of participants was estimated following discussions with the participant organization as to the number of resources that would be required to complete the study. This convenience sample represented more than 20% of the 150 individuals in the cybersecurity department at the beginning of the study, which has grown by more than 350 individuals since then. In the end, the number of participants grew (n = 48), as there were opportunities to add participants that were justified from a research perspective. A sample size of n = 32 was determined as the minimum number that would make it possible to include a few participants from different teams in the cybersecurity area and different work roles.

The participants covered the cyber-defense, cyber-intelligence, strategic, and operational groups, and all work roles that can be found in the NCWF (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017). It must be noted that the sample size increased as opportunities arose in the multiple ADR BIE cycles, as previously mentioned. Furthermore, in the early steps of the process, it was determined that some of the validation activities and discussions would involve external participants from a Canadian cybersecurity industry association of which the financial institution in question is a member. This was somewhat made more difficult because of the coronavirus disease 2019 (COVID-19) crisis, which made some of the planned discussions unfeasible. However, presentations were made, and valuable feedback was received and integrated into this study. As well, including another financial institution was excluded due to COVID-19 and because of confidentiality concerns of the participating organization.

4.6.1 Inclusion criteria

Only the employees of the cybersecurity group of a specific large Canadian financial institution were included in this study. All the participants were under the CISO at various hierarchical levels. Consultants or external resources were excluded and did not contribute to the interviews and focus groups. During the validation phases, participants from the Human Resource Department involved in cybersecurity hiring and members of the cybersecurity industry association were involved.

4.6.2 Recruiting participants

As the participation of the organizations was secured before this study was initiated, an opportunistic sample was used, thereby leveraging opportunities that arose and considering the agreement with the participating organization that made this study possible. Recruitment was initially done through email, a copy of the initial emails sent to the proposed participants are included in Appendix G. With this email, participants were directly invited by the principal investigator of the organization. In some cases, the team managers proposed participants as per their availability, experience, and interest in participating in this study, in addition to their knowledge in the concerned domain. The managers were also participants in this study and involved in the artifact validation in various stages. Once the initial email contacts were made, the participants involved in the early data collection, in the early BIE iterations, were met in person for an initial interview. After the explanation of the study and its objectives, they were provided with the choice to accept or refuse to participate in the study. Those who accepted were asked to sign an informed consent form that was approved by UQO's Research Ethics Committee. Details of the participants profiles are included in Appendix T.

At subsequent stages in this study, when group discussions took place or in the validation stages of the artifact, the additional participants were contacted, informed about the goals of the study, and asked to sign the informed consent.

4.7 Data Collection

This section presents the main elements of the data collection strategy and the underlying theories that underpinned the same for several months. The ADR research protocol, designed to achieve the construction of the artifact, the constructs, the models, and the methods were based on Hevner et al. (2004), Kukulies et al. (2016), Mullarkey and Hevner (2019), and Niederman and March (2012). There are several data collection activities in this research. As the ARD method was used to execute the study, there were ongoing activities and interactions between the researchers and the participants that were all opportunities for data collection, including interviews, job postings, internal job descriptions, participant observation, and internal documents on SharePoint and shared network drives. Furthermore, the researchers kept notes and logs to systematically document the interactions and observations. The notes and logs are not shared in this dissertation because of confidentiality concerns and requirements of the participating financial institution. As this is a form of a cyclical process, like what was done in the organization in continuous improvement activities, collecting data at each BIE cycle informed the knowledge that was being created. As the study evolved from the initial definition of the problem to eventually develop the proposed solutions and validation, realized in the multiple BIE cycles of ADR, the data collection evolved over the many months of this study until there was a consensus in the research team that data saturation had been achieved for producing meaningful and reliable evidence.

4.7.1 Interviews

The data collection was initially conducted through semi-structured individual interviews with the divergent groups identified. The purpose of these interviews is to understand the emic point of view of the participants about information risk while seeking to develop an understanding of the systems, individuals, and relationships between the variables (Savoie-Zajc, 2009). This is particularly important in the first step, namely problem determination. Additionally, the interviews were preceded by contact with each participant to explain the purpose of this study and how the selection of the participants

was made, to assure them of the confidentiality of the data and obtain their written consent.

An interview semi-directed questionnaire was drawn up with a list of themes that were drawn from the literature review and the initial analysis. It included the topics to be addressed using open questions related to the concepts associated with risk and the competencies expected from various workers in their roles related to the cybersecurity and the management of risk. This interview grid was pre-tested with the help of an external group of workers with extensive experience in information risk management. These workers are practitioners in this field with over 10 years of relevant field experience and have completed graduate studies in either computer engineering or administration.

All the interviews were conducted by the principal investigator using a semi-structured interview guide, included in Appendix H. The order of interviews was determined by the researchers and the availability of the participants. The number of interviews was determined by the results. As the interviews were being done, the data was analyzed and the information that was obtained was integrated into the mind map that was gradually growing. The analysis of the data being performed as the interviews progressed served a number of purposes. First, this allowed the researchers to continuously improve the interview process by ensuring that the interview guide did allow the flow of relevant data for the study. As well, it allowed to evaluate if data saturation was approaching, as it was possible to compare the results from the field with the concepts that emerged from the literature review. At the same time, new subjects or emerging knowledge could also be identified.

All the interviews took place on the premises where the informants work, who agreed to provide us with a room for this purpose at an appropriate time for the duration of this study. The interviews were also recorded, with the authorization of the informants, using a portable digital device to facilitate the cutting of the interviews and the use of citations when presenting the results and for the future use of the results. Additional data, on the environmental context, the language, and the evolution over time of the organization

were collected by observing the living environment in relation to the variables concerning the cultural and socio-structural contexts and all the variables of the conceptual diagram.

4.7.2 Workshops

As the study moved in various steps, three group discussion workshops were conducted to assist in the BIE cycles of the research methodology. The first workshop was done in-person on the premises of the organization. Using the data gathered from the semi-structured interviews, the first group discussion workshop was done with a group of managers representing various teams in the cybersecurity department. This workshop made it possible to identify the strategic and organizational significance of cybersecurity in the participating organization while getting insights on the work roles that are required. The second and third workshops included cybersecurity workers in different work roles areas from different teams. This made it possible to better understand the tasks and competencies required to perform in these areas. The second workshop involved workers in the more technical areas of cybersecurity. The third workshop had individuals in the more business-related cybersecurity roles. Appendix x also presents

It is worth noting that the last two workshops were performed online using Microsoft Teams because of the severe pandemic situation that coincided with the schedule for these activities in this study. Further details on these workshops are presented in Section 9.1.3, Step 3: Workshop. As well, the use cases and other material used for the workshops are presented in Appendix R (ADR methodology) and Appendix S (BIE cycles).

4.8 Ontology Validation and Testing

One of the key results of this study is the new cybersecurity competency ontology, developed as an artifact using OWL in Stanford Protégé. As proposed, using a middle-out approach, it is based on the job requirements of competent cybersecurity workers in an organizational setting that are identified in the NICE framework (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017), and the information gathered in the field during this study. The ontology can work as an effective data collection tool as it

was employed in the ADR cycles to illicit new data in subsequent cycles to be used in an organizational context and tested to determine its use as a predictive tool.

This section provides an outline of the validation and test processes that were conducted to ascertain the pertinence of the ontology with regard to the needs of the target organization and its use as a decision support system (DSS) instrument in guiding talent management functions for cybersecurity workforce planning.

4.8.1 Validation process

The proposed validation process was conducted in two steps. In the first step, the ability to use the ontology as a query tool was evaluated. Finally, in the second step, formal knowledge extraction queries onto the ontology were utilized and compared with the human-bound process. Given the highly focused qualitative data gathering and analysis, with successful cybersecurity professionals, reflecting on the core competencies required, the ontology should, in principle, embody these same qualities and be valid, as is discussed later.

4.8.2 Testing the ontology as a query tool

This first series of tests was conducted to evaluate the use of the ontology as a query tool, an important element of the expected outcome from this study. Here, once the ontology was developed and integrated into a Resource Description Format (RDF) storage plugin within Stanford Protégé, cybersecurity scenarios were developed and expressed in the Semantic Query Language for data stored in RDF format, better known by the acronym SPARQL. These formal expressions made it possible to test if the ontology could be employed as a tool to find factually relevant competency recommendations as per the desired outcomes of cybersecurity management.

To estimate the quality of the results of these formal queries, these results were submitted to a group of workers including cybersecurity managers, human resource team members who specialize in cybersecurity placements, and workers who are currently in cybersecurity roles in the target organization. These results were analyzed using scoring techniques described later in this dissertation.

4.8.3 Testing the ontology as a management tool

The second series of tests were executed to evaluate the quality of the results to help financial organizations manage the fit between required competencies and available talent. As initial research at the onset of the study indicated what cybersecurity competencies were included in the ontology, the validation was used to confirm what parts accurately reflect the numerous categories of activities and tasks that could be viewed as attributes of a successful cybersecurity function. These include detailed and structured information about cybersecurity job positions, knowledge, skills, abilities, soft skills, certifications, and education. As such, the testing process focused on using the proposed ontology as an HRM tool, primarily as a DSS for talent management and team staffing.

These tests used a compilation of cybersecurity job descriptions. These descriptions represent the "gold standard" of formal and well-fitted team compositions and job attributes for cybersecurity studies and functions. The quality of the inference queries was then evaluated to test to what extent the ontology can be reliably employed to guide talent staffing.

4.9 Research Calendar

The phases followed the four steps of ADR. This study started in 2019 upon obtaining the approval from Ethics Committee and after the principal researcher completed the course requirements of the Ph.D. program. The first steps of the study started in June 2019, with securing organizational commitment and agreement on this study.

The main research phase is step 2, with the BIE cycles of ADR starting only after March 11, 2020, the date the study obtained the Ethics Committee approval. This concluded in December 2020, nearly on the original schedule. The completion of this study on the original, rather optimistic, schedule was aided by the current global health crisis, which allowed the main researcher to have more time to dedicate to the study while working from home and retaining remote access to the participants and to the organization. Three BIE cycles were required to gather all the data required and allow this study to proceed to

the next steps. In line with the three phases, sections of this dissertation were completed. These sections correspond to the three initial phases of the study and a fourth phase that was subsequently added for external validity. In the first cycle, the ontology was built, and the ontology engineering process was documented. In the second BIE cycle, it was mapped to cybersecurity competencies, populated the ontology, and the researcher documented the process. In the third BIE cycle, the building of the ontology was completed and later transferred to Protégé to allow the use of SPARQL queries, validate the ontology, and document the process. The details of what is summarized here are presented in Sections 9, 10, 11, and 12.

5 Designing the ontology for cybersecurity requirements

This section presents the ontology design approach followed for the construction of a cybersecurity ontology that represents the cybersecurity competencies, skills, and abilities required of IT professionals to accomplish the tasks expected of them in the field. This study was realized using ADR innovative methods in requirement elicitation, representation, and validation.

5.1 Ontology design

Following the ADR methodology, BIE cycles were utilized for gradually building the ontology. An iterative approach permitted the researchers to gradually gain a better understanding of ontology design and cybersecurity competencies. This enabled the integration of the two into artifacts. An early artifact of the BIE cycles is the ontology design strategy used to develop the ontology, the subject of this section.

Capitalizing on the tools available and the existing organizational knowledge, this process initiated with building a mind map of the problem domain using MindManager, a commercial software. Mind maps were chosen because the researchers involved in this study have been using them for many years. MindManager was used as it was already used in the participating organization and by the researchers, who already had the license for this product. The initial intuition was that there was an alignment with map development and knowledge representation with ontologies. As the initial study evolved, this intuition was corroborated with evidence, as reported in many articles (Křemen, Mička, Blaško, & Šmíd, 2012). An early stage of development included a literature review, information from international cybersecurity conferences, and discussions with colleagues. All these sources provided valuable inputs for the early versions of the mind map in the first BIE cycle, which led to the initial semi-structured interviews of stakeholders feeding off the early knowledge as a source for the interview guide. This was followed by workshop sessions with cybersecurity managers and cybersecurity workers in the organization. Throughout this process, the initial data collected was integrated as the mind map artifact was gradually developed into a larger mind map, shared, and discussed with the stakeholders as a new artifact.

The mind map was then used to create the UML models of the primary knowledge areas that interact with the problem domain. The different knowledge nodes in the mind map became the individual classes in the UML model. These UML models allowed us to obtain a better understanding of what later became the objects and classes that composed the ontology. For this, the UML model's classes becoming the classes in the ontology. Both these UML models were created by the researchers and are original creations. The initial version of the mind map used the data from the literature review, which evolved with the data from the data collection and consensus building. This led to the UML model of competency and the definitions of competency. Figure 4 presents the UML model of cybersecurity competency categories that emerged.

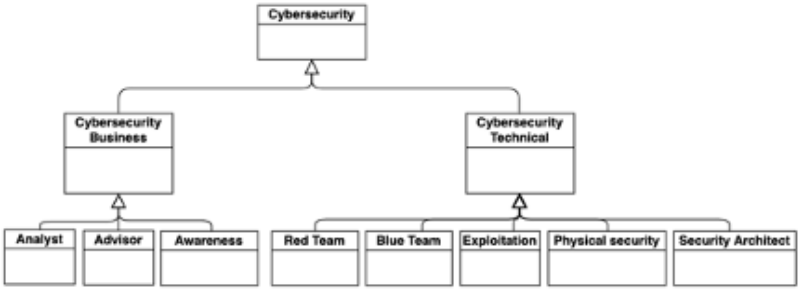


Figure 4: UML model of cybersecurity competency categories

Finally, from the mind map and the UML models, the knowledge was integrated into Protégé OWL as the ontology, connecting the objects together to indicate how different competencies and competency levels connect to specific KSAs and the tasks of the cybersecurity workers. In every step, the materials were shared with the stakeholders in the development of a consensus around the artifact following the ADR process. Several iterations of the ontology were conducted until the consensus of the stakeholders, researchers, and participants was reached. Subsequently, the ontology was employed as an input into the final phase of the study, the validation, and then a subsequent phase where it was tested for usability using queries and validity tests. Those other phases are not discussed in this section.

5.1.1 Design approach

To develop the cybersecurity competency ontology described in this section, a middle-out approach was used. This approach contains the elements of two approaches, namely the bottom-up and top-down analysis approaches. The former starts with the existing sources that can be integrated into the new ontology, while the latter starts from the needs of the stakeholders, who will eventually use the resultant ontology. For instance, they would or could ask questions from the ontology and tool built using the ontology and the other sources used as inputs (Ahmed-Kristensen et al., 2007). In the middle-out approach utilized, the existing sources and the needs of the users were considered together, working from the middle of the two information sources to construct the ontology.

Such analyses lead to the formulation of questions that the ontology should answer to provide expected value to the stakeholders. To this end, such questions may be considered as queries, which will also be utilized to validate the ontology and determine its quality as a decision support tool. The queries are inputs in the top-down analysis and are a validation procedure indicating when the construction of the ontology is satisfactorily complete for a specific development stage. Use cases and scenarios can also be employed. A development stage was considered to be complete when the queries provide accurate results with a usable level of information (Uschold & Gruninger, 1996). For the success of this ontology, information from the cybersecurity domain, including its entities, existing frameworks, relationships, properties, and business rules, were crucial. The cybersecurity ontology design strategy used in this study highlights the principles of reuse and parsimony (Gavrilova, Leshcheva, & Strakhovich, 2015; Gruber, 1993, 1995; Keet, 2020; Uschold & Gruninger, 1996). By reuse, research, and collaboration with stakeholders, the researcher investigated reusing existing ontologies in the cybersecurity domain where applicable. At the same time, following the parsimony principle, attempts were made to simplify the ontology so that all the crucial information was present and all the superfluous information was removed. The following steps were used to gather the initial data required for this study and acquire a better understanding of the subject area.

5.1.2 Step 1: Gathering of the initial data

1. Existing candidate ontologies, reference models, and frameworks that could provide value and meaning to the cybersecurity competency ontology were identified and evaluated for inclusion.
2. A literature review, searching academic sources, web sources, and various databases contributed to provide additional information. Some of the resources identified are described further in this section.
3. Available cybersecurity certifications, training programs, courses, and post-secondary diplomas were inventoried to identify the competencies that were included. When possible, journal articles or the published listings of training programs, such as the Serene-RISC Cybersecurity course directory were used (SERENE-RISC, 2020).
4. Existing ontologies and the materials gathered in the first steps that demonstrated potential value were gradually integrated into a mind map with sufficient details to further determine value and usefulness. As the mind map grew, the respective relationships between the information and the concepts were added.
5. Early discussions among the researchers and the stakeholders were held to assess the relevance of this new information into the model and continued throughout the study. For this purpose, the mind map was regarded to be a useful tool.
6. The mind map was updated to reflect the comments and the growing consensus. Including information into the mind map data for which there was an agreement that it could provide value and meaning. Those for which there was no agreement were removed from the mind map, and no additional information was added. In cases where partial information was considered meaningful, the map remained unchanged at this stage, understanding that this could be further refined later.
7. The Human Resource Department of the participating organization provided internal job descriptions and postings. This information was integrated into the

mind map to provide additional information. In particular, it provided insights into cybersecurity competencies required and valued by the organization when hiring new employees into cybersecurity positions.

8. Internal shared network drives, intranet, Microsoft SharePoint pages, and other sources were searched for additional information and previous work and reports on the topic of cybersecurity competencies and job descriptions that could be added to the mind map.
9. Saved searches in the job posting websites that are popular for advertising cybersecurity job openings in the local market were created to gather information on the requirements of organizations.
10. Discussions with stakeholders and sharing the mind map was ongoing as it helped building consensus. Alterations were made as required to maintain the consensus.
11. From the information gathered, a semi-structured interview guide was developed, and this concluded with this top-down analysis.

At this stage in this study, the bottom-up ontology strategy was used, starting with the need of the users. Semi-structured interviews were conducted in the first BIE cycle with the managers of all the different specialty areas in the cybersecurity department who were available and consented to participate in the study. This corresponded to eight individuals (**n = 8**). Using the semi-structure interview guide as a starting point, discussions focused on the divergent work roles and tasks performed in those roles by the individual team members. From there, information was obtained on the competencies, training, education, and certifications required to successfully perform these roles. Finally, issues related to talent movement were discussed within the different cybersecurity groups, recruitment, and retention. As competency and talent management are important shared concerns of the participating organization, this facilitated obtaining the consent of all mid-level cybersecurity managers to participate in this study. As the interviews progressed, the following steps were used to pursue data collection for this study and increase the understanding of the subject area.

5.1.3 Step 2: Data collection and integration

1. Following the interviews, notes and information gathered were incorporated into the concept map.
2. Appropriate changes were made to the map to clarify or update it in accordance with the new information.
3. As more interviews were conducted, the information was again updated until all the interviews were completed.
4. The results were shared as discussions continued with stakeholders and the mind map enabled building consensus. Changes were made as required to maintain the consensus.
5. An initial cybersecurity competency model was developed.
6. Material was prepared for a group workshop to be conducted with the senior cybersecurity management of the organization.

In total, three workshops were held, with a convenience sample of 34 (**n = 34**) participants that were available and agreed to participate from 50 staff members who were invited from the cybersecurity department that had a total of about 240 staff at the time these took place. To get a good representation of the different areas, a few individuals from every team were invited, which is how the number 50 was achieved. A first group workshop was done with 12 participants (**n = 12**), all in leadership positions in the cybersecurity department. This included the CISO, senior directors, directors, and managers. For this workshop, a brainstorming technique based on design thinking was used, which works well to generate group discussions on a particular topic. This was selected as the facilitator of the activity was very familiar with this technique. Furthermore, a use case was presented that included a persona, and the story of a young student, a 19-year-old male, residing in an affluent suburb, starting university in the Fall, wanting to join the organization upon graduation, and asking the participants to help them decide. To help develop empathy for the persona, a stock photography picture of the student was found, which was shared with the participants at the workshop. The use case was built to provide a personal story that made sense and was a likely situation for the participants: a young person in his extended network, named Philippe, who considered a

job in their field and came to them for advice on career choices and post-secondary options. The 2.5-hour workshop followed the steps enumerated below.

5.1.4 Step 3: Workshop

1. The plan and modus operandi of the workshop were explained.
2. The use case was presented to all the participants.
3. Each participant was asked to have an individual reflection on the possible cybersecurity job opportunities in the organization for Philippe and write down their proposals using the post-it notes provided to them.
4. After 20 minutes, when no significant activity could be observed by the facilitator and the participants began conversing with each other, the facilitator started to go around and ask the participants to name their ideas and post them on a very large whiteboard on the wall, grouping similar ideas, fostering discussions, and generating new ideas. This continued until all the post-it notes were up on the wall. After a few iterations, similar ideas appeared; the participants were then asked that if a participant presented an idea that they also had, to be added to the wall in the same area.
5. The next phase was performed in three groups of four participants, and the individuals sitting nearby were also grouped.
6. The groups were asked to discuss among themselves for approximately 20 minutes to identify the possible groupings of the probable cybersecurity job opportunities for Philippe into cybersecurity job categories. They utilized large post-it notes for documenting this.
7. Each group presented the outcome of their discussions and added their large post-it notes.
8. This was followed by a consensus-building discussion among all the participants to explain or defend their proposal, which was led by the facilitator. After nearly 30 minutes, a consensus emerged on the two categories of cybersecurity positions: business and technical.
9. This workshop concluded with a group discussion on possible cybersecurity job opportunities in the organization for Philippe.

Following the first workshop, the information gathered was, here again, integrated into the mind map and shared with the stakeholders. The initial model was updated in accordance with what was learned. The emerging consensus was that there were two main categories of cybersecurity positions with some common cybersecurity competency requirements and some that were distinct. The first category, cybersecurity business, was composed of individuals who are experts of the business aspects, with more of a strategic outlook and understand how cybersecurity provides value to the organization and its stakeholders. The second category, cybersecurity technical, included IT workers in various technical aspects, hardware, software, and cybersecurity tools. However, considering this, the participants in the workshops also acknowledged that business experts needed some technical knowledge and know-how, while the technical experts also needed to understand the business before gaining competence in their job. There is a continuum of cybersecurity competency from business roles to technical roles, with many work roles in between. This continuum leads to identify specialties within the two categories. It was suggested that there were specialty areas that should be identified in the model. Accordingly, two additional workshops were held: one for business category workers and the other for technical workers. These workshops followed the procedure described above, but each focused on its category. One workshop included staff in various technical cybersecurity roles (**n = 12**), and the other workshop included staff from the business-oriented cybersecurity roles (**n = 10**). Here, as described previously, a convenience sample of participants who were available and agreed to participate from staff members invited from the cybersecurity department that had a total of about 240 staff at the time these took place. To get a good representation of the different areas, a few individuals from every team in the desired areas (business or technical) were invited. The first group workshop was done with 12 participants (**n = 12**), all in leadership positions in the cybersecurity department. Following the two additional workshops, the mind map and model were updated and shared with the stakeholders. The workshop documents are shown in Appendix K.

There was also some interesting information about cybersecurity competency, work roles, and talent management that was mentioned by participants in the workshops. While some

will be useful in this study, some of the information may plausibly be used in future studies. Some participants at a management level in the organization indicated that academic degrees and diplomas, such as a bachelor or masters in cybersecurity, are no longer considered a guarantee of competence or even a condition of employment. Nonacademic cybersecurity certifications, such as CISSP, Security+, or CISA, can be good indicators of a minimum level of knowledge that can contribute to indicating a minimal competency. Another useful measure of competency is the participation, and more particularly, the victory in cybersecurity competitions, such as capture-the-flag (CTF) events. This is also viewed as a very good indicator of advanced technical cybersecurity competencies. Competency demonstration, such as badges, micro-accreditation, or recognition of prior learning and skills in accelerated training programs may be interesting avenues for the organization. As well, life or work experience in cybersecurity and related fields is an essential component of competency. The participants in the workshops mentioned that knowledge of the company's strategic business activities and the banking industry is an important requirement in considering the competency of a cybersecurity worker. Many participants, both managers and workers, expressed in the workshop the importance of soft skills for cybersecurity workers. Personal skills, interpersonal skills, and emotional intelligence are viewed as important requirements of cybersecurity workers. The next sections of this dissertation will present the resultant data, and how this information is integrated into the cybersecurity competency ontology.

5.1.5 Cybersecurity work roles identification

On the basis of a thorough literature review, an initial model was developed that included six cybersecurity work roles. This model was used to help us understand not only the requirements but also design some of the original data collection instruments. However, after the interviews and, most importantly, the workshops, done in the first BIE cycle, a more elegant model was proposed. The outcome based on the data from all three workshops was the identification of the two main categories of cybersecurity work roles with eight specialty areas (Figure 1). A ninth specialty area was identified as managers, crossing over management KSAs with the different specialty areas. However, to focus on

workers' competencies, the management category was not included in this study. Moreover, within each specialty, there were varied levels of competency requirements identified, depending on the role. For instance, there could be a justified need for a range from junior, or a specialist entering a work role, all the way to a more senior, experienced, and highly competent worker.

The results were consolidated into an interim report and a PowerPoint presentation that was presented at an online conference to all the internal stakeholders, followed by discussions with the stakeholders in the first BIE cycle. As needed, changes were made to the mind map to integrate new information emerging from the conference and maintain a consensus on the results. Working from the mind map and the data collected in the workshops, a UML model was created by the researchers, as previously shown in Figure 4, to define competency and schematic diagrams representing possible training and career paths into and within cybersecurity in accordance with the information gathered in the workshops. A worksheet describing the cybersecurity job categories was also prepared, as this requested by the organization, which later proved to be useful for discussions in the BIE cycles.

At this point in the study, all the elements of the cybersecurity competency ontology were present and could be integrated into a coherent ensemble, as described in Step 4.

5.1.6 Step 4: Integration of the data into a coherent ensemble

- a cybersecurity competency work role, as presented in the UML format in Figure 1, with two categories and nine specialties that were identified in the workshops;
- a comprehensive list of cybersecurity tasks and KSAs from the NIST NICE framework (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017) and other sources with additional information on KSA;
- an inventory of cybersecurity certifications;
- an inventory of cybersecurity training programs, courses, and post-secondary diplomas;

- a mind map depicting the consensus of the stakeholders and researchers on how these different elements are connected at this point in this study, as presented in Appendix J.

5.1.7 Cybersecurity job posting

Saved searches were created in the job posting websites using keywords such as cybersecurity, information security, Canada. This was done manually by creating saved searches, as it was sufficient for the intended use. If there were no confidentiality issues, this would have been done using job postings and CVs of current staff in the organization. Other possible ways to gather the data are web scraping or buying databases from data providers. These options were not possible, as this could have created ethical and legal concerns. Websites popular for advertising cybersecurity job openings in the local market, such as LinkedIn and Indeed, were used to gather information on the requirements of organizations. The motivation behind creating the saved searches, but more so saving the results, is that these job postings represent the actual requirements of organizations for cybersecurity professionals. This also furnishes meaningful insights into the competency requirements and how these competencies are described and named in a particular context. It was initially believed that this could likely complement documented frameworks, such as the ones presented in the following sections. The process followed to collect the job postings is described in step 5.

5.1.8 Step 5: Collection of the job posting data

1. First, job posting advertising websites that are being used by the participant organization to publish cybersecurity positions in the market were identified, and these are Indeed (https://emplois.ca.indeed.com/) and LinkedIn (https://www.linkedin.com/jobs/).
2. Then saved searches were created on the websites using “cybersecurity” and “information security” as search themes and used an email alert to send the regular alerts of new postings. The location setting was Canada.

3. Once an alert was received, posted positions were reviewed and confirmed that a particular posting was relevant. To be relevant for potential inclusion in the ontology, the job posting had to include cybersecurity as a principal requirement or component of the work role that was being recruited.
4. Confirmed postings were printed as PDF files and saved in a directory to build a repository that was later loaded into a database as will be explained later.

5.1.9 Integrating the results

Following the workshops and using the updated mind map and model that were shared and discussed with the participants in the first BIE cycle, this study demonstrated that all the elements required for constructing the ontology were available. The resultant data, presented in the previous section, could now be integrated into the cybersecurity competency ontology. Accordingly, WebProtégé, available online (https://webprotege.stanford.edu/) was used. WebProtégé is a web-based ontology development environment that was developed by the Protege team in the Biomedical Informatics Research Group at Stanford University. There is also a desktop, standalone version that is available for Windows and macOS platforms. WebProtégé supports the OWL 2 Web Ontology Language and data formats used for ontology upload and download. It does not have SPARQL query support, only Protégé. However, OWL data can be used and exchanged on both the web and desktop versions. Once an account was created on the WebProtégé site, Step 6 was followed.

5.1.10 Step 6: Integration of the model into the ontology

1. Using the Create New Project button, a new project named Cybersecurity Competency was created. The language was selected as en-US (US English), and a description of the study was added.
2. The integration of the data from the mind map as classes in the ontology, started with the central concept. Two branches, related to the research methodology and possible solutions proposed by the researchers, were not integrated as they were more related to the empirical research aspects of this study and not the ontology.

3. Working along the different branches, the class hierarchy with the classes and divergent subclasses can be added until all the concepts were integrated.
4. When possible, multiple classes were created in bulk as WebProtégé allows for multiple entries in the class entry window.
5. It is possible to copy concepts in the mind map and paste them as multiple classes in WebProtégé. This makes it easy to reorganize the classes using drag and drop features to put them in the correct tree structure, which saved time and keystrokes during the creation of the ontology.
6. When possible, WebProtégé annotations were added when the objects were created. Otherwise, this was done later in the process. RDS: label was created when the class was created, and the following labels were added, as the ontology was being populated to complement the information:
    a. RDS: comment
    b. RDS: seeAlso
7. When possible, WebProtégé relationships were added when the objects were created to help unlock the meaning of the ontology. Otherwise, this was done later in the process. The following ontological relationships were added:
    a. isIssuedBy: indicating the body that issues a degree
    b. performsTask: tasks performed by an actor in a work role
    c. requiredKnowledge: knowledge required by an actor in a work role
    d. requiredSkill: skill required by an actor in a work role
    e. requiredAbility: ability required by an actor in a work role
    f. requiredCertification: professional certification required by an actor in a work role
    g. requiredDegree: academic degree required by an actor in a work role
    h. desirableKnowledge: knowledge that is desirable for an actor in a work role
    i. desirableSkill: skill that is desirable for an actor in a work role
    j. desirableAbility: ability that is desirable for an actor in a work role
    k. desirableCertification: professional certification that is desirable for an actor in a work role

l. desirableDegree: academic degree that is desirable for an actor in a work role

The result from these steps is the completed cybersecurity competency ontology engineering process. Following this process, the ontology can be created using the selected ontology tool, WebProtégé. Using this tool, all the data was identified to be included in the ontology itself. This is a gradual process that may take time, depending on the depth required. When this was done, this study was working with the underlying hypothesis that the ontology would continuously evolve as the competencies required of workers in an evolving domain of knowledge, and in this case, cybersecurity would evolve. However, from the ontology engineering point of view, the process was completed at this stage, as the first BIE cycle was completed.

The ontology was then ready to be used in the subsequent phases of this study, to validate the ontology in the second BIE cycle and test its usefulness as a management tool for financial organizations in the third BIE cycle. It was then possible to continue to populate the ontology with additional data as it emerged through further research. In general, adding further information is an ongoing process that could continuously add to the value of the ontology as an efficacious tool for organizations.

5.1.11 Review of the design process

Creating the ontology is an iterative process, and initially, several mistakes were made while creating the ontology by thinking that it would be a simple sequential process. It was necessary to go back and forth a few times to revise the engineering strategy. What is presented here is the result of this iterative process. The proposed cybersecurity competency ontology design approach has six steps.

- Step 1: Gathering of the initial data
- Step 2: Data collection and integration
- Step 3: Workshop
- Step 4: Integration of the data into a coherent ensemble
- Step 5: Collection of the job posting data

- Step 6: Integration of the model into the ontology

Each step has additional steps that are described in this section. Following this approach, it was possible to develop a cybersecurity competency ontology that can be used by a financial institution. The initial data suggests that helping the organization in the areas of talent and training management is useful, but this will be discussed further in a subsequent section.

While the study intended to reuse previous ontologies, it was not possible to identify existing ontologies that could be reused. In some cases, where some materials could be found, the links to the information had not been updated, as the information was no longer available online. This did not imply that they do not exist or are available, but none could be located. Perhaps, there is work required to provide the up-to-date directories of available materials.

Having a shared agreement with the different stakeholders on the ontology and its contents is a time-consuming process. Consensus-building activities in general take time, and this study was no exception. Several back-and-forth discussions were held with the stakeholders to obtain a shared agreement. In this study, as mind maps were used as a tool and as the participants in the target organizations were very familiar with the use of mind maps, it became a highly helpful tool. However, the were many iterations of the mind map in the study, thereby leading to the point when the ontology could start to be built.

Bearing in mind that this section presents the ontology engineering phase of a larger study, and that further information will be presented in other sections on the contents and the validation of the ontology, the results and discussion on the process of creating the ontology are presented here.

A cybersecurity ontology that not only focuses on the technical knowledge of professionals but also considers the importance of social and organizational traits would enable the future cyber workforce to be effective in their field. The approach for constructing an ontology representing the cybersecurity workforce has been described.

The ontology must confront the intricacy of the cyber system and adapt to the complexity of the domain. Overall, the approach should consider the combination of technical and social behavior skills on the network (Fontenele & Sun, 2016). National frameworks, such as the NICE, provide the base for the ontology (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017).

Cybersecurity workers have many duties and responsibilities that emphasize the execution of divergent tasks, detect intrusions, and attain the desired outcomes. The minimum levels of analytical and technical rigor are essential skills to ensure that the provided solutions are practicable for employers and organizations. By solely focusing on the technical knowledge and failing to emphasize organizational traits, various security, knowledge, and retention gaps are spawned that make it inadequate to develop an ontology that represents an effective cybersecurity workforce. Furthermore, a cybersecurity ontology should represent the pattern matching and good mental flexibility abilities and the situational awareness of the workforce.

Given that the cyber system is complex, the ontology represents a teamwork, as cybersecurity professionals will need to work in teams with diverse talents. A cybersecurity ontology represents the future association with the value system, which prevents professionals from exploiting the lack of expertise by their employers. Besides the social and technological skills, a cybersecurity ontology represents trustworthiness and reliability. In addition, a cybersecurity ontology representing the workforce should account for the diversity of organizations contained within the cyber network. The key traits represented by a cybersecurity ontology should include systemic thinking, teamwork, technical and social traits, civic duty, continued learning, and communication.

5.2 Ontology alignment

This section presents the mapping of a cybersecurity competency ontology onto cybersecurity competency reference models and the requirements of a large Canadian financial institution done to test the internal validity of the ontology. The ontology was necessary to provide the data required for a human resource management tool to assist

financial sector organizations in adequately managing IT security and risks, addressing a major concern in matching competent cybersecurity workers with cybersecurity work roles. Initial data gathered in this study was used to populate the WebProtégé ontology, as described in this section.

This section addresses the following research question:

> **How can the structure and contents of an OWL ontology represent the core competencies of the cybersecurity domain?**

Following the cybersecurity competency ontology design approach developed in the research study, the artifact, the ontology constructed using OWL in the Stanford Protégé application, was produced. This artifact is based on the cybersecurity requirements of a large Canadian financial institution, aligned with the NCWF (Newhouse et al., 2017a).

Answering this research question enabled mapping the cybersecurity ontology with cybersecurity competency reference models and the requirements of financial institutions in a manner that can provide a practical application of the ontology for human resource management.

5.2.1 Cybersecurity ontology mapping process

Mapping the cybersecurity competency ontology with cybersecurity competency reference models and the requirements of a large Canadian financial institution began by creating the ontology and defining the data collection process for the contents. These are described in Section 9. Following the ADR methodology, BIE cycles were utilized to gradually build the ontology design approach and the cybersecurity competency ontology. An iterative approach to building the ontology permitted the researchers to gradually gain a better understanding of ontology engineering and cybersecurity competencies. The process can be summarized in six steps, with subprocesses that go into further detail. The main steps have been explained in previous sections.

The said process yielded the core components of the cybersecurity competency ontology. However, the initial data from the competency models and the literature, the data

gathered in the field, and data from the cybersecurity stakeholder workshops needed to be aligned with the actual work roles and the in vivo requirements of the Canadian financial institution. This alignment is the topic of this section.

5.2.2 Cybersecurity ontology alignment

A strategy for alignment was required in the study, as multiple elements need to be in some form of equilibrium. This section presents in detail the process for a single actor and a single role. In reality, for the study as a whole, the process needs to be repeated for all seven categories, 33 specialties, 52 work roles, 1,007 tasks, 630 knowledge elements, 374 skills, and 176 abilities of the NCWF (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017). Additionally, these were aligned with the two categories (cybersecurity business and cybersecurity technical) and the nine specialty areas of the model. In the cybersecurity business category, the study identified the following specialties from the analysis of the data collected during the study:

1. Cybersecurity manager
2. Cybersecurity analyst
3. Cybersecurity advisor
4. Cybersecurity awareness and education specialist

In the cybersecurity technical category, the following specialties were identified:

5. Offensive cybersecurity specialist
6. Defensive cybersecurity specialist
7. Cybersecurity exploitation specialist
8. Cybersecurity systems architect
9. Physical security specialist

It might come to the attention of the reader that there is a substantial difference between the seven categories, 33 specialties, and 52 work roles as compared to the two categories and nine specialties of the model. From this study, the consensus emerged that the NCWF is much more complex than what the target organization could reasonably use in the field,

suggesting a more concise model may be more applicable. This more elegant model could achieve better results as a contributing factor to risk reduction. Furthermore, there is some interest among cybersecurity managers within the organization in adding another layer of expertise within the specialties, thus increasing the level of detail and the number of identified work roles and bringing the proposed model closer to the NCWF in the number of work roles. This will be investigated in future research.

5.2.2.1 Actors have roles

This process was eventually repeated for all actors in the organization. However, before creating the other actors, the CISO actor was further defined using relationships to connect it to the tasks performed by the CISO in accordance with the reference models and the collected data, including interview data.

5.2.2.2 Roles involving tasks, knowledge, skills, and abilities

The next step involved linking the KSA of the NCWF to the EXL role. This was done in a similar fashion to the tasks shown in Table 2. A noteworthy difference was the nature of the relationship created for this content, as presented in Table 3. In the same manner, the required skills for the EXL role from the NCWF were captured and integrated into the ontology, as summarized in Table 4. Likewise, the abilities attached to the EXL role were captured and integrated into the ontology (Table 5).

With these tasks, knowledge, skills, and abilities integrated into the ontology, the basic elements of the NCWF for the EXL role were integrated. As mentioned, this was needed to be repeated for all 52 work roles in the 33 specialties. They were then connected to the 1,007 tasks, 630 knowledge elements, 374 skills, and 176 abilities of the NCWF. It should be noted that individual tasks and KSA were shared by multiple work roles.

5.2.2.3 Know-how for the role

Using know-how as an element of competency is one of the contributions of the proposed cybersecurity competency framework. Know-what and know-how-to-be were also used as competency elements. Starting with know-how, the required know-how for all

cybersecurity roles in the financial organization were determined from multiple sources. First, by looking at all the NCWF roles corresponding to the organizational categories, it was possible to identify the most common items for each role (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017). These were then combined with data from the interviews, workshops, and discussions with stakeholders during the consensus-building process. Discussions were held with industry stakeholders as well. Finally, with the data integrated into the mind map, a consensus was built among different stakeholders on the know-how to include.

This led to the identification of two principal categories of cybersecurity work roles: the cybersecurity business role and the cybersecurity technical role. In the cybersecurity business role category, four specialties were identified: cybersecurity manager, cybersecurity analyst, cybersecurity advisor, and awareness specialist. The cybersecurity technical category has five specialties: offensive cybersecurity, defensive cybersecurity, IT security exploitation specialist, cybersecurity architect, and physical security specialist. Some KSA were found to be common to all roles and were defined as cybersecurity baseline competencies, as shown in Table 6.

Accordingly, a baseline cybersecurity know-how was determined that would be required for all cybersecurity workers and a minimal level of competency that included these six knowledge elements. To these, a know-how element (BKH0007) was added, and subsequently a second know-how element (BKH0023) was added in accordance with additional data and feedback gathered during the consensus building. These can be seen as the core competencies common to all cybersecurity workers. These competency elements were defined as the baseline know-how or BKH, with a sequential number for coding. The resulting competency elements are presented in Table 7. The know-how elements could then be added to the cybersecurity role class in the ontology as required knowledge (Table 8).

The common tasks and KSA for the entire cybersecurity business category were then added, starting with the NICE roles identified with this category (Table 9). This led to identifying the baseline cybersecurity know-how that would be required from all

cybersecurity business workers, as along with a minimal level of competency. These competency elements were defined as the baseline know-how or BKH, with a sequential number for coding, as was done for the previous more general category. This is presented in Table 10. The relationships were then added to each appropriate subclass; those for cybersecurity business are shown in Table 11. The same was done for the cybersecurity technical subclass, which resulted in the identification of common know-how required by this group of workers (Table 12), and added relationships, as was done for the other category.

More specifically, the CISO role, for example, corresponded to the cybersecurity manager specialty that was identified in this study. As per the data provided in the NCWF, eight roles were identified for this specialty and are enumerated in Table 13. These roles combine 242 KSAs. The KSAs that were the most prevalent were identified, common to four or more roles (Table 14).

From there, a more neutral know-how description was proposed that would be compatible with the requirements of this study. As mentioned previously, know-how is knowing how to do something. It can be acquired via education and requires high-level problem-solving abilities. This was proposed to the stakeholders and discussed to establish a consensus. The result was then added to the ontology, as shown in Table 15. Note that elements were numbered to facilitate their identification, using the acronym AKH for acquired know-how, followed by a sequential numerical value starting at 0001.

Know-what could then be added. Know-what is the ability of competent individuals who demonstrate practical knowledge of the work, tasks, techniques, business, and ecosystem of an organization. When individuals have know-what, they possess in-depth mastery of how their domain of competency operates as a coherent system. Effective performance and career efficiency can also be linked to emotional, social, and cognitive intelligence of individuals, or know-how-to-be (Bacigalupo et al., 2016; Boyatzis, 2008, 2008; Man et al., 2002a). These elements were numbered to facilitate their identification using the acronym DKW (for acquired domain know-what), followed by a sequential numerical value starting at 0001 (Table 16).

Finally, know-how-to-be for the cybersecurity manager specialty was proposed. Know-how-to-be is a characteristic of competent individuals with high emotional and human relations abilities, mental and physical capacities, basic sense attitudes, strong value systems, and behaviors compatible with the organization's culture and the dominant socio-cultural values of the various internal and external stakeholders, including abilities to interact with colleagues (Bacigalupo et al., 2016; Boyatzis, 2008, 2008; Draksler & Širec, 2018b; Man et al., 2002a; McClelland, 1973; Mitchelmore & Rowley, 2010). These were numbered to facilitate their identification, using an acronym and a sequential number starting at 0001. Several categories of know-how-to-be were created, as per Table 17.

Once these categories were created and the different competency elements of this category, or subclasses, were captured in the ontology, it was possible to add relationships to appropriate roles in the ontology. This information came from the literature review, the NCWF, and from the interviews after the consensus-building efforts, as with the other competencies described previously. In this case, relationships were added, as presented in Table 18, to the cybersecurity manager specialty that was identified in the study.

5.2.2.4 Competency acquisition

As was mentioned previously, competency can be acquired in many ways. The most common manner of acquiring a competency is by training or education. For example, the course DAT813 from the University of Sherbrooke, located in Sherbrooke (Québec, Canada) in the Governance, Audit, and Information Security program offers cybersecurity competency acquisition as learning outcomes as described on the program's public website. This information was added to the ontology, as presented in Table 19. From there, it became possible to update the requiredKnowHow AKH0003, knowledge of risk management practices, to include the information summarized in Table 20.

Hence, to better understand the approach, it can be described as such: A CISO, who is a cybersecurity manager, in the cybersecurity business role category, must have the required acquired knowledge of risk management practices (AKH0003). This knowledge could be acquired by completing the University of Sherbrooke course DAT813, which includes learning outcomes for (LO0004) risk assessment, (LO0005) risk management, and (LO0008) risk management methodologies that are included in AKH0003.

Again here, a single example of how this ontology can be used is illustrated in this dissertation. It is necessary to identify all learning outcomes for the different cybersecurity and information technology courses to be included in the completed ontology. Furthermore, these learning outcomes must be linked to the know-how, know-what, and know-how-to-be. Thus, as they all become connected, the ontology becomes a useful tool to help organizations fulfill these requirements.

5.2.3 Results

Once again, the reader is reminded that the data for a single work role is presented, the CISO, which was the cybersecurity manager specialty in this study. However, the process was the same for all NICE roles in the two principal categories of cybersecurity work roles. This study aimed to repeat what was described here to complete the creation of the cybersecurity competency ontology.

The more time-consuming portion of the study and activities described in this section was data gathering. Once the data was gathered and consensus-building on the elements and the meaning of competency for a particular work role was achieved, it became a tedious, repetitive, data-entry process to capture all information accurately in the ontology. Potential technical challenges of using WebProtégé made it necessary at times to ensure that the data is properly saved. There was also a need to ensure that regular backups were made. Once these were completed, a fully functional ontology could be used.

5.2.4 Review of the alignment process

At this stage, the study was aimed to describe the structure and contents of an OWL ontology representing the core competencies of the cybersecurity domain in accordance with the cybersecurity requirements of a large Canadian financial institution aligned with the NCWF (Newhouse et al., 2017a) with the additional material and data resulting from the study, providing clarification and domain expertise that applies to the financial sector. More specifically, structuring the NCWF work roles into specialty areas and top -level domains (Business and Technical) mentioned in this section. The improvement to the NCWF is the addition of the two top-level cybersecurity domains, the cybersecurity business role and the cybersecurity technical role. Each specialty area includes NCWF work roles, as presented in Appendix N.

The ontology was created using OWL in the application Stanford Protégé, following the design approach presented in a previous section. This allowed the researchers to map the cybersecurity ontology onto cybersecurity competency reference models and the requirements of financial institutions in a way that could provide a practical use of the ontology for cybersecurity talent management, a current concern for financial institutions in Canada and elsewhere in the world. The next section explores the use of the ontology to answer organizational concerns and, eventually, serve as the basis for a machine learning algorithm that may help automate some of the tasks related to competency management, talent management, and recruitment of new talent.

This research demonstrated that the creation of the ontology is feasible and would appear to produce the expected organizational benefits. This was later validated, as will be described now.

6 Cybersecurity competency ontology internal validity test

This section presents the validation of a cybersecurity competency ontology as a reliable competency and talent management tool for Canadian financial organizations. It describes systematically the process and the various steps taken to ensure that the ontology allowed organizations to answer queries concerning the alignment of competencies and the associated key elements with the tasks and the work roles of individuals in the organizations.

Information security in a connected world, customarily referred to as cybersecurity, is a prevailing and growing concern and a major challenge for all organizations (Banham, 2017; Callen-Naviglia & James, 2018; Cleveland & Spangler, 2018; Cleveland & Cleveland, 2018; Hiller & Russell, 2013; Schatz et al., 2017; van Kessel, 2018). In the last couple of years, the world has witnessed a surge in cybercrimes, involving IT as the target as well as an instrument. As most organizations rely on IT to create and maintain a competitive advantage, they have a strategic requirement for governance, risk, and compliance (GRC) management programs. For the researchers involved in this study, this is a crucial area requiring thorough investigation. Of the multifarious dimensions of the GRC domain, the issues that were found noteworthy were the role of actors, the individuals in the cybersecurity workforce in the organizations, and their competency in the roles that the organization requires them to play to maintain IT-related risks at an acceptable level.

Where technologies and systems are dependent on a finite number of variables, individual actors are social and cultural animals that evolve, change, and can be influenced in unpredictable ways. They are often described in the cybersecurity field as the greatest vulnerability in information systems, typically referred to as the “human factor.” Nonetheless, such individuals in the organizations are also considered one of their most valuable assets. This was a contributing factor to support cybersecurity training, awareness, and educations programs as the perfect instruments to mitigate the negative effects of the human factor on GRC. However, to utilize this instrument optimally, it is required to understand the way a cybersecurity competency is created, gauged,

maintained, and nurtured. The ontology was designed by effectively modeling these competencies and creating a management tool proficient in matching the competency requirements of the organization with the capabilities of the actors in place and a strategy to fill the gaps when compared to GRC frameworks and the best practices.

This section describes how the ontology was validated and tested in a large Canadian financial institution. The ontology that was designed and built can allow organizations to identify the most adequate individual workers for their different cybersecurity roles. This section describes how queries in the ontology can be used to answer the real-world organizational questions and addresses the following research question: **What is the level of validity of the ontology in accurately representing the cybersecurity domain, and to what extent is it effective as a talent management decision tool?**

In essence, answering this research question facilitates using the cybersecurity ontology, constructed with cybersecurity competency reference models and the requirements of financial institutions, in a way that can make it viable for talent and competency management, contributing significantly to enterprise risk management.

6.1 Ontology design

Following the ADR methodology, BIE cycles were employed to build the cybersecurity competency ontology. The process was initiated in the first BIE cycle by building a mind map, which is presented in Appendix J. This involved semi-structured interviews, followed by workshop sessions with cybersecurity managers and finally group discussions with cybersecurity workers in the organization. Throughout this process, the initially collected data was integrated as the mind map artifact, developing gradually, and was subsequently shared and discussed as it evolved into a new artifact. In the next iteration, the mind map was utilized to create UML models of what eventually became the objects constituting the ontology. The models were used in the validation of the ontology, gradually evolving from the mind map and the UML models as they evolved in the multiple research cycles. These models made it possible to visualize the data and make sense of how this was all connected as the understanding of the research problem

grew. Once it was felt that a sufficient level of information had been gathered, the information was integrated into Protégé OWL as the basis of the ontology and the mind map was no longer used. From that point, the ontology in WebProtégé was used. Later, where it became possible to start to perform SPARQL queries, the desktop version, called Protégé, was used. Then, once all the data was integrated into the ontology the work continued using Stardog, as described later in this dissertation.

6.2 Ontology validation requirements

The cybersecurity competency ontology is useful to organizations as it can potentially enable them to resolve major difficulties in matching individual actors, work roles, and the competencies required in the execution of the tasks that ought to be performed to support risk management activities adequately. Actors, as individual workers, have roles in the organizations and require competencies to perform efficiently in their roles. This is also compliance, regulatory, legal or contractual requirement in the case of financial institutions and in various other industries. In addition, principles of good governance, best practices, and due diligence are the other crucial reasons that substantiate the need to find the optimal fit between cybersecurity workers' competencies in supporting business technologies and the related business processes supporting strategic business units. The quality of an ontology denotes how accurately it provides a suitable description of the problem domain, while being syntactically correct, precise, and semantically suitable (Vyšniauskas et al., 2012). Thus, it was required to affirm if the ontology that was developed addressed the problem domain. Once it was ascertained that the ontology met the fundamental semantic criterion for an ontology, the goal of validation was to verify if the ontology could effectively represent and contribute to the optimization of the cybersecurity actor-role-competency fit.

6.3 Using the ontology as a query tool with SPARQL queries

The next step in the validation of the ontology required the use of queries. The goal was to determine if the OWL ontology could be utilized to answer specific management questions. In the initial phases of the research study, three questions were created for the

validation process in the form of queries, described in a phased manner in the next sections of this section using the SPARQL query language. At this stage, a problem with one element in the ontology was identified.

As object properties had been defined in WebProtégé, it was also necessary to define data properties associated with individuals in the ontology. In the cybersecurity competencies ontology, individuals in the ontology refer to specific work roles of workers. It was necessary to add these to the ontology, as displayed in Table 6. Furthermore, it was required to match the individuals to their tasks, and to the know-how, know-what, and know-how-to-be that were required to accomplish the tasks. In the ontology, these were already present as relationships associated with the classes. As depicted in Table 2, this relationship can be established with the individuals by using data properties that were created. This was rather simple to add as it was simple to edit the ontology OWL file in XML format with a text editor and copy the data from the relationships to the data properties section of the OWL file, which can be found at the end of the file. A minor edit was required to get them in the correct format. For future studies, using these data properties when designing the ontology would be recommended.

6.4 Applicability scenarios

To test the usefulness of the ontology as a query tool, as discussed in the previous section, risk scenarios that represent likely situations in a financial institution were used in accordance with the MITRE ATT&CK framework. This became hypothesis H3 that the ontology would allow organizations to match work roles to risk scenarios. These are presented here:

1. A cybersecurity analyst in a governance role supporting risk assessment and management advisory. This is used in queries 1, 5, 9, and 16.
2. A system security analyst supporting vulnerability mitigation and cyber defense role following a security incident. This is used in queries 2, 6, 10, 17, and scenario 1. Specifically, the incident considered in scenario 1 is:
    - An adversary exploiting software vulnerabilities for privilege escalation as described in the MITRE ATT&CK framework https://attack.mitre.org/techniques/T1486/ done in order to extract

monetary compensation from a victim in exchange for decryption or a decryption key, commonly referred to as a ransomware attack.

- by LockerGoga https://attack.mitre.org/software/S0372/
- which required application isolation and sandboxing https://attack.mitre.org/mitigations/M1048/ and software update https://attack.mitre.org/mitigations/M1051/

3. A red team analyst performing penetration testing and vulnerability identification. This is used in queries 3, 7, 11, and 18.
4. A blue team analyst responding to a security incident. This is used in queries 4, 8, 12, 19, and scenario 2. Specifically, the incident considered in this scenario is:
    - Data encryption for impact as described in the MITRE ATT&CK Framework https://attack.mitre.org/techniques/T1068/
    - by APT28 https://attack.mitre.org/groups/G0007/
    - which required activating the incident management and disaster recovery plans as well as data backup https://attack.mitre.org/mitigations/M1053/ and vulnerability scanning https://attack.mitre.org/mitigations/M1016/

6.5 Preparing for the use of queries

The queries were used in two different ways; first, in Protégé, as there was no support for SPARQL queries directly in WebProtégé. The steps for executing queries on protégé are as follows:

1. Open the ontology in protégé.
2. Select the SPARQL Query tab
3. If the SPARQL query tab is not displayed, go to Windows - Tabs - SPARQL Query
4. Click on SPARQL Query Tab to access the query editor.
5. Type the desired query in the window.
6. Press **Execute**

Apache Jena Fuseki server was installed and configured as a precursor to a prospective management information system and to validate the queries in a second environment. The procedure to install Apache Jena on macOS version 11.0.1 is as follows:

1. Press **Command** + **Space** and type **Terminal** and press **enter**.
2. Copy and paste the following command in the Terminal app:
   ruby -e "$(curl -fsSL
   https://raw.githubusercontent.com/Homebrew/install/master/install)"
   </dev/null 2> /dev/null
3. Press **enter** to execute

When the screen prompts the user for the password, Mac's user password is required to be entered to continue. When the user types the correct password, it does not get displayed on the screen, though the system accepts it. After typing the password, the user must press **Enter** and wait for the command to finish. Upon completion, the terminal is used to install Fuseki**:**

4. **Brew install fuseki**

The installed Fuseki can be started with a macOS terminal with the command: **fuseki start**. Subsequently, an SPARQL endpoint will run on http://localhost:3030/. Using a web browser, it is possible to load the ontology in Fuseki and execute SPARQL queries to test the queries in both environments, each possessing a copy of the same ontology extracted from WebProtégé to validate the results. The objective of this validation was to verify if management questions could be readily answered with the cybersecurity competency ontology.

The next step involves the execution of the queries and the analysis of the results from the queries. The researchers also performed and discussed queries that correspond to two risk scenarios based on the MITRE ATT&CK framework that describe specific incidents that could take place in a financial organization.

Query 1: What is the know-how, know-what, and know-how-to-be required for the cybersecurity analyst role?

Answering query 1 requires the Uniform Resource Identifier (URI) of the ontology being searched and assigning it a PREFIX, as shown in the first part of the query with the creation of the prefix CyberSecOnto. The Internationalized Resource Identifiers (IRI) for subject of the search was also required in this query, in addition to the work role under consideration, cybersecurity analyst, shown in the query by the variable ?analyst. In the Protégé ontology, this information can be found at: http://webprotege.stanford.edu/RDT0Br1r5F6LW8Aeq9bAcuh. Table 1 displays all the work roles that were identified in the study, making it simpler to perform queries in the ontology by using the IRI. This was done rather expeditiously by accessing the OWL file for the ontology with a text editor and by creating a table. This was repeated for the roles used in the other queries.

Table 1: Work roles

Individuals	IRI
Baseline	http://webprotege.stanford.edu/RBEc9x0WqhBg6Eh3IwsQfQ9
Business	http://webprotege.stanford.edu/R7lFVb3vIO2RMlQT3MeJVIp
Advisor	http://webprotege.stanford.edu/RpOWBN8QQUwdUWc5NUyI8t
Analyst	http://webprotege.stanford.edu/RDT0Br1r5F6LW8Aeq9bAcuh
Awareness specialist	http://webprotege.stanford.edu/RDyxNdbDkdXMODobLAGtE97

Manager	http://webprotege.stanford.edu/RGoO8ECZBJ5O7JwGI5Ohng
Technical	http://webprotege.stanford.edu/R8RCGG1zuRQdMPtbV3Zb3D
Blue team	http://webprotege.stanford.edu/R8HVb43o1m56KihScFoTPFx
Red team	http://webprotege.stanford.edu/RJznoCjLO8BK6O53PwZq4Y
Exploitation	http://webprotege.stanford.edu/RzLdQxNvODbOk4724Zh9UE
Security architect	http://webprotege.stanford.edu/RBpmMPf3tgGc0JcLW1hsMq2
Physical security	http://webprotege.stanford.edu/RByUeFEhyzCO0nPFg6wec58

In the WHERE statement of the query, the competency element that was being examined in the query was identified. In the first query, this was the know-how, know-what, and know-how-to-be, which in the ontology were described by hasKnowHow, hasKnowWhat, and hasKnowHowToBe, respectively. Table 2 encapsulates all the data properties that were used in this ontology. The SELECT statement uses the variable that was being searched, which was also assigned as KnowHow, KnowWhat, and KnowHowToBe.

Table 2: Data properties

Label	Description
doesTask	Identifies a task that is accomplished by an individual

hasAbility	Identifies an ability that an individual should have
hasKnowHow	Identifies know-how that an individual should have
hasKnowHowToBe	Identifies know-how-to-be that an individual should have
hasKnowWhat	Identifies know-what that an individual should have
hasKnowledge	Identifies knowledge that an individual should have
hasName	Identifies a role name for an individual
hasRole	Identifies a work role of an individual
hasSkill	Identifies a skill that an individual should have

The following query was used:

```
PREFIX CyberSecOnto:
<http://webprotege.stanford.edu/study/DpeZFWjjTksyFnolG1b70o#>

SELECT ?KnowHow ?KnowWhat ?knowHowToBe

WHERE {

    # find something that is of type `RDT0Br1r5F6LW8Aeq9bAcuh`

    ?analyst a
    <http://webprotege.stanford.edu/RDT0Br1r5F6LW8Aeq9bAcuh>.

    # which has a property `hasKnowHow` whose value is something

    ?analyst CyberSecOnto:hasKnowHow ?KnowHow.

    # which has a property `hasKnowWhat` whose value is something

    ?analyst CyberSecOnto:hasKnowWhat ?KnowWhat.

    # which has a property `hasKnowHowToBe` whose value is something

    ?analyst CyberSecOnto:hasKnowHowToBe ?knowHowToBe

    }
```

Once query 1 was executed, multiple results appeared to be correct, though difficult to interpret owing to a high number of elements. It was determined that the best strategy would be to run the entire query as three separate queries, for the know-how, know-what, and know-how-to-be, which provided the same results but in more manageable, which allowed to validate the results by manually comparing them to the actual data in the database and the data collected in the field. Furthermore, the DISTINCT instruction was added to eliminate duplicates. The results of the query are presented in Table 3.

The resultant query for know-how is as follows:

```
PREFIX CyberSecOnto:
<http://webprotege.stanford.edu/project/DpeZFWjjTksyFnolG1b70o#>

SELECT DISTINCT ?Know_How

WHERE {

    # find something that is of type `RDT0Br1r5F6LW8Aeq9bAcuh`
    ?analyst a
    <http://webprotege.stanford.edu/RDT0Br1r5F6LW8Aeq9bAcuh>.

    # which has a property `hasKnowHow` whose value is something

    ?analyst CyberSecOnto:hasKnowHow ?Know_How.

}
```

Table 3: Query 1 results

Query results
PEC0021 Following directions
BKH0014 Knowledge of management frameworks, best practices and standards used in the industry
WPC0013 Identifying the problem
BKH0005 Knowledge of the specific operational impacts of cybersecurity gaps

BKH0007 Knowledge of computer network concepts and protocols
WPC0015 Locating, gathering, and organizing relevant information
BKH0004 Knowledge of cyber threats and vulnerabilities
BKH0012 Knowledge of organizational supply chain and value chain management policies, requirements and procedures
BKH0001 Knowledge of cybersecurity and confidentiality principles
WPC0014 Implementing the solution
WPC0018 Troubleshooting and maintenance
BKH0010 Knowledge of the organization's information classification program
BKH0003 Knowledge of risk management processes
BKH0002 Knowledge of cybersecurity methodologies
BKH0011 Knowledge of the requirements and implementation of the risk management framework
WPC0017 Selecting tools
WPC0012 Generating alternatives

BKH0009 Knowledge of emerging technologies in cybersecurity and IT
BKH0008 Knowledge of the organization's mission and business processes
WPC0016 Keeping current
BKH0023 Knowledge of international and industry standards, common frameworks, and best practices in cybersecurity and business technologies
BKH0013 Knowledge of cybersecurity laws, policies, procedures, and governance that apply to the organization
WPC0011 Choosing a solution
BKH0006 Knowledge of laws, regulations, policies, and ethics relating to cybersecurity and the protection of personal information
WPC0019 Using tools
BKH0018 Knowledge of the stages of cyber attacks
BKH0020 Knowledge of personal data security standards
BKH0019 Knowledge of network security architecture concepts
BKH0017 Knowledge of physical components and architectures, peripherals, and operating systems

BKH0021 Knowledge of PCI-DSS standard.
BKH0016 Knowledge of concepts, terminology, operations, network, and telecommunications protocols
BKH0022 Ability to identify cybersecurity and privacy issues arising from connections with internal and external customers and partners
BKH0015 Knowledge of cybersecurity and privacy principles and organizational requirements of CIA and data classification

In query 1, the results provided useful answers that faithfully represented the data collected. Identical results were observed upon the comparison of the answers from both strategies, Protégé and Fuseki. As mentioned, there was a match with the data found directly in the WebProtégé ontology. This validation test can therefore be asserted as a success, as the query to determine the know-how, know-what, and know-how-to-be required for the cybersecurity analyst role is successful.

Query 2: What knowledge is required for the role of system security analyst?

In the same manner as query 1, answering this query required identifying the PREFIX, CyberSecOnto. The value of the system security analyst was assigned to the variable ?analyst, as shown in Table 4. This work role comes from the NIST cybersecurity framework. In the WHERE statement, the competency element that the study attempted to identify in the query is indicated, the knowledge required of the work role. These have been defined in the NIST framework integrated into the ontology as described earlier.

Table 4: NIST cybersecurity framework work roles

Work Role	Work Role ID	IRI
Authorizing Official/Designating Representative	SP-RSK-001	http://webprotege.stanford.edu/R7nTf0B3U0DJXtuIppFjfhM
Security Control Assessor	SP-RSK-002	http://webprotege.stanford.edu/R8gXL5AKGYc24dWbrEZLsML
Software Developer	SP-DEV-001	http://webprotege.stanford.edu/R90ocX8MaJkvScUDGWDHFuz
Secure Software Assessor	SP-DEV-002	http://webprotege.stanford.edu/R71rbHuUM6eggx7VcBASudO
Enterprise Architect	SP-ARC-001	http://webprotege.stanford.edu/RBgfoHn9KaUEYyXJmT9WV73
Security Architect	SP-ARC-002	http://webprotege.stanford.edu/RBpmMPf3tgGc0JcLW1hsMq2
Research & Development Specialist	SP-TRD-001	http://webprotege.stanford.edu/R8Rq6u74ze49G7GQ0uJLfkO
Systems Requirements Planner	SP-SRP-001	http://webprotege.stanford.edu/RBq24PstN7TIehfOCt37D5n

System Testing and Evaluation Specialist	SP-TST-001	http://webprotege.stanford.edu/ReHV6ES15q2WE4nqGExS7h
Information Systems Security Developer	SP-SYS-001	http://webprotege.stanford.edu/RDHEAlaxaRAeSpKJe9VcjSO
Systems Developer	SP-SYS-002	http://webprotege.stanford.edu/R85IBSKMWNKliTCtzvZGeR6
Database Administrator	OM-DTA-001	http://webprotege.stanford.edu/R8UJsrntYIXaypVKqotgny1
Data Analyst	OM-DTA-002	http://webprotege.stanford.edu/RDHZg9qWOFyA2F4bRgGrLhZ
Knowledge Manager	OM-KMG-001	http://webprotege.stanford.edu/RDnMTkIk8Ks7X4dlbzB8BSB
Technical Support Specialist	OM-STS-001	http://webprotege.stanford.edu/RBCYBLnATjduJOSsV4qPnE
Network Operations Specialist	OM-NET-001	http://webprotege.stanford.edu/R7BVKer03kw8OBJxdzdS0t
System Administrator	OM-ADM-001	http://webprotege.stanford.edu/RC1SWHhTYNKAdoNqFzN2fBB
Systems Security Analyst	OM-ANA-001	http://webprotege.stanford.edu/RDR1JnLSCzLfwVkxsVDjaVu

Cyber Legal Advisor	OV-LGA-001	http://webprotege.stanford.edu/RCyaKG66hoMNbO5fXU8xtSu
Privacy Officer/Privacy Compliance Manager	OV-LGA-002	http://webprotege.stanford.edu/RYKxd7UovKCw5E3qc80K3V
Cyber Instructional Curriculum Developer	OV-TEA-001	http://webprotege.stanford.edu/R9cf5CvhKoAFb4jkIXIOH4x
Cyber Instructor	OV-TEA-002	http://webprotege.stanford.edu/R7cdvRCcduvVmLzkByiRNyo
Information Systems Security Manager	OV-MGT-001	http://webprotege.stanford.edu/RCqYF1Bg9rxoec7JU3b1BBR
Communications Security (COMSEC) Manager	OV-MGT-002	http://webprotege.stanford.edu/R7AkrDK01WKCJDV8BeNQBty
Cyber Workforce Developer and Manager	OV-SPP-001	http://webprotege.stanford.edu/RDo8fHaIEtSQ41w3b4HGSgI
Cyber Policy and Strategy Planner	OV-SPP-002	http://webprotege.stanford.edu/R8ZmgcvgRt6GiUpsPuAjx46
Executive Cyber Leadership	OV-EXL-001	http://webprotege.stanford.edu/RBUftyHIgVJXn3KbSGUt2gs
Program Manager	OV-PMA-001	http://webprotege.stanford.edu/Rc9P1VTzjkFzf9DKO1p1NE

IT Project Manager	OV-PMA-002	http://webprotege.stanford.edu/RueUEJ2PWxV6gsp1SHlAn8
Product Support Manager	OV-PMA-003	http://webprotege.stanford.edu/RCnTFlek8VzZWIbwN0TGAFU
IT Investment /Portfolio Manager	OV-PMA-004	http://webprotege.stanford.edu/RCu3m8fbU5Sw5oyiQWmIc9z
IT Program Auditor	OV-PMA-005	http://webprotege.stanford.edu/R71Lwmb94XIySUCyoJXRbcQ
Cyber Defense Analyst	PR-CDA-001	http://webprotege.stanford.edu/RCKdg3t1APK7KW3LmYFR1OO
Cyber Defense Infrastructure Support Specialist	PR-INF-001	http://webprotege.stanford.edu/R961XQSISXnZ3wQR6NEQ5w4
Cyber Defense Incident Responder	PR-CIR-001	http://webprotege.stanford.edu/R8zzty8eMlgQ9gHLbMpuaG5
Vulnerability Assessment Analyst	PR-VAM-001	http://webprotege.stanford.edu/R7UTorAluYDcMMjMnLasxAT
Threat/Warning Analyst	AN-TWA-001	http://webprotege.stanford.edu/RkhOcp5iVKBr3bWzI3MzvW

Exploitation Analyst	AN-EXP-001	http://webprotege.stanford.edu/RDxGTcCnUywn4MwZtesktqK
All-Source Analyst	AN-ASA-001	http://webprotege.stanford.edu/R7Vigr5NPmwmlKfArHfGyMs
Mission Assessment Specialist	AN-ASA-002	http://webprotege.stanford.edu/RDxF56UC2HNIhuqa2i6DrL6
Target Developer	AN-TGT-001	http://webprotege.stanford.edu/RblhRGYbHdMP10lu9OFhAm
Target Network Analyst	AN-TGT-002	http://webprotege.stanford.edu/RByZtTskUBLsiuHliO1Rn7n
Multi-Disciplined Language Analyst	AN-LNG-001	http://webprotege.stanford.edu/RBTAMJ0vvP5QkIUUUXiBTrp
All Source-Collection Manager	CO-CLO-001	http://webprotege.stanford.edu/R7geWUTAGj6Phfwo0jSWe7v
All Source-Collection Requirements Manager	CO-CLO-002	http://webprotege.stanford.edu/R9kOf2Hi82BeNprBlyE9F0v
Cyber Intel Planner	CO-OPL-001	http://webprotege.stanford.edu/R8XpPCNxs50E1xy3xcT9TNM
Cyber Ops Planner	CO-OPL-002	http://webprotege.stanford.edu/R94RuI6YUc31HeoHPOs7mZn

Partner Integration Planner	CO-OPL-003	http://webprotege.stanford.edu/RmsGCz9eErw37YK72GOfUd
Cyber Operator	CO-OPS-001	http://webprotege.stanford.edu/RDe2OvHioGr9iBJRE8iVi02
Cyber Crime Investigator	IN-INV-001	http://webprotege.stanford.edu/RvDM2IYR7BDijL00lmPVaT
Law Enforcement /Counterintelligence Forensics Analyst	IN-FOR-001	http://webprotege.stanford.edu/R7rEvwCbieRmd87OkGB1vs2
Cyber Defense Forensics Analyst	IN-FOR-002	http://webprotege.stanford.edu/R88wGayotVrLflikfr2GSuU

Table 5 encapsulates all the object properties that were used in this ontology needed for this type of query, used in query 2. In the SELECT statement of the query, the variable relating to the object property that corresponds to what is being searched is used. In this query, this is the property that identifies the knowledge of the objects, which can be seen in Table 5 as the required Knowledge object property. To allow the search for other object properties, in a future implementation of the results of this study, for example, the other possible object properties that could be used with a similar query are presented in Table 5. The query results are presented in Table 6.

Table 5: Ontology object properties

Label	Description
requiredKnowledge	Used in a relationship to identify a knowledge element required for a particular cybersecurity actor or role
isRecognizedBy	Identifies an organization that formally recognizes a certification or degree
isIssuedBy	Name of the organization that issues a certification or a degree
educationalRequirement	Used in a relationship to identify an educational requirement for a particular cybersecurity actor or role
desirableKnowledge	Used in a relationship to identify a knowledge element that is desirable for a particular cybersecurity actor or role
desirableSkill	Used in a relationship to identify a skill that is desirable for a particular cybersecurity actor or role
isSimilarTo	Used to identify a similarity
requiredDegree	Used in a relationship to identify a degree that is required for a particular cybersecurity actor or role
requiredKnowHow	Used in a relationship to identify a know-how element that is required for a particular cybersecurity actor or role

performsTask	Indicates the tasks performed by an actor in a role
TrainingActivity	Used to identify a training activity that is intended for a particular work role or to achieve a competency or competency element
desirableAbility	Used in a relationship to identify an ability that is desirable for a particular cybersecurity actor or role
requiredSkill	Used in a relationship to identify a skill that is required for a particular cybersecurity actor or role
requiredKnowWhat	Used in a relationship to identify a know-what that is required for a particular cybersecurity actor or role
TrainingFor	Used to identify what competency or competency element a training activity
requiredCertification	Used in a relationship to identify a certification that is required for a particular cybersecurity actor or role
requiredKnowHowtoBe	Used in a relationship to identify a know-how-to-be element that is required for a particular cybersecurity actor or role
hasRole	Role of an actor
isSimilarToRole	Used in a relationship to identify similarities between cybersecurity actors or roles

	May also be used to create a relationship between similar roles in different competency frameworks
desirableDegree	Used in a relationship to identify a degree that is desirable for a particular cybersecurity actor or role
desirableCertification	Used in a relationship to identify a cybersecurity certification desirable for a particular cybersecurity actor or role
requiredAbility	Used in a relationship to identify an ability required for a particular cybersecurity actor or role
isMitigatedBy	Used to match the MITRE ATT&CK Tactics to the Mitigation measures

The following query was entered:

```
PREFIX CyberSecOnto:
<http://webprotege.stanford.edu/project/DpeZFWjjTksyFnolG1b70o#>

# show only unique rows (do not repeat rows with the same cell values)
SELECT DISTINCT ?knowledge

WHERE {

    # find something that is of type `RDR1JnLSCzLfwVkxsVDjaVu`
    ?analyst a
    <http://webprotege.stanford.edu/RDR1JnLSCzLfwVkxsVDjaVu>.

    # which has a property `hasKnowledge` whose value is something
    ?analyst CyberSecOnto:hasKnowledge ?knowledge.

}
```

Table 6: Query 2 results

Query results
K0048 Knowledge of Risk Management Framework (RMF) requirements
K0146 Knowledge of the organization's core business/mission processes
K0056 Knowledge of network access, identity, and access management (e.g., public key infrastructure, Oauth, OpenID, SAML, SPML)
K0090 Knowledge of system life cycle management principles, including software security and usability
K0287 Knowledge of an organization's information classification program and procedures for information compromise
K0169 Knowledge of IT supply chain security and supply chain risk management policies, requirements, and procedures
K0101 Knowledge of the organization’s enterprise IT goals and objectives
K0060 Knowledge of operating systems
K0044 Knowledge of cybersecurity and privacy principles and organizational requirements (relevant to confidentiality, integrity, availability, authentication, non-repudiation)

K0019 Knowledge of cryptography and cryptographic key management concepts
K0170 Knowledge of critical infrastructure systems with information communication technology that was designed without system security considerations
K0200 Knowledge of service management concepts for networks and related standards
K0018 Knowledge of encryption algorithms
K0059 Knowledge of new and emerging IT and cybersecurity technologies
K0126 Knowledge of Supply Chain Risk Management Practices (NIST SP 800-161)
K0267 Knowledge of laws, policies, procedures, or governance relevant to cybersecurity for critical infrastructures

In query 2, the results provided useful answers. Here as well, when compared to the answers from performing the query in Protégé, Fuseki, and the source data in the WebProtégé ontology, a match was observed, indicating that the validation test can be deemed successful. Table 6 presents the knowledge that is required for the role of system security analyst, retrieved with query 2.

Query 3: What are the know-how elements required for the red team analyst?

For query 3, the PREFIX CyberSecOnto was employed and assigned the value red team analyst to the variable ?analyst, as shown in Table 7. The work role red team analyst corresponds to an individual in the ontology, and this is what makes this query different from the two previous queries focuses on ontology classes. In the WHERE statement, the

competency element was used to identify the know-how required for the work role. These were clearly defined in this study and are presented in Table 2. In the SELECT statement, the required variable was used, which was also assigned as knowledge, similar to the previous queries.

Table 7: Individuals identified in the ontology

Individuals	IRI
Senior Advisor	http://webprotege.stanford.edu/R3AvgajJEn7hWvEbnQSLDs
Senior Director Information Security Mechanisms	http://webprotege.stanford.edu/R7Mgp5vzs29h2uQi8EWLplq
Director Information Security Defense	http://webprotege.stanford.edu/R7Sd1Hvp9IHglZaV0SS7GkN
Red Team Analyst	http://webprotege.stanford.edu/R7bGEqPo69Ee7iBx0OvVlJD
Investigation Technician	http://webprotege.stanford.edu/R7esP8mf4ULvfWHwa2KvEld
Director Information Security Offense	http://webprotege.stanford.edu/R7goEAjNVevNvcMLFUUNbqq
GRC Analyst	http://webprotege.stanford.edu/R8JjBjZKvoUBa8GJ0pdIjFB
Senior Director Information Security Strategy	http://webprotege.stanford.edu/R8M2tw4J23PXR6RhmMrGRFu

Director Information Security Projects	http://webprotege.stanford.edu/R8WnxTJnc6FjDeYfhM6CPMr
Director Identity and Access Management	http://webprotege.stanford.edu/R8hNA8A7eF5YX0SxN4TZCbA
Senior Director Information Security Surveillance and Response	http://webprotege.stanford.edu/R8q1OStIPyoLCgXiMKMuj52
Manager Security advisory services	http://webprotege.stanford.edu/R93r1vJ90j0ZM5HSdlQAvJ1
Senior Intelligence Advisor	http://webprotege.stanford.edu/R9MXbEgtYhDvMibLBLEdpM1
Director Information Security Intelligence and Analytics	http://webprotege.stanford.edu/R9s1GXco0WUwsHN1W8WlXzA
Director Governance	http://webprotege.stanford.edu/RBG63DuERGNF6SvQAmvrswv
Awareness Advisor	http://webprotege.stanford.edu/RBhquySh0Mv6sMKZuEZOMOj
Director Physical Security	http://webprotege.stanford.edu/RBjHnikoBPUrhGjJ8hUN3c8
Manager Information Security Awareness	http://webprotege.stanford.edu/RBq8NHePp8kcwegfOkhvdMW

Director Security Integration	http://webprotege.stanford.edu/RCBqdDuiO1P16y4xWnWCPlM
Intelligence Intern	http://webprotege.stanford.edu/RCHlnlK3f1EYor4DF1i7ylD
Chief Information Security Officer	http://webprotege.stanford.edu/RCMdmanN73dxmttxGgq1dgX
Director Information Security Planning	http://webprotege.stanford.edu/RCOKfq0VNwOsMqqdk00YUjf
Blue Team analyst	http://webprotege.stanford.edu/RCfgS40OzkThXlfF7h8NCS4
Senior Director Governance	http://webprotege.stanford.edu/RDQYyx9ahyKnHFH5AA0zzgW
Vulnerability Analyst	http://webprotege.stanford.edu/RDThlSfsHyxywmrjpksbbnc
Investigation Analyst	http://webprotege.stanford.edu/RECLyWpZ7KvzLjzoZWOPzh
Business Information Security Officer	http://webprotege.stanford.edu/RSvmJa0e4HbrOKbYnenhBj
Intelligence Advisor	http://webprotege.stanford.edu/RWhL0QH7zttGf7yENj7Mt0
Cyber Instructional Curriculum Developer	http://webprotege.stanford.edu/Rhox1vp4wx9D7eHs6xbD4U

The query that was used to identify the know-how is as follows:

PREFIX CyberSecOnto:
<http://webprotege.stanford.edu/project/DpeZFWjjTksyFnolG1b70o#>

SELECT ?knowhow

WHERE { <http://webprotege.stanford.edu/R7bGEqPo69Ee7iBx0OvVlJD>

find exactly the resource `R7bGEqPo69Ee7iBx0OvVlJD` with the property `hasKnowHow` whose value is something

<http://webprotege.stanford.edu/R7bGEqPo69Ee7iBx0OvVlJD> CyberSecOnto:hasKnowHow ?knowhow.

}

Table 8: Query 3 results: work role know-how

Work role know-how
WPC0016 Keeping current
BKH0023 Knowledge of international and industry standards, common frameworks, and best practices in cybersecurity and business technologies
WPC0012 Generating alternatives
WPC0017 Selecting tools
BKH0002 Knowledge of cybersecurity methodologies

BKH0016 Knowledge of concepts, terminology, operations, network, and telecommunications protocols
BKH0021 Knowledge of PCI-DSS standard
BKH0003 Knowledge of risk management processes
PEC0021 Following directions
BKH0018 Knowledge of the stages of cyber attacks
BKH0006 Knowledge of laws, regulations, policies, and ethics relating to cybersecurity and the protection of personal information
WPC0019 Using tools
BKH0017 Knowledge of physical components and architectures, components, peripherals and operating systems
WPC0011 Choosing a solution
BKH0020 Knowledge of personal data security standards
WPC0013 Identifying the Problem
BKH0001 Knowledge of cybersecurity and confidentiality principles

BKH0019 Knowledge of network security architecture concepts
BKH0005 Knowledge of the specific operational impacts of cybersecurity gaps
BKH0007 Knowledge of computer network concepts and protocols
BKH0022 Ability to identify cybersecurity and privacy issues arising from connections with internal and external customers and partners
WPC0018 Troubleshooting and maintenance
BKH0015 Knowledge of cybersecurity and privacy principles and organizational requirements of CIA and data classification
WPC0014 Implementing the solution
WPC0015 Locating, gathering, and organizing relevant information
BKH0004 Knowledge of cyber threats and vulnerabilities

The results from query 3 are presented in Table 8. For this query as well, the results from Protégé and Fuseki were identical and provided meaningful answers from management perspective in accordance with the original requirements, further substantiating the validation test. It is therefore possible to successfully query individuals in the ontology, as shown, and obtain an answer allowing to identify the know-how elements required of the ontology individual red team analyst.

Query 4: What are the know-what elements required of the blue team analyst?

This SPARQL query does not bring added value when compared with the three previous ones. However, this query was used to prepare the external validation tests by creating a risk scenario involving the blue team analyst, a critical risk mitigation role in the cybersecurity team of a financial institution. For query 4, the PREFIX CyberSecOnto was used and assigned the value of the blue team analyst to the variable ?analyst, as shown in Table 7. Here as well, this work role corresponds to an individual in the ontology. In the WHERE statement, the competency element used in the query is identified, the know-what required for the work role. These were clearly defined in the study and are presented in Table 2. In the SELECT statement, the variable is defined as:

```
PREFIX CyberSecOnto:
<http://webprotege.stanford.edu/project/DpeZFWjjTksyFnolG1b70o#>

SELECT ?knowwhat

WHERE { <http://webprotege.stanford.edu/RCfgS40OzkThXlfF7h8NCS4>

        # find exactly the resource `RCfgS40OzkThXlfF7h8NCS4` with the
property `hasKnowWhat` whose value is something

        <http://webprotege.stanford.edu/RCfgS40OzkThXlfF7h8NCS4>
CyberSecOnto:hasKnowWhat ?knowwhat.

}
```

Table 9: Query 4 results: work role know-what

Know-what elements required for the blue team analyst
WPC0001 Acknowledging team membership and role
WPC0026 Safeguarding one's person
WPC0009 Generating innovative solutions
WPC0008 Employing unique analyses
WPC0022 Global awareness
WPC0006 Planning
WPC0024 Situational awareness
WPC0020 Business ethics
WPC0023 Market knowledge
WPC0010 Seeing the big picture
WPC0003 Identifying with the team and its goals
WPC0007 Prioritizing

WPC0002 Establishing productive relationships
WPC0025 Maintaining a healthy and safe environment
WPC0005 Managing projects
WPC0004 Resolving conflicts
WPC0021 Business practices

The results from query 4 are encapsulated in Table 9. In this query as well, the query results from Protégé and Fuseki were identical and provided meaningful answers from management perspective, thereby substantiating the validation test.

6.6 The ontology as a talent management tool for financial institutions

After the completion of the validation tests, the use of the ontology as a management tool for financial institutions can be further explored. To perform this investigation, a simple tool that allowed us to query the ontology was developed. In the future, this is anticipated to enable organizations to connect the ontology with a tool to facilitate the creation of management reports to be used with a management dashboard application. The use of Python and Java applications were explored, selecting to develop in Java using the Eclipse integrated development environment with an Apache Tomcat back end, as it was the most feasible option that allowed to create a viable product. This would prove to be sufficient to evaluate the ontology in this study. To use this application, it is necessary to install Tomcat (see: https://medium.com/@ngotantien/how-to-install-apache-tomcat-9-on-windows-mac-os-x-ubuntu-and-get-started-with-java-servlet-45f959d7ee0a). Upon installation, it is necessary to start Tomcat, which can be easily accomplished by copying the following instructions into the macOS Terminal window:

cd /usr/local/apache-tomcat-9.0.40/bin

./startup.sh && tail -f ../logs/catalina.out

Installed on a standalone computer, this application can then be accessed through the link: http://localhost:8888/Maleger. Figure 5 presents an example of the knowledge competency element search query for the Cybersecurity role of Senior Advisor.

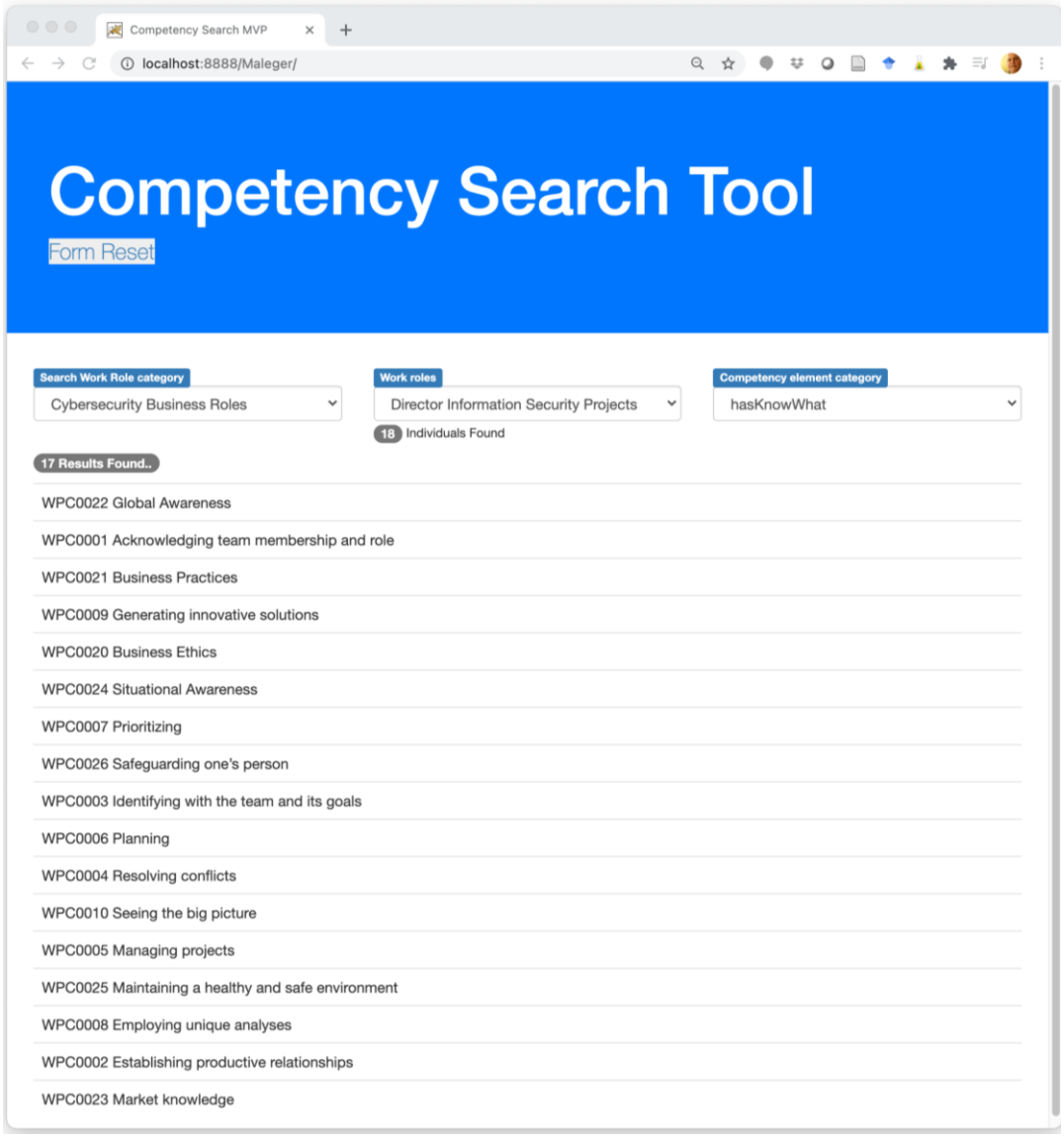


Figure 5: Competency search application example

6.7 Validity of the ontology

An ontology to represent cybersecurity competencies expected from cybersecurity workers in their work roles in a Canadian financial institution was designed and built. The foremost aspect that was explored in this section was the degree of validity of the ontology in accurately representing the cybersecurity domain. The validated cybersecurity competency ontology would be a valuable solution to help resolve the problem of matching risk management roles and competencies with actors. As the quality of an ontology indicates how well it provides a suitable description of the problem domain while being syntactically correct, precise, and semantically suitable (Vyšniauskas et al., 2012), it was necessary to assert if the ontology that was developed addressed the problem domain. Once this it was ensured that it met the fundamental semantic criterion for an ontology, the goal of the ontology validation was to verify whether it could represent and contribute to optimizing the cybersecurity actor-role-competency fit. For this purpose, the results of three different queries based on realistic questions that may be asked in a cybersecurity competencies ontology were compared. These queries considered various competency elements required from cybersecurity workers in a financial institution. Subsequently, the results of the queries using four different tools were compared: WebProtégé, Protégé, Fuseki, and a custom Java application. In all the four tools, the same results were observed, i.e., a list of the competency elements that corresponded to the search query. When this list was compared with the actual cybersecurity competency requirements in the field, it was possible to confirm the correspondence between the results from the query and the requirements identified by the participants in this study. Hence, it can be deduced from the findings of this study that the proposed ontology was able to fulfill this requirement, as shown in tests 2 and 3. In these tests, the ontology was queried to extract information about the competencies required in different work roles. To make the queries possible, it was necessary to define the individuals and the data properties in WebProtégé, which had not been done initially. This, however, did not change the data in the ontology required to make it usable but rather affected how the information is stored. It was also discovered that in many cases, it is more effective to edit the ontology with an XML-enabled text editor directly. However,

in doing so, meticulous efforts are required for the maintenance of integrity of the ontology.

The other crucial aspect that was investigated to assess the validity of the cybersecurity competency ontology was the effectiveness of the ontology as a talent management decision tool. For this purpose, a simple prototype was developed and tested, which can be readily expanded into a more comprehensive management tool and dashboard. However, at this stage in the study, the objective was to evaluate the usability and practicability of this tool and whether it provided organizational value by utilizing the proposed ontology. With the same queries that were used in the validation, it was realized that it is possible to develop a simple interface that can provide answers from the ontology that do not require any knowledge of SPARQL from the user. Therefore, this validation test was considered a success.

6.8 Conclusion of this section

This section elucidates how a cybersecurity competency ontology was validated and tested in a large Canadian financial institution, supporting hypothesis H2. Cybersecurity competencies are pivotal to financial organizations as competency gaps create vulnerabilities that, in turn, contribute to unacceptable risks. This study proposed a partial solution to reducing these competency gaps by employing an ontology. Following a brief overview of the existing literature, this study demonstrated how the proposed ontology can potentially assist organizations in filling cybersecurity roles with competent individuals. It was also observed how a simple interface could be developed to help managers answer real-world organizational questions. In relation to the initial goals, this can be considered an accomplishment.

Future work is anticipated to be established on the observations of the study and the results reported in this section. In accordance with that, it is expected to develop a complete management information system and dashboard, which the participant financial organization has just agreed to initiate, upon the conclusion of this study. A first version of this management information system and dashboard is presented in Appendix Q.

7 External validity test

External validity relates to generalizability of the results. It concerns how the results of a study can be used to explain and understand the reality of other organizations than the one where the study was performed. Of course, they would need to share some characteristics – in this case, a good level of external validity that would allow the results to create value for other Canadian organizations in the financial sector would be excellent.

More specifically, the goal was to perform an external validation test on the ontology by doing two tests:

1. Matching cybersecurity work roles posted in online job offers on to work roles in the ontology, namely the work roles used in previous queries.
2. Matching cybersecurity work roles in an employee CV with the work roles used in previous queries to find matches.

If these tests can be performed for various work roles on multiple documents and reliably obtain usable results, then this test would be considered successful. This section describes the tests that were performed to achieve this, gather data on the reliability of the tests, presents the analysis of the results, and conclude with an assessment of the external validity of the cybersecurity competency ontology.

7.1 Choosing a tool to perform the test

This study investigated the possibility of using different off-the-shelf solutions to assist in performing the external validity test. Developing and building a custom solution that would allow us to integrate the ontology was not considered to be a viable option at this point in the study for many reasons. Principally, the time required to do would make this very difficult in relation to the benefits, considering that there might be ready-to-use alternatives that could achieve the desired results, even if a few relatively simple additional steps could be necessary. As well, this did not lie within the purview of this study. Appendix B presents the pros and cons of the most likely available solutions that

were investigated. In some instances, software was installed and a few initial trials were performed to help in this evaluation to determine the information. On the basis of this evaluation of the pros and cons list that is presented, it was decided to use Stardog for the external validity test. As it was available at no cost for the study and was already installed and configured, this seemed like a reasonable choice. There was no reason or missing features that would justify using a different product for the required work in the study.

Stardog (https://www.stardog.com/) is a commercial enterprise knowledge graph platform. Knowledge graphs have become in recent years as a popular IT tool to manage unstructured data, such as RDF triple stores and heterogeneous data.

7.2 Enabling the Stardog search

To perform the test, the data had to be imported into the Stardog database system. This was done in several steps. First, owing to a limitation in Stardog, it was necessary to ensure that all filenames were less than 100 characters in length with no spaces in the names. As the cybersecurity CVs and job postings were gathered over several months using LinkedIn, to be used at this stage in the study, and some filenames generated automatically were much longer. In the future, this could be avoided by ensuring the names are of a usable length when they are gathered; however, this was not known during this stage in the study.

Once the files were ready to be loaded, they needed to be parsed into an RDF format. Initially, this was performed using a software tool called the Stanford CoreNLP Natural Language Processing Toolkit with the English-language jars files. This allowed us to quickly produce annotated text files from the dataset. While this was not the ideal tool to use for all the tests that were performed or to provide a completely automated solution for the organization, it was suitable for the external validity that is described in this section. It was later realized that it had some limitations, which are discussed here.

The files were added to the Stardog document store on the Stardog server available in the lab on the university's campus. The intervention of a lab assistant was required as a

privileged local access to the server's file system was required to perform this task. Adding the files required the following steps:

1. Copy the files to a local directory on the server – in this case /home/ubuntu/CV_CyberJobs. Use a different directory for different categories of files to allow for better file management and maintenance.
2. Create the Stardog database, using the following commands:

   $STARDOG_HOME = /root/stardog-7.3.0/

   $STARDOG_HOME/bin/stardog-admin db create –o search.enabled=true –n myDb
3. Move the folder with required files under the newly created database:

   mv /home/ubuntu/CV_CyberJobs $STARDOG_HOME/myDb
4. Load the files in the database:

   stardog doc put -u username -p Password myDb *.pdf
5. Create a separate database in Stardog using the **CoreNLPEntityLinker** custom RDF extractors using the following command:

   stardog doc put --rdf-extractors CoreNLPEntityLinker name -u username -p Password myDb File1.pdf

Using this process, 363 job postings that were collected from LinkedIn and Indeed over the duration of the study were loaded. The employment data of 57 research participant profiles from the CV posted in their public LinkedIn profile were also added. A review of the posting was also done to evaluate them in relation to the categories used in the queries. This resulted in the identification of 267 security analysts, 14 red team and 106 blue team job offers, some being in two categories. Of the collected job postings, 139 security job postings were in none of these categories. For the 57 CVs, 29 security analysts, two red teams, 10 blue team, and 48 for management positions, some being in two categories, were identified. Of the collected job postings, 42 were found to be in none of these categories. This information was used to evaluate the accuracy and repeatability of the results as it enabled the researchers to identify false positives and false negatives in the query results.

To perform the tests and compare the results after various tests, two databases in Stardog were created. The databases created were:

1) CyberSec003, with the ontology, job postings, and CVs in a single database loaded with the default Stardog entity linker, Tika. This database was used in queries 5, 6, 7, and 8.
2) CyberSec004, with the ontology, job postings, and CVs in a single database using the CoreNLP entity linker. This database was used in queries 9, 10, 11, 12, 16, 17, 18, and 19. This database was also used for the queries used for the risk scenarios SC1 and SC2.

While developing these queries, additional queries were also performed to test various hypotheses, which were not successful, and the failed queries, 13, 14, and 15 are not presented here. Stardog offers predictive or probabilistic inference support through supervised machine learning; typically, this is performed as either a classification (categorical) or regression (numerical) problem-solving operation. Performing this analysis requires a structured definition of a concept with properties or attributes, which can be found in the cybersecurity competency ontology. In this study, the job postings and the CVs required in the analysis process were also used. As this large selection of unstructured PDF documents was collected over several months, these had no clear boundaries for the different data elements that would easily be used to create a defined structure. Extracting or creating structure out of free-form text is a distinct problem. Learning ontology from that structure is another. Named entity recognition, which comes with Stardog, can only help establish which ontological terms appear in a document.

Several options to perform this using tools, off-the-shelf software, and the opportunity of creating a custom Python data mining application to have a mechanism to pre-process the documents to extract the data prior to integration into the Stardog database were investigated. Pre-processing requires multiple steps, such as tokenization, removal of stop words, lemmatization, tagging of nouns, verbs, and adjectives, and filtering to reduce sparsity and noise (Bernabé-Moreno, Tejeda-Lorente, Herce-Zelaya, Porcel, & Herrera-Viedma, 2019). The use of various ontology learning techniques from the fields of natural

language processing, machine learning, information retrieval, data mining, and knowledge representation that have contributed to the improvement of ontology development were also considered (Asim, Wasim, Khan, Mahmood, & Abbasi, 2018).

What would best serve the study is to identify, or define, the abstract model or an ontology for the documents. However, the documents collected are free-form text retrieved over a long period from LinkedIn and Indeed, as mentioned previously. It would be possible to use the tools that were identified if a shared or common structure to the text was found. However, the initial analysis of the documents indicated this would be difficult. This is something that could have been done differently if this requirement was identified earlier in the study. This could have allowed for a different retrieval strategy, indexing the documents gradually as they were added. However, as this was a final external validation after nearly two years into the study, it was decided not to redo this part of the data collection unless it would become necessary, which it was not. One approach would be if the named entity recognition tool could extract full concepts with the class and its properties from each of the documents. However, it was possible to obtain extracted classes and properties separately directly in Stardog with a simple process using internal features, text parsing, and analytics.

While pre-processing of the documents could work, there would be a number of work, testing, and validation activities required. Within the limited scope of the validation activities, a simpler strategy was used that yielded similar results in less time overall but required additional steps and using Excel. In that way, a job profile document could establish that it has an analyst role described in accordance with the properties that are found. Initial tests on a small sample indicated that it was possible to perform an analysis of the documents. It was observed that if there is a job post document or a LinkedIn post that establishes the same, an ontological match could be identified with the data available. This could be achieved by performing a review of the documents to determine what category should be. The detailed query results and analysis table are presented in Appendix A.

7.3 Stardog matching

The text matching and analytics functions of Stardog use the Apache Lucene syntax and structure (StarDog, 2021). The Lucene scoring is based on similarity based on probabilistic models and on how the documents are indexed. This can be configured and optimized, as explained in the Lucene documentation ("Lucene," 2021). Lucene indicates that specific scores cannot be compared across different searches. However, the score was not used in the study other than in early ranking of results within a query result in initial tests, as significantly more data and tests are needed to be able to automate the processing of documents and use this score in any reliable manner. The score, combined with a validation process, was initially used to confirm the validity of the results in the limited context of this study. There were only very small differences in the results between the data using Stardog markup (CyberSec003) and the data using CoreNLP (CyberSec004). These differences can be seen in the tables summarizing the different test queries, which are presented later in this dissertation. However, before the tables can be presented, there is more information required on the tests and their validity using the F1-scores.

7.4 F1-score

F1-scores have become a popular evaluation metric for the evaluation of classification problems (Sokolova & Lapalme, 2009). They are used by first determining the precision and recall value of the results of a classification to assess performance (Hand, Christen, & Kirielle, 2021; Tharwat, 2020; Van Rijsbergen, 1974). Precision is defined as the number of correctly classified positive examples, or true positives (TP), divided by the number of examples erroneously labeled by the system as positive, or the false positive results (FP). The value of TP is an indication of the classification capacity of the queries. The higher the value of TP, the better. While a coin toss or random sample could be expected to provide 50%, an F1 score above 0.5 would indicate a better classification than random. False positives are classification errors. Recall is a measure of the repeatability of the classification process. Recall is calculated by using the number of correctly classified positive examples divided by the number of positive examples in the data. F1 scores

provide a measure of the retrieval of positive examples in a classification problem but neglect the correct classification of negative examples. This study used the F1-score, which has the following formula:

$$F1 - Score = 2 \times \frac{(TN \times TP)}{(TN + TP)}$$

7.5 Matthews Correlation Coefficient

The Matthews Correlation Coefficient (MCC) provides information regarding the correlation between the observed and predicted classifications (Tharwat, 2020). It is considered a good indicator of the quality and accuracy of classification (Chicco, Tötsch, & Jurman, 2021). It provides a more informative and truthful score to evaluate and compare the reliability of classification predictions (Chicco & Jurman, 2020). The expected values are from +1 to −1, where a coefficient of +1 indicates a perfect prediction and −1 a total disagreement between prediction and true values. A result of zero means that the predictions is no better than what could be expected if a random classification was performed. The MCC calculation uses the following formula:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

7.6 Document review

To calculate the F1-score and MCC, it was necessary to have data that could allow the evaluation of true and false nature of the positive or negative answers that was needed to use to calculate the F1-scores. For this purpose, a document review of all the job postings was performed. As will be explained later, this was not done for the CVs, as it was found there was not enough data to provide reliable or significant F1-scores. This process was manual, meaning that every document was opened and reviewed, and then classified into relevant categories related to the queries. The detailed results for the 363 job postings were integrated into the complete table presented in Appendix A. There are cases where two or more categories were applicable. For example, a system security analyst with a

management role would be in two categories. In summary, the document reviewed identified:

- 270 Cybersecurity analysts, used in queries 1, 5, 9, and 16.
- 270 System security analysts, used in queries 2, 6, 10, 17, and scenario 1.
- 16 Red team analysts, used in queries 3, 7, 11, and 18.
- 109 Blue team analysts, used in queries 4, 8, 12, 19, and scenario 2.
- 48 Managers
- 42 Other posting not in any of these categories

The language of the job postings, either in French (66), English (301), or both (13), was identified. As both are official languages in Canada, this was done to investigate if the language would be an issue in the relevance of the results. In the analysis of the data, the questions of the language of the document affecting the precision of the classification arose. By having this information, it was possible to examine this possibility. The study demonstrated that it had no significant impact on the results.

7.7 Performing the Stardog classification

The Stardog classification was performed by using SPARQL queries in Stardog system with the database that was created, as described previously. On the basis of the results from the previous queries, 1 through 4, a strategy was developed to perform the queries. In the remainder of this section, the process to query the databases to get answers to address the research questions is described.

7.7.1 Connecting to Stardog

Prior to executing the queries, it was necessary to connect to the server. This was done by following the steps enumerated as follows:

1. Start Stardog studio and connect to Stardog server; in this study, the server is http://dev2.gagnontech.org:5820. A valid user account and password are required.
2. Connect to the ontology database instance.

3. From there, the Workspace is accessed by using the left-side menu. This is where SPARQL queries can be created and executed. This requires users to select the database against which to execute the query. In this study, as mentioned, there are two, using Stardog markup (CyberSec003) and the data using CoreNLP (CyberSec004).

7.7.2 Executing the queries with Stardog search

Once connected to Stardog, the queries that were used in the first test could be executed. As indicated, the external validity test compared the CV of individual actors, extracted from the LinkedIn profiles of participants in the study, all of which had previously signed informed consents to participate in this research. This was done by analyzing the CV in a text format, and in this case extracted as a PDF file and processed using CoreNLP, as explained previously. The same query also made it possible to obtain the data for the second test, which compared cybersecurity job posting from Canadian companies to the ontology to classify them. For this purpose, the same work roles used previously in the study were used, such as in Section 10:

1. The cybersecurity analyst role, previously used in query 1, was used in Q5, Q9, and Q16
2. The System security analyst role, previously used in query 2, was used in Q6, Q10, and Q17
3. The red team analyst role, previously used in query 3, was used in queries 7, 11, and 18
4. The blue team analyst role, previously used in query 4, was used in queries 8, 12, and 19

Three different sets of queries were used:

1. Queries 5, 6, 7, and 8 used the database CyberSec003 (using Stardog markup), using the ontology classes and the Stardog textMatch function.
2. Queries 9, 10, 11, and 12 used the database CyberSec004 (using CoreNLP text annotation) and the Stardog textMatch function.

3. Queries 16, 17, 18, and 19 used the database CyberSec004 (using CoreNLP text annotation), with the Stardog linker advanced search function with dc:reference.

In Stardog, the document store full-text search can retrieve the textual contents of a document for indexing. Once a document is added to the Stardog database using BITES, the contents of the documents can be searched using the standard textMatch predicate in SPARQL queries. This is different and depends on the strategy used; in this study, three different strategies were explored: the default Stardog (Lucene), CoreNLP, or the advanced linker functions. When the textMatch command is used by itself, as in these first queries, the documents are searched for matches against the text elements included in the query itself and classes and subclasses of this element in the ontology that has been loaded.

From this work, SPARQL queries with the roles described were used for further queries. First with **Run**, adding the Limit 10 on the last line to quickly test the query. There could be a typographical or a transcription error that could be fixed quickly rather than having to run a query that could take as much as 30 seconds to execute. Once the query was ready, the limit was removed and **Run to file** was used, as there is a limit of 1000 in Stardog studio to run a query on the web interface. The file produced by the **Run to file** command was in a CSV format, which could be imported and manipulated in Microsoft Excel. As stated previously, in an eventual management implementation of a tool, this could be automated in Python, Java, or using other tools. However, after a short investigation, no benefits were found in developing an application for this study considering the effort required compared to the benefits when there was a simpler alternative of Microsoft Excel. The query was executed in the two databases that were created, as mentioned previously, which included the cybersecurity competency ontology, the CVs, and the job postings that were collected. CyberSec003 (Stardog markup) was used for queries 5, 6, 7, and 8, and CyberSec004 for queries 9, 10, 11, and 12 with the very similar queries to investigate the effect of the entity processing on the quality of the results. The results of these queries are integrated using Excel into a large table, presented in Appendix A. The general structure of the queries is as presented

below, with some changes for the specific element, that is, the object of the query, show here in italics:

```
prefix fts: <tag:stardog:api:search:>

SELECT * WHERE {

   # find something that is a Class and has a label of something
   ?class a owl:Class ;

        rdfs:label ?keyword .

  # search the text index for the string `object of the search` and return the search score and result (which can be a document or another resource)
  service fts:textMatch {

      [] fts:query "object of the search" ;

        fts:score ?score ;

        fts:result ?result ;

  }

}

order by desc(?score)
```

Once the queries were executed, the results were found in the CSV files. The query results from the CSV files created by Stardog could not be imported into Excel, as Excel has an import limit of a maximum of 1,048,576 rows. Using a free tool, OpenRefine, it was possible to import the CSV files by creating a study for each query. Once all the duplicate rows were removed, all the information was added to a table. The results from this are integrated into Appendix A, where the results also show additional treatment explained later in this section.

To remove the duplicates in OpenRefine the following procedure was used:

1. Export a copy of the data as a CSV file to have a backup in case of problems.
2. Identify the column with duplicate information that could be used – in this case column 4, named **result**.
3. Sort the data on this column by clicking the triangle left of **result**, and then choose **Sort** and select the **text** bullet.
4. Once the sorting process is completed, reorder permanent by using **Sort** menu on the top-left side of OpenRefine and choose **Reorder rows permanently**.
5. Next, blank the duplicate result rows by clicking on the **result column triangle**, then choose **Edit cells** followed by **Blank down**.
6. Eliminate the rows with blank results by clicking on the **result column triangle**, then **Facet**, followed by **Customized facets**, and lastly **Facet by blank**.
7. Finally, in the left side panel, select **false**. What is now displayed should be the unique **result** rows.
8. Then export these in OpenRefine as **Excel 2007+** in the top right side menu **Export**.

Once the documents were in Excel format, they could then be integrated into a table. The results from the query were compared with the document review and the data was labeled, as presented in Table 10. For example, when the document review and the query results matched, the data was identified as true positive and labeled TP.

Table 10: Result label confusion matrix

Confusion matrix	**Query positive**	**Query negative**
Review positive	True positive (TP)	False negative (FN)
Review negative	False positive (FP)	True negative (TN)

To analyze the results, a large table was created with all the data from the queries used in the external validity tests for all the job postings, which is presented in detail in Appendix A. In the following sections, a summary of this table is presented.

7.7.3 Cybersecurity analyst role

Query 5 (Q5), query 9 (Q9), and query 16 (Q16) investigated a cybersecurity analyst in a governance role supporting risk assessment and management advisory. This is the same role that was previously used in queries 1, which was used in the test for queries 5, 9, and 16. As mentioned previously, Q5 used textMatch and the CyberSec003 (Stardog markup) database in Stardog, while Q9 used textMatch with CyberSec004 using CoreNLP text annotation and Q16 used dc:reference with CyberSec004 with CoreNLP text annotation. The results of the queries were processed with OpenRefine, as described previously, and the results integrated into a table. Finally, they were compared to the document review results to determine if they should be evaluated as TP, FP, TN, or FN. From there, the precision and recall values were calculated, which made it possible to calculate the F1-score, as can be seen in Table 11.

Table 11: Cybersecurity analyst role query results

	Q5	**Q9**	**Q16**
No. of documents	363	363	363
True positives	211	146	123
True negatives	65	100	108
False positives	57	28	22

False negatives	30	89	110
Document review count	270	270	270
Precision	0.79	0.84	0.85
Recall	0.88	0.62	0.53
F1-score	0.83	0.71	0.65
MCC	0.44	0.38	0.35

The reader will notice that the F1-scores for Q5, Q9, and Q16 are high, which indicates that in this situation, all three queries performed on the ontology and the documents are successful predicators. In all three queries, high precision can be observed. While the recall is higher with textMatch than dc:reference, because of the high number of false negatives, the results indicate that the queries are successful in reference to the intended purpose. The MCC indicates that the results are better than random but not a prefect prediction.

7.7.4 System security analyst role

Q6, Q10, and Q17 investigated a different role – that of a system security analyst supporting vulnerability mitigation and cyber defense activities following a security incident. This was done in the same manner as the cybersecurity analyst role presented in the previous section, only by changing the object of the query. Here as well, the query results were processed with OpenRefine and compared with the document review to determine positives and negatives, leading to the determination of the F1-Scores presented in table 12.

Table 12: System security analyst role query results

	Q6	Q10	Q17
No. of documents	363	363	363
True positives	262	5	217
True negatives	7	93	27
False positives	86	0	70
False negatives	8	249	49
Document review count	270	270	270
Precision	0.75	1.00	0.76
Recall	0.97	0.02	0.82
F1-score	0.85	0.04	0.78
MCC	0.10	0.07	0.10

For this group of queries for the system security analyst, the F1-scores for Q6, Q10, and Q17 are high, indicating successful predicators. However, in the case of Q10, this is misleading because of a low number of true positives and high number of true negatives

and false negatives, supporting a very low recall value. The Q10 results were rechecked several times to ensure this was not caused by an error in the query or in the analysis of the results, but this was not the case. There were no false positives; however, there were many false negatives, which explained the low recall value for Q10. The recall is higher with Q6, using textMatch and Q17, using dc:reference, thereby producing high F1-scores for those two queries. The MCC values indicate the results are only slightly better than a random classification.

7.7.5 Red team analyst role

Q7, Q11, and Q18, investigated a red team analyst performing penetration testing and vulnerability identification. This is the same role that was previously used in Q3. As mentioned previously, Q7 used textMatch and the CyberSec003 database in Stardog (with Stardog document markup), while Q11 used textMatch with CyberSec004 with CoreNLP text annotation and Q18 used dc:reference with CyberSec004 with CoreNLP text annotation. As in the previous queries, the results of the queries were processed with OpenRefine, integrated into a table, and compared with the document review evaluation. From there, the precision and recall values were calculated, which made it possible to calculate the F1-score, as presented in Table 13.

Table 13: Red team analyst role query results

	Q7	**Q11**	**Q18**
No. of documents	363	363	363
True positives	13	13	11
True negatives	56	55	111

False positives	293	293	230
False negatives	1	2	11
Document review count	16	16	16
Precision	0.04	0.04	0.05
Recall	0.93	0.92	0.5
F1-score	0.08	0.08	0.08
MCC	0.05	0.01	−0.09

Q7, Q11, and Q18 exhibited low precision, with a very number of true positives. This being a more specialized work role and considering that the document review count indicated a low number of actual job postings (16), the high number of false positives helps to understand the results. It was observed that some of the ontology terms used in the construction of the concept were often mentioned in cybersecurity job postings, thereby contributing to the high number of false positives and causing a low precision but with a high recall value and thus low F1-scores. The MCC results indicate that these results are close to what we could expect if this was done at random.

7.7.6 Blue team analyst role

Q8, Q12, and Q19 investigated a blue team analyst responding to a security incident. This is the same role that was previously used in Q4. As mentioned previously, Q8 used textMatch and the CyberSec003 database in Stardog (with Stardog document markup), while Q12 used textMatch with CyberSec004 with CoreNLP text annotation, and Q19

used dc:reference with CyberSec004 with CoreNLP text annotation. Likewise the previous queries, the results of these queries were processed with OpenRefine, integrated into a table and compared with the document review. From there, the precision and recall values were calculated, which made it possible to calculate the F1-score, as can be seen in Table 14.

Table 14: Blue team analyst role query results

	Q8	**Q12**	**Q19**
No. of documents	363	363	363
True positives	100	98	77
True negatives	36	45	96
False positives	218	211	158
False negatives	8	9	32
Document review count	109	109	109
Precision	0.31	0.32	0.33
Recall	0.93	0.92	0.71
F1-score	0.47	0.47	0.45

MCC	0.09	0.12	0.08

For Q8, Q12, and Q19, as in the previous queries for the red team role, low precision was observed, along with a very few true positives. This is also a specialized work role with a high number of false positives. A similar situation was observed for the red team where ontology terms used in the construction of the concept were often mentioned in cybersecurity job postings, contributing to the high number of false positives, causing the low precision but also a high recall value and thus F1-scores. The MCC values indicate the results are only slightly better than a random classification.

7.7.7 Stardog classification query results

The queries encapsulated in Table 15 are grouped by the three different SPARQL query models that were used. Group 1 included the queries that used textMatch and the CyberSec003 database with Stardog document markup (Q5 to Q8); group 2 comprised those queries that used textMatch with CyberSec004 with CoreNLP text annotation (Q9 to Q12); lastly, group 3 included those queries that used dc:reference with CyberSec004 with CoreNLP text annotation (Q16 to Q19). This was already presented in the previous tables the results of all these queries groups by query object, which is basically the term that was being searched in the query. As six of the twelve queries had high precision values, Q5 (0.79), Q6 (0.75), Q9 (0.84), Q10 (1.00), Q16 (0.85), Q17 (0.76). The other six of twelve have low precision values, Q7 (0.04), Q8 (0.31), Q11 (0.04), Q12 (0.32), Q18 (0.05) and Q19 (0.33). This test can be considered a success as it can be observed that when there is enough data for different roles, the ontology can be used with queries to reliably categorize the documents. As stated earlier, three different strategies were explored: the default Stardog (Lucene) for group 1, CoreNLP for group 2, and the advanced linker functions for group 3. The results showed that the basic Stardog function and advanced search functions both provide good F1-scores.

Table 15: Summary of query results

Group	Query	Query object	Precision	Recall	F1-score	MCC
1	Q5	Cybersecurity analyst	0.79	0.88	0.85	0.44
	Q6	System security analyst	0.75	0.97	0.85	0.10
	Q7	Red team analyst	0.04	0.93	0.08	0.05
	Q8	Blue team analyst	0.31	0.93	0.33	0.09
2	Q9	Cybersecurity analyst	0.84	0.62	0.53	0.38
	Q10	System security analyst	1.00	0.02	0.04	0.07
	Q11	Red team analyst	0.04	0.92	0.08	0.01
	Q12	Blue team analyst	0.32	0.92	0.71	0.12
3	Q16	Cybersecurity analyst	0.85	0.53	0.65	0.35
	Q17	System security analyst	0.76	0.78	0.78	0.10
	Q18	Red team analyst	0.05	0.5	0.08	−0.09

	Q19	Blue team analyst	0.33	0.47	0.45	0.08

7.8 Work role CV classification

This section presents the results of the queries of the participants' CVs in relation to the same queries that were presented in the previous section. It should be noted that for test involved the public LinkedIn CVs of only 57 participants. This is relatively small, as compared to larger number of job postings in the previous test. As well, as can be observed, the number of CVs that results from the queries did not make it possible to perform the same F1-score analysis shown for the job postings. The calculation could still be done but would provide unreliable results that could not be used in this study, and thus, this was not done. As the number of results was very small, all of them are presented in this section. The data presented results from the same queries that have been presented Q5 to Q8.

Table 16: Cybersecurity analyst role (Q5)

Employee CV code	Profile
#54	Senior Director – Cybersecurity Strategy and Transformation, 14 years cybersecurity experience with multiple graduate degrees, manages teams of analysts
#29	Cybersecurity advisor, eight years cybersecurity experience, post-secondary degree, performs analyst duties
#53	Project manager

This study showed the complete job profile results from this query. In this case, the query results identified three CVs, and all of them were relevant. When the files that correspond to the code were manually verified, it was observed that the LinkedIn profile extract of the participant that corresponded to the work role being searched in this query, i.e., the Cybersecurity Analyst role defined as a class in the ontology resulting from the research. This role is part of the cybersecurity business category.

Table 17: System security analyst role (Q6)

Employee CV code	Profile
#29	Cybersecurity advisor, eight years cybersecurity experience, post-secondary degree, performs system security analyst duties
#54	Senior Director – Cybersecurity Strategy and Transformation, 14 years cybersecurity experience with multiple graduate degrees, manages system security analyst teams
#52	Business Information Security Officer, 18 years cybersecurity experience, with graduate degree, performs system security analyst duties

These results show the complete job profile results from this query on the system security analyst role. In these, three CVs were identified. Two of them were also identified in the previous query and one was not. As the work roles bear high similarity, this was not a surprise. A document verification confirms that the files that correspond to the code that were attributed. It was observed that the LinkedIn profile extract of the participant corresponded to the work role that was searched in this query, the System Security

Analyst role from the NIST NCWF that was defined as a class in the ontology (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017).

Table 18: Red team analyst role (Q7)

Employee CV code	Profile
#29	Cybersecurity advisor, eight years cybersecurity experience, post-secondary degree, and acted as a team lead for the red team
#54	Senior Director – Cybersecurity Strategy and Transformation, 14 years cybersecurity experience with multiple graduate degrees who managed the red team in the past

The complete job profile results from this query on the red team analyst role are shown in the Table 18. The number of results is very limited. When these were verified, the LinkedIn profile extracted corresponded to the work role being searched in this query, the red team analyst role, which was defined as an individual in the ontology. This shows that it is possible to search for classes, as in the previous queries, for individuals in the ontology.

Table 19: Blue team analyst role (Q8)

Employee CV code	Profile
#29	Cybersecurity advisor, eight years of cybersecurity experience, post-secondary degree, and was involved in Blue team for a period.
#54	Senior director – Cybersecurity Strategy and Transformation, 14 years cybersecurity experience with multiple graduate degrees, and was involved as a blue team analyst earlier in their career.
#53	Project manager who is involved in blue team activities

Here as well, the results show the complete job profile results from this query. When the results were manually verified using the files that correspond to the code, it showed that the LinkedIn profile extract of the participant that corresponded to the work role that is searched in this query, the blue team analyst role that was defined as an individual in the ontology.

7.9 Summary of the work role CV classification

In this test, the results were observed to be less revealing than the previous results. The hypothesis is that these results are less significant because of the smaller dataset that was used for the test, in addition to the quality of the data available. If this test had been envisioned at the onset of this study, full-length CVs in Word format could have been obtained directly from the participants, but this was not possible at the time the tests were

added to this study. However, while the results cannot be considered a strong proof, it was shown that the employee CVs that have been identified to contain the work role expertise that was searched. When compared to the previous test with a dataset six times larger, this would seem reasonable.

It would be suggested that implementing a solution in an organizational setting would involve adapting or creating a business process to ensure that job descriptions and CVs are gathered in a standardized format or in a format that would enable these documents to be encoded in an ontology-friendly manner that would markedly improve the processing. An interesting approach using automatic skills standardization approach based on subject matter expert knowledge extraction and semantic matching that could help is based on data science could gather data by crawling job postings from websites (Bernabé-Moreno et al., 2019). Some interesting strategies using average word embeddings and Principal Component Analysis that were used to retrieve CVs on the basis of job description could also be considered (Fernandez-Reyes, 2017). There are also commercial offerings for a job and resumé parsing, such as R-Chilli (https://www.rchilli.com/) or Textkernel (https://www.textkernel.com/) that could be investigated in the future.

Another possible solution that was considered would be to develop a business solution to expedite job posting, and CV integration process would be to create a Java or Python app to integrate the Lucene scoring and automate the process. This could also be combined with Apache Tika, a contents analysis toolkit for developing an in-house solution (https://tika.apache.org/). As this would require a full-fledged study on its own to implement, it was not integrated into the current study, as simple work-around that would satisfy the goals was found.

With the data resulting from the external validation tests, resulting from the queries 4, 5, and 6, based on the initial queries 1, 2, and 3 and comparing to the job postings and CV data from LinkedIn, it was possible to conclude that the ontology does allow to successfully perform the two tests, namely:

1. The ability to identify a job posted on LinkedIn.

2. The ability to identify a cybersecurity role of an individual by his CV.

These further support hypothesis H2. However, the reliability of the CV tests is limited owing to the data quality that was used, as discussed later in this dissertation.

8 Risk scenario test

The purpose of this test was to ascertain the applicability of the proposed ontology as a tool to help Canadian financial institutions match work roles to the requirement of a cybersecurity function. This is one of the external validity tests that were done to explore the applicability of the results of this study to other organizations than the organization in which the study is performed. This test also looks to provide information on hypothesis H3, that the ontology would allow organizations to match work roles to risk scenarios. For performing the test, two risk scenarios based on real-world cases were developed. The cases are neither specific to Canada nor to financial organizations. The scenarios apply to the participating organization but could also find applicability in other large organizations. In this section, these tests and the results are presented. As previous results with the SPARQL queries showed excellent results by using the advanced queries and that these are best adapted to run complex queries using the graphical database capabilities of Stardog, the queries have been structured using these features. The documentation is presented in Appendix C.

8.1 Test protocol

The goal of this test is to perform a one-sample t-test to evaluate the overall accuracy of the classification. Initially, test 1 demonstrated the classification ability of the ontology. From there, several tests were developed. Test 2 investigated the ability to use the ontology with Stardog, to identify potential roles defined as classes and positions to handle a security incident. In test 3, a similar test was performed using an individual in the ontology estimating the quality of the results. In test 4, a sample of 50 queries based on realistic cases, or scenarios, representing possible cybersecurity incidents that could occur in the participating organization were developed using the MITRE ATT&CK cybersecurity framework. By performing the classification multiple times, once per scenario, and comparing the results, made it possible to statistically determine the predictive value of the cybersecurity competency ontology using Stardog and job postings and assess the repeatability of the process. Finally, in test 5, the ontology with Stardog was used to identify potential roles defined as classes and positions to handle a

security incident in an automated manner by only using the scenario tactic from MITRE ATT&CK cybersecurity framework rather than by explicating the mitigation measures as done in test 2, 3, and 4.

This can provide an answer to research question RQ-1.3, indicating that whether ontology can accurately represent the cybersecurity domain and be used as a talent management decision tool.

8.2 Test 1: Using virtual graph and advanced search for classification

The risk scenario tests presented in this dissertation use the advanced linker function combined with the document markup that were included in the CyberSec004 database. This is required with other queries for estimating the predictability value of the ontology using the values of F1-scores. In Stardog, the document store full-text search is used to retrieve the textual contents added to the Stardog database using BITES. The contents of the documents can then be searched for matches against the text elements using the virtual graph capabilities of Stardog, with the Graph command. Virtual graphs enable Stardog to map various data types and sources to the RDF graph of the ontology. In using this feature, the classes and subclasses in the ontology are used for the search and enable ontological matching of the constructs in the data using the textMatch elements with the documents indexed using BITES. This is described int the Stardog documentation in https://docs.stardog.com/virtual-graphs/#querying-virtual-graphs and in Advanced Search, presented in Appendix C. A classification query with only the GRAPH ?doc instruction is presented in Appendix L. This resulted in 292 of 363 documents.

8.3 Test 2: Identification using classes

This first scenario-based test, SC1, was done to evaluate the quality of the results for the intended purpose. As initial research at the onset of this study should indicate what cybersecurity competencies must be part of the ontology, the validation could confirm what parts were accurately reflective of the numerous categories of activities and tasks that could be viewed as attributes of a successful cybersecurity function. These include detailed and structured information about cybersecurity job positions, knowledge, skills,

abilities, soft skills, certifications, education, and others. As such, this part of the testing process focused on using the ontology as an HRM tool, primarily as a tool for talent management and team staffing. This requires the organization to equate the successful individual for a work role by using a combination of the attributes, within certain ranges of acceptable minimal and optimal combinations, given typical cybersecurity scenarios with specific requirements based on real-world situations. Top-10 threats agents and risk scenarios were used to identify potential scenarios from the MITRE ATT&CK framework that could be used for the test to represent likely situation in financial organizations. From these lists' scenarios, SC1 and SC2 were chosen.

In both scenario-based tests, applicability scenarios introduced in a previous section were used as the starting point. Scenario SC1 considered the role of system security analyst, which was also used in queries Q2, Q6, Q10, and Q17. The queries used in the test were along with the model of Q17, which had an F1-score of 0.78, with a precision of 0.76 and a recall of 0.78. In scenario SC1, the study investigated the system security analyst supporting vulnerability mitigation activities as a member of a cyber defense team following a security incident that has occurred. Specifically, the incident considered in SC1 is:

a. An adversary exploiting software vulnerabilities for privilege escalation as described in the MITRE ATT&CK framework https://attack.mitre.org/techniques/T1486/ done to extract monetary compensation from a victim in exchange for decryption or a decryption key, commonly referred to as a ransomware attack.
b. by LockerGoga https://attack.mitre.org/software/S0372/
c. which required application isolation and sandboxing https://attack.mitre.org/mitigations/M1048/ and software update https://attack.mitre.org/mitigations/M1051/

Such an incident would likely be the result of a successful phishing attack, where the threat, in this case LockerGoga, was able to get an employee of the organization to click on a link in an email, which initiated the events leading to the incident. The scenario was

considered on this premise. This scenario was linked to the previous query 2, which was presented, looking for the knowledge required for the role of system security analyst, the results of which are presented in Table 6. In accordance with the MITRE ATT&CK information available online, the main elements of this ransomware incident were identified, namely:

- Changing account passwords and logging off current users;
- Encrypting files and demanding Bitcoin for the decryption key;
- Disabling anti-virus;
- Deleting the original launcher after execution;
- Moving around the victim network and copying files from computer to computer instead of self-propagating;
- Using stolen certificates to make it look more legitimate;
- Shutting down infected systems.

In a system security analyst role in a financial institutional, the expected supporting vulnerability mitigation and cyber defense role following a security incident will be in assistance with the risk mitigation measures. It was indicated in the scenario that it would be more specifically looking to application isolation and sandboxing https://attack.mitre.org/mitigations/M1048/ and software update https://attack.mitre.org/mitigations/M1051/. This would require know-how and know-what of these mitigation measures. The system security analyst will also require project management know-how and software update know-what.

The participants in the study indicated in the interviews that the System Security Analyst supporting vulnerability mitigation and cyber defense role has the following responsibilities:

- Detection of technical process and procedure vulnerabilities
- Analysis and contextualization of the vulnerabilities
- Collaborating with solution experts to determine and implement required corrective actions and measures, including:

    - Business impact analysis
    - Proposing of solutions
    - Making recommendations
    - Advising business lines
    - Validating that the job has been done and the vulnerabilities have been addressed
- Supporting internal and external audits

As the role is the NIST NCWF role OM-ANA-001 system security analyst role, it can directly use the description of the role as including the competency elements that correspond to the organizational requirements (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017). In the NIST NCWF, the competency elements of this role include 31 tasks, 45 knowledge elements, nine skills, and two ability elements (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017). These can be viewed by consulting the class that corresponds to the role OM-ANA-001 system security analyst role directly in Protégé. It can also be retrieved from the ontology using the query Q2 with task, skill, and ability, along with the knowledge already in this query. Of these tasks and KSAs, the tasks that correspond the most to the scenario used for the test were selected. The results of the query in the database for the role OM-ANA-001 system security analyst role are presented in Table 20 for the tasks and Table 21 for the KSAs.

Table 20: Tasks of the system security analyst for scenario 1

Task	Description
T0086	Ensure that the application of security patches for commercial products integrated into system design meet the timelines dictated by the management authority for the intended operational environment.

T0088	Ensure that cybersecurity-enabled products or other compensating security control technologies reduce identified risk to an acceptable level.
T0123	Implement specific cybersecurity countermeasures for systems and/or applications.
T0128	Integrate automated capabilities for updating or patching system software wherever practical and develop processes and procedures for manual updating and patching of system software based on current and projected patch timeline requirements for the operational environment of the system.
T0169	Perform cybersecurity testing of developed applications and/or systems.
T0309	Assess the effectiveness of security controls.
T0344	Assess all the configuration management processes.
T0485	Implement security measures to resolve vulnerabilities, mitigate risks, and recommend security changes to system or system components as needed.
T0545	Work with stakeholders to resolve computer security incidents and vulnerability compliance.

Table 21: KSAs of the system security analyst for the scenario

KSA	Description
K0040	Knowledge of vulnerability information dissemination sources.
K0060	Knowledge of operating systems.
K0290	Knowledge of systems security testing and evaluation methods.
K0339	Knowledge of how to use network analysis tools to identify vulnerabilities.
S0024	Skill in designing the integration of hardware and software solutions.
S0027	Skill in determining how a security system should work (including its resilience and dependability capabilities) and how changes in conditions, operations, or the environment will affect these outcomes.
S0031	Skill in developing and applying security system access controls.
S0036	Skill in evaluating the adequacy of security designs.
S0141	Skill in assessing security systems designs.
S0167	Skill in recognizing vulnerabilities in security systems. (e.g., vulnerability and compliance scanning).

From the information presented in Tables 20 and 21, an SPARQL query that was used to produce the results was developed. The query for SC1:

Query SC1

```
prefix stardogapi: <tag:stardog:api:>

select ?entity ?doc ?mention ?type ?label where

{

   # find something that has a textual match with any one or more of the given string values
?doc stardogapi:property:textMatch

   "System Security Analyst,",

   "Ensure that the application of security patches for commercial products," ,

   "Integrate automated capabilities for updating or patching system software," ,

   "Implement security measures to resolve vulnerabilities, mitigate risks, and recommend security changes to system or system components as needed," ,

   "Knowledge of vulnerability information dissemination sources," ,

   "Knowledge of operating systems," ,

   "Knowledge of systems security testing and evaluation methods," ,

   "Knowledge of how to use network analysis tools to identify vulnerabilities," ,

   "Skill in designing the integration of hardware and software solutions," ,

   "Skill in determining how a security system should work," ,

   "Skill in developing and applying security system access controls," ,
```

```
    "Skill in evaluating the adequacy of security designs," ,

    "Skill in assessing security systems designs," ,

    "Skill in recognizing vulnerabilities in security systems." .

  # get the subgraph (aka named graph) for that something
graph ?doc

  {

      # where it has the property `hasEntity` whose value is something

            # i.e. where BITES has created a new resource matching one or more concepts in
      the database
      ?doc stardog:docs:hasEntity ?entity .

      # and find the mention of that resource by the `references` property
      ?entity <http://purl.org/dc/terms/references> ?mention

   }

           # outside of that subgraph, get the type and label of the found mention

            # (if we do this within the subgraph we may not get any results as these
          statements are not specific to the document but to the database/ontology)
          ?mention a ?type ; rdfs:label ?label

}

# order rows by this variable (default ascending)
Order by ?doc
```

Using this query, the results were exported to a CSV format, loaded into Excel and integrated into the detailed results, which are presented in Appendix A, in the column SC1. As before, a validation of the results was performed to identify true positives, true

negatives, false positives, and false negatives, the summary of which is presented in Table 22.

Table 22: results of Scenario 1 (SC1)

	SC1
No. of documents	363
True positives	165
True negatives	175
False positives	14
False negatives	9
Precision	0.92
Recall	0.95
F1-score	0.93
MCC	0.87

It can be noticed that the results show high precision, recall, F1-scores and MCC. To be certain that there was no error a review of all the documents in relation to the query results was performed, which confirmed that the results were an accurate presentation of the results of the query that was performed. This also supported the researcher's intuition that with a good number of quality documents, quality results could be obtained, in this case by performing the query against the large number of jobs posting in the Stardog database.

8.4 Test 3: Scenario-based identification using individuals

The second scenario-based test investigated the role of a blue team analyst responding to a security incident. In the ontology this role is described by the ontology class blue team, a subclass of the cybersecurity Technical class in the ontology and described by the blue team analyst, defined as an ontology individual. This is linked to query 4 looking for the know-what required for the role of blue team analyst, the results of which are presented in Table 9. The blue team analyst was also used in Q8, Q12, and Q19. Q19 has an F1-score of 0.45, a precision of 0.33, and a recall of 0.47. Specifically, the incident considered in scenario SC2 is:

a. Data encryption for impact as described in the MITRE ATT&CK Framework https://attack.mitre.org/techniques/T1068/
b. by APT28 https://attack.mitre.org/groups/G0007/
c. which required activating the incident management and disaster recovery plans as well as Data Backup https://attack.mitre.org/mitigations/M1053/ and Vulnerability scanning https://attack.mitre.org/mitigations/M1016/

Based on the MITRE ATT&CK data, the main elements of this Ransomware incident have been identified, namely that APT28, a threat group that has been attributed to Russia's General Staff Main Intelligence Directorate (GRU) 85th Main Special Service Center (GTsSS) military unit 26165, has exploited CVE-2014-4076, CVE-2015-2387, CVE-2015-1701 and CVE-2017-0263 to escalate privileges. APT28 is alleged to be the group that compromised the Hillary Clinton campaign, the Democratic National Committee, and the Democratic Congressional Campaign Committee in 2016 to interfere with the U.S. presidential election. This scenario considers that the attackers are possibly intending to disrupt the results of an upcoming Canadian federal election by targeting and attacking Canadian financial institutions. As described in the Common Vulnerabilities and Exposures (CVE) database, the vulnerabilities in the scenario are:

- **CVE-2014-4076:** Microsoft Windows Server 2003 SP2 allows local users to gain privileges via a crafted IOCTL call to (1) tcpip.sys or (2) tcpip6.sys, aka "TCP/IP

Elevation of Privilege Vulnerability." (https://nvd.nist.gov/vuln/detail/CVE-2014-4076)

- **CVE-2015-2387:** ATMFD.DLL in the Adobe Type Manager Font Driver in Microsoft Windows Server 2003 SP2, Windows Vista SP2, Windows Server 2008 SP2 and R2 SP1, Windows 7 SP1, Windows 8, Windows 8.1, Windows Server 2012 Gold and R2, and Windows RT Gold and 8.1 allows local users to gain privileges via a crafted application, aka "ATMFD.DLL Memory Corruption Vulnerability." (https://nvd.nist.gov/vuln/detail/CVE-2015-2387)
- **CVE-2015-1701:** Win32k.sys in the kernel-mode drivers in Microsoft Windows Server 2003 SP2, Vista SP2, and Server 2008 SP2 allows local users to gain privileges via a crafted application, as exploited in the wild in April 2015, aka "Win32k Elevation of Privilege Vulnerability." (https://nvd.nist.gov/vuln/detail/CVE-2015-1701)
- **CVE-2017-0263:** The kernel-mode drivers in Microsoft Windows Server 2008 SP2 and R2 SP1, Windows 7 SP1, Windows 8.1, Windows Server 2012 Gold and R2, Windows RT 8.1, Windows 10 Gold, 1511, 1607, 1703, and Windows Server 2016 allow local users to gain privileges via a crafted application, aka "Win32k Elevation of Privilege Vulnerability." (https://nvd.nist.gov/vuln/detail/CVE-2017-0263)

In this scenario, based on the information provided by the MITRE ATT&CK framework T1068 and the descriptions that are provided in the scenario, a blue team analyst role in a financial institutional, the expected role in dealing with the scenario will be of activating the incident management and disaster recovery plans as well as perform or support data recovery (backup) activities (https://attack.mitre.org/mitigations/M1053/) and perform or coordinate vulnerability scanning and remediation activities (https://attack.mitre.org/mitigations/M1016/).

The participants in the study indicated that the blue team analyst performed tasks related to the cyber defense activities of the financial organization. The role was described as like

that of the system security analyst role used in SC1, presented previously, but added the following KSA to the competency elements:

- Investigative abilities
- Passion and curiosity
- Being a quick learner
- Able to adapt to quickly changing solutions
- Strong technical abilities, including programming, computer systems, servers, operating systems, and networking

In the ontology, the role blue team analyst role is described by the class blue team and by the blue team analyst individual. On this basis, as for the SC1 scenario, the tasks and competency elements, the KSAs, that apply to this scenario and are required to implement the mitigation and response measures identified in the description of the scenario presented were identified. The tasks are presented in Table 22 and the KSAs are presented in Table 23.

Table 22: Tasks of the blue team analyst for the scenario

Task	Description
T0065	Develop and implement network backup and recovery procedures.
T0162	Perform backup and recovery of databases to ensure data integrity.
T0306	Supports incident management, change management, release management continuity management, and availability management for databases and data management systems.
T0477	Ensure the execution of disaster recovery and continuity of operations.

Table 23: KSA of the blue team analyst for the scenario

KSA	Description
K0021	Knowledge of data backup and recovery.
K0026	Knowledge of business continuity and disaster recovery continuity of operations plans.
K0040	Knowledge of vulnerability information dissemination sources.
K0210	Knowledge of data backup and restoration concepts.
K0292	Knowledge of the operations and processes for incident, problem and event management.
K0373	Knowledge of basic software applications (e.g., data storage and backup, database applications) and the types of vulnerabilities that have been found in those applications.
S0001	Skill in conducting vulnerability scans and recognizing vulnerabilities in security systems.
S0158	Skill in operating system administration. (e.g., account maintenance, data backups, maintain system performance, install and configure new hardware/software).

S0242	Skill in interpreting vulnerability scanner results to identify vulnerabilities.
A0015	Ability to conduct vulnerability scans and recognize vulnerabilities in security systems.

From the tasks and KSAs identified the SPARQL query was created, which is presented below:

prefix fts: <tag:stardog:api:property:textmatch:>

prefix stardogapi: <tag:stardog:api:>

select ?entity ?doc ?mention ?type ?label where

{

find something that has a textual match with any one or more of the given string values
?doc stardogapi:property:textMatch

"Blue Team Analyst,",

"Develop and implement network backup and recovery procedures," ,

"Perform backup and recovery of databases to ensure data integrity," ,

"Supports incident management, change management, release management continuity management, and availability management for databases and data management systems,"
,

"Ensure the execution of disaster recovery and continuity of operations," ,

"Knowledge of data backup and recovery," ,

"Knowledge of business continuity and disaster recovery continuity of operations plans," ,

"Knowledge of vulnerability information dissemination sources," ,

"Knowledge of data backup and restoration concepts," ,

"Knowledge of the operations and processes for incident, problem and event management," ,

"Knowledge of basic software applications and the types of vulnerabilities that have been found in those applications," ,

"Skill in conducting vulnerability scans and recognizing vulnerabilities in security systems," ,

"Skill in operating system administration," ,

"Skill in interpreting vulnerability scanner results to identify vulnerabilities," ,

"Ability to conduct vulnerability scans and recognize vulnerabilities in security systems," .

```
 # get the subgraph for that something
 graph ?doc

 {

    # where it has the property `hasEntity` whose value is something

          # i.e. where BITES has created a new resource matching one or more
       concepts in the database
       ?doc stardog:docs:hasEntity ?entity .

    # and find the mention of that resource by the references property
    ?entity <http://purl.org/dc/terms/references> ?mention
```

```
  }

        # outside of that subgraph, get the type and label of the found mention

        # This needs to be done here as if we do this within the subgraph we may not
        get any results as these statements are not specific to the document but to the
        database/ontology
        ?mention a ?type ; rdfs:label ?label

}

Order by ?doc
```

Table 24: results of Scenario 2 (SC2)

	SC2
No. of documents	363
True positives	86
True negatives	241
False positives	19
False negatives	17
Precision	0.82
Recall	0.83

F1-score	0.83
MCC	0.76

This second scenario-based test also resulted in high precision, recall, and F1-scores. In this case, as for SC1, a review was performed and confirmed that the results were accurate. In the results, in addition to the job postings, there was also a CV, document number 29, that was included. You may recall from Q8, presented in Table 19, that this is for a Cybersecurity Advisor, who performs a system security analyst, including in the blue team, which can explain why it appeared in the results.

As mentioned at the beginning of this section, the purpose of this test was to test the applicability of the ontology by matching work roles to the requirement of a cybersecurity function in a scenario representing a realistic risk scenario. In both cases, the results had high precision, recall, F1-scores, and MCC, as presented in Table 25.

Table 25: Scenarios 1 (SC1) and 2 (SC2) summaries

	SC1	**SC2**
Precision	0.92	0.82
Recall	0.95	0.83
F1-score	0.93	0.83
MCC	0.87	0.76

These results are indicative of a successful classification of the documents with a high repeatability. This would indicate a high the level of validity of the ontology in accurately representing the cybersecurity domain. As shown above, using the ontology with Stardog would enable the organization to match work roles and job posting to individuals to perform effective talent management. Using the ontology with the SPARQL queries can form the basis of a decision support tool for financial organizations. To further investigate the effectiveness of the tool, additional tests were performed, which are described in the next section.

8.5 Test 4: F1-scores of scenario-based queries

Further tests were performed to evaluate the predictive qualities of the classification that have been performed with the two MITRE ATT&CK risk scenarios, SC1 and SC2. The use of the first scenario (SC1) is described here as a model to be repeated multiple times to make it possible to perform a one-sample one-tailed t-test. The objective was to compare population mean with predefined value using one sample of data. One-tailed is directional test as this study was focused on determining if the populational mean is above a predetermined value. In this case, we sought to determine if the results obtained using the cybersecurity competency ontology are better that what could be expected in a random choice of 50%. The value for the null hypothesis was therefore set at 55%, or 0.55. As the test showed, the 95% confidence interval that could estimate the true accuracy of classification (F1-score at the population) is 68% ± 5% or 0.68 ± 0.05. We are 95% confident that the true (population level) F1-score is between 0.63 and 0.73. We present the details of this test and how we arrived at these results.

While the initial determination was that 32 tests would be a minimum to achieve a statistically significant sample, once the process was elaborated and a process for the tests developed, 50 tests were executed. After the initial runs, it took only 30 minutes per test. While adding SC1 to this sample, we would end up with 50 observations, which helped improve the confidence interval of the test.

- H0: $\mu \leq 0.55$ null hypotheses states that the classification is not accurate

- H1: $\mu > 0.55$ alternative hypotheses states that the classification is accurate, indicating that the classification is above random chance

As mentioned earlier, a sample of 50 queries was used. These represented the population, meaning that the queries were based on risk scenarios randomly chosen from the 206 scenarios in the MITRE ATT&CK matrix for the enterprise, available online at https://attack.mitre.org/ . To make the selection, a draw was done to randomly select scenarios. All the selected scenarios were validated using the inclusion criteria, listed below:

- The scenario contained mitigation measures' recommendations.
- The scenario was plausible in the participating organization.

For example, a scenario that concerned macOS would be excluded as the participating organization did not use macOS-based computers. In total, three scenarios that were randomly chosen were excluded; had they been included, the sample mean F1-score value from the test would be 0.01 lower at 0.65 and the sample standard deviation 0.16 rather than 0.11.

The resulting 50 queries are described in Appendix D. Using the scenario description from the MITRE TT&CK matrix, the recommendation measures were selected and integrated into the query, as in the example below, which represents query 1 of 50. This query uses the scenario Exploit Public-Facing Application, in the initial access category, https://attack.mitre.org/techniques/T1190/. Table 26 shows the six mitigation measures from the MITRE site. These measures were then used to form the query for this scenario test, and the same process was repeated for all the other scenarios that were included in the test. To avoid SPARQL errors in Stardog, some mitigation measures did require some corrections to remove special characters, such as slash (/) or brackets ([or]). As well, numbered references, such as [4] had to be removed. We also investigated whether the text "System Security Analyst" could skew the results, which we found that removing this part of the query had no effect on the results. However, as the scenario is for a system security analyst involved in the implementation of the mitigation measures that for the

remainder of the query, it was decided to include it, as we had done in the previous test queries that were performed in this study, to ensure consistency.

Mitigations

Mitigation	Description
Application Isolation and Sandboxing	Application isolation will limit what other processes and system features the exploited target can access.
Exploit Protection	Web Application Firewalls may be used to limit exposure of applications to prevent exploit traffic from reaching the application.
Network Segmentation	Segment externally facing servers and services from the rest of the network with a DMZ or on separate hosting infrastructure.
Privileged Account Management	Use least privilege for service accounts will limit what permissions the exploited process gets on the rest of the system.
Update Software	Regularly scan externally facing systems for vulnerabilities and establish procedures to rapidly patch systems when critical vulnerabilities are discovered through scanning and through public disclosure.
Vulnerability Scanning	Regularly scan externally facing systems for vulnerabilities and establish procedures to rapidly patch systems when critical vulnerabilities are discovered through scanning and through public disclosure.[6]

Table 26: mitigation measures from the scenario Exploit Public-Facing Application

From these, the data in the MITRE page, the data that will be used in the query, was extracted. The goal of the query was to identify the documents in the data store that best matched using the ontology the tasks required to be performed by a system security analyst in addressing the recommended risk mitigations measures for this risk scenario. Here is the resultant data:

- Application isolation will limit what other processes and system features the exploited target can access.

- Web application firewalls may be used to limit exposure of applications to prevent exploit traffic from reaching the application.
- Segment externally facing servers and services from the rest of the network with a DMZ or on separate hosting infrastructure.
- Using least privilege for service accounts will limit what permissions the exploited process gets on the rest of the system.
- Regularly scan externally facing systems for vulnerabilities and establish procedures to rapidly patch systems when critical vulnerabilities are discovered through scanning and through public disclosure.
-

This data was used to create the following Stardog query based on the model in the augmented search functionality of Stardog, as presented in Appendix C. The query started by using the entity references to products stored for each document that has been processed with the ontology and added to the Stardog database. This data was then combined with an external data source mapped into the virtual graph database in Stardog providing product details and availability, as already described in Section 13.1. In this specific case, we first identified the textual elements from the documents in the document database using the textMatch command with the ?doc variable to select the entities in the query, with the ?entity variable, and compared them to the virtual graph, using the graph command, executing the query against the documents, which resulted in the best matches being identified in the query results. The results from the query being the best matches between the search terms, in this case the mitigation measures and the job postings that were loaded into Stardog. This was similar to the example proposed in the Stardog documentation, as shown in Appendix C.

```
prefix stardogapi: <tag:stardog:api:>

select ?entity ?doc ?mention ?type ?label where {

  # find something that has a textual match with any one or more of the given
  string values
  ?doc stardogapi:property:textMatch
```

```
"System Security Analyst,",

"Application isolation will limit what other processes and system features
the exploited target can access.",,"

"Web Application application Firewalls firewalls may be used to limit
exposure of applications to prevent exploit traffic from reaching the
application.",,"

"Segment externally facing servers and services from the rest of the
network with a DMZ or on separate hosting infrastructure.",,"

"Use Using least privilege for service accounts will limit what
permissions the exploited process gets on the rest of the system.",,"

"Regularly scan externally facing systems for vulnerabilities and establish
procedures to rapidly patch systems when critical vulnerabilities are
discovered through scanning and through public disclosure.",,"

"Regularly scan externally facing systems for vulnerabilities and establish
procedures to rapidly patch systems when critical vulnerabilities are
discovered through scanning and through public disclosure.".,"

# get the subgraph (aka named graph) for that something
graph ?doc  {

    # where it has the property `hasEntity` whose value is something
    ?doc stardog:docs:hasEntity ?entity .

    # and find the mention of that resource by the `references` property
    ?entity <http://purl.org/dc/terms/references> ?mention }

    # get the type and label of the found mention
    ?mention a ?type ; rdfs:label ?label }

# order rows by this variable (default ascending)
Order by ?doc
```

As described for scenario SC1, the results were exported to a CSV format, using the **run to file** option. The CSV file was then loaded into Excel. It was then necessary to clean the results by deleting the unrequired columns, as the test was only looking at the predictive nature of the basis of the job postings. From there, duplicate rows could be removed, as some job postings could be repeated. This could have been avoided by using the **Distinct** command as we had used in previous queries, but the time taken for processing was less

than the additional time it took to run the queries with this enabled. Finally, the results were sorted. From there, the MITRE ATT&CK validation scenarios query results could be integrated into the detailed results table, which is presented in Appendix E. Likewise the previous instances, a validation of the results was performed to identify true positives, true negatives, false positives, and false negatives. This allowed the calculation of the precision, recall, and F1-scores, the summary of which is presented in Table 27. The same process was repeated for all 50 scenarios shown in Appendix E.

Table 27: Summary of F1-scores

	Precision	**Recall**	**F1-score**	**MCC**
Sample mean	0.77	0.63	0.67	0.06
Sample SD	0.02	0.16	0.13	0.03
Sample variance	0.000	0.027	0.018	0.001

Table 27: Summary of F1-scores mean values for 50 scenarios

As mentioned, in this study, a one-sample one-tailed t-test was performed to compare population mean with the predefined value of 0.55 using one sample of data. In this case, the one-tailed was a directional test as we were interested in population mean being above 0.55. The t value indicates the amount of evidence against the null hypothesis, H0, or in support of H1. The p-value evaluates the probability that null is true, the probability of observing a sample like the one we are evaluating if null was true.

For this, the t and p values were calculated using the Excel spreadsheet:

t =(BD10-0,55)/(BE10/RACINE(50))

t = 6.47

p = =LOI.STUDENT.DROITE(BD14;49)

p-value = 0.0000000221

Since the p-value < 0.05, or the level of statistical significance, 1–0.95 confidence, the null hypothesis (H0) was rejected, and the alternative (H1) was accepted. It can thus be concluded that the classification performed by the ontology using the query is accurate, as the average F1-score > 0.55 with a mean value or 0.67.

A one-sample t-test was performed to evaluate the overall accuracy of the classification. We found that a sample of 50 queries had:

- F1-score = 0.67, which is significantly higher than 0.55
- SD = 0.13
- t(49) = 6,47
- p-value = 0.0000000221

We conclude that the test was successful and the predictive value of the cybersecurity competency ontology using Stardog and job postings, as presented above, is statistically significant. This would support the research question RQ-1.3, indicating that the ontology can accurately represent the cybersecurity domain and be used as a talent management decision tool.

8.6 Test 5: Scenario-based identification of roles

For this test, four (4) MITRE ATT&CK risk scenarios also used in test 4 were used. As previous tests had demonstrated the reliability of the proposed ontology, it was not useful to have a higher number of queries, as is explained later. Four was chosen to demonstrate the repeatability. This test was performed to investigate if the ontology could be used to obtain recommendations for suitable candidates to assist in the implementation of the mitigation measures. It is basically the same as test 4 with an additional nested query. Unlike the previous tests, the mitigations measures related to the MITRE ATT&CK scenarios were explicit in test 5, rather that explicated, as in test 4. In this case, the

researchers only used the MITRE ATT&CK code that corresponds to a specific scenario, which is called a tactic in the framework. For example, in the query described below, attack tactic T1199 is used. This tactic can also be identified as a rdfs:label in the ontology. For this, an SPARQL query was developed, which is presented below.

```
# these are prefixes used throughout the query
PREFIX : <http://webprotege.stanford.edu/> # this is the "default" one, from the ontology
PREFIX fts: <tag:stardog:api:property:textmatch:> # not used here, the property form of search
PREFIX fots: <tag:stardog:api:search:> # used here, the service form of search
PREFIX  dc: <http://purl.org/dc/terms/> # for one of BITES' properties, from a popular vocabulary

SELECT
# show or "project" the variables we are interested in
?tactic ?mitigation ?doc ?score

# group multiple occurrences of a given variable into a single row, with given separator
# see GROUP BY for the grouping criteria (by which a row is projected)
# otherwise all mentions for the same doc would repeat rows for the doc
(GROUP_CONCAT(?mention; SEPARATOR=",") as ?mentions)

        WHERE {
          FILTER(?tactic = "T1199") # provide the tactic label used in the query
          FILTER(?property = :R9FIH0JGU0cP0BhCagaD9vG) # the property URI for isMitigated
          # (this workaround was required since there is a bug using this label in StarDog)

          # The tactic is related to a mitigation in the ontology via this thing called blank nodes, which are what they sound like empty nodes which get random identifiers.

          # When a class is defined with some kind of  "axiom,", say a subclass, a blank node is instantiated.  This is specific to OWL, so the RDF syntax for navigating this complex structure is complex.

          ?subject rdfs:subClassOf ?bNode ;
                rdfs:label ?tactic .

          # in this case, it's not that complex, we can find the mitigation from
```

```
  # the tacting by the subClassOf -> [] -> onProperty|someValuesFrom
chain, where [] is a blank node.

  # the `onProperty` would be the property which was considered a
"relationship" by the ontology, i.e., `isMitigated`, and the
someValuesFrom the value for that property.

  # this is just a class definition which says "tactic has isMitigated
something"

  ?bNode owl:onProperty ?property ;
        owl:someValuesFrom ?value .

  ?value rdfs:label ?mitigation .

  # as we have the mitigation match we need, we can now supply all
matches to the text index and search by them, yielding the documents

  service fots:textMatch {
    [] fots:query ?mitigation ;
       fots:score ?score ;
       fots:result ?doc .
  }

  # when we ask for the graph, we are automatically restricting the result from the
search to documents processed by BITES.

  # each document already has the entities instantiated, so we skip the `hasEntity`
path and directly ask for the reference/mention, which is a concept in the
ontology.

  graph ?doc {
    ?entity dc:references ?reference .
  }

 ?reference rdfs:label ?mention .
}
GROUP BY ?tactic ?mitigation ?doc ?score # group all rows by these
ORDER BY desc(?score) # sort in descending order by this
#LIMIT 10 # this is used to limit the final result when testing the query ONLY
```

While executing the queries, the code of the tactic used was changed for each query. In the example above the variable ?tactic is equal to T1199. This is the value changed in each individual query in test 5. The query result from each query was then sorted and

duplicates were removed, as in previous tests. The tactics T1199, T1569, T1525, and T1078 were used. The test demonstrated that the ontology can be used for the intended purpose. This test also allowed us to identify a few bugs with the Stardog system that made it difficult to execute the queries. Workarounds were finally identified that it was possible to perform test 5. However, to perform prescriptive tests using Stardog, the strategy used in test 4 should be preferred because of its simplicity and reliability. Once this was determined, it became obvious that running a larger number than the four queries already performed would not generate additional useful information for this study.

9 Discussion

The primary aim of this study was to help improve information security management in Canadian financial organizations by contributing to a better alignment between the competencies of actors in cybersecurity work roles and business requirements. Misalignment of the competencies with the cybersecurity work roles required by an organization increases risks, which can result in additional expenses or negative impacts on the organizations. A better understanding of the competency needs of financial institutions is expected to help reduce vulnerabilities. A major contribution of this study is a better understanding of the alignment of these competencies and work roles.

As mentioned, creating a tool based on this ontology that will help match individuals, competencies, competency frameworks, organizational requirements, and obligations, can considerably improve the effectiveness of risk management activities and the efficiency of cybersecurity. As risk and uncertainty are different in nature, developing dynamic capabilities with cybersecurity competencies will help organizations address uncertainty that relate to human aspects of cybersecurity.

The problem of cybersecurity management is also compounded by a growing cybersecurity workers shortage, which is creating additional vulnerabilities and increasing cybersecurity risks (Furnell & Bishop, 2020; ISC2, 2020b; Oltsik, 2019; van Kessel, 2018). Previous studies have confirmed the usefulness of ontologies to help organizations. In particular, this has been demonstrated for cybersecurity (Fenz et al., 2007b), competency gap (Bouras & Zainal, 2016a, 2016b; Fontenele, 2017; Fontenele & Sun, 2016; Zainal, 2017), and other business domains. It is shown here that the ontology may further help organizations, by helping them match the competency elements with the roles as shown with the various queries that can be performed using the ontology which can produce usable data. A better understanding of the contributing elements, such as specific knowledge, know-how, and know-what that make up cybersecurity competency, will contribute to improving dynamic capabilities in organizations, as explained by dynamic capabilities theory, presented in the literature review in Section 3.1 (Eisenhardt & Martin, 2000; Teece et al., 1997). Considering that cybersecurity competencies have

become core competencies of financial institutions, the proposed ontology contributes to the organizations' ability to quickly reconfigure competencies to adapt to its rapidly changing environment, thereby becoming more agile. This contributes to creating and maintaining a competitive advantage, in this case, by helping organizations adapt to rapidly evolving threats and risks. The ontology thus becomes a strategic tool to foster dynamic capabilities in cybersecurity.

Using the ontology, an organization can perform queries, in the same manner as was done in queries 1, 2 and 3, that could identify the chief competency elements, such as the know-how and other elements, that are required for an actor in a particular role. This substantiates hypothesis H2, which states that the cybersecurity competency ontology representing work roles, tasks, and competency elements in Canadian financial institutions can be designed, constructed, and populated. It must be noted that the reliability of the CV tests is limited because of the data quality that is used.

In a similar manner, using the more advanced queries that were demonstrated in tests 4 and 5, an organization can find the best individuals to participate in mitigation activities that concern particular risk scenarios. The findings of this study revealed that matching competency elements to work roles can be done with a high level of precision and recall when there is sufficient data in the database that is used, further supporting hypotheses H1 and H3; the latter states that the ontology would allow organizations to match work roles to risk scenarios. However, it was also noted that when the amount of data or the quality of the data is not sufficient, the results are less reliable. With the low number of CVs, it was not possible to demonstrate reliable results in matching work roles to individuals. However, the larger number of more detailed job posting documents showed excellent results that can be used in an organizational setting. With the job posting, the results have shown that the ontology can be used to analyze risk scenarios to identify the work roles that are best suited to participate in the remediation activities.

There were several challenges confronted at different stages of this study, in particular the global pandemic that started just after completing the data collection and field work. In this regard, the researchers were very lucky, although in a different context, a larger

database of more detailed CVs could perhaps have been collected to investigate the effect of the volume and the quality of documents. However, this was not practicable in the foreseeable future within the research project, and it was necessary to move on. This would be done in the future and a graphical database model based on the ontology, presented in Appendix O, which could be used to help structure large datasets for this purpose. The ontology can become the starting point to develop management information systems to help organizations optimize human resources and cybersecurity capabilities.

9.1 Results and contributions

With regard to the research questions, this study can be deemed successful. First, it was possible to propose an effective approach to develop a new cybersecurity ontology that represents the competencies, skills, and abilities, and effective practices of cybersecurity professionals in a financial institution (RQ-1.1). From this ontology design, it was possible to gather the contents from recognized cybersecurity frameworks, such as the NIST NCWF and from cybersecurity practitioners in the field through interviews and observation (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017). The contents were integrated into an OWL structure, first with WebProtégé and then with Protégé. In this process, it was then possible to create the ontology representing the core competencies of the cybersecurity domain (RQ-1.2) by employing the approach that was defined. Within the limits of the data available, the study supported hypothesis H2, proposing that the cybersecurity competency ontology representing work roles, tasks, and competency elements in Canadian financial institutions can be designed, constructed, and populated. The ontology became the main artifact of this study. The ontology has been presented in this dissertation and can also be viewed online on WebProtégé. As such, the cybersecurity ontology was mapped to cybersecurity competency reference models and the requirements of financial institutions to give a practical use of the ontology for human resource management. A contribution of this project is a simpler model of cybersecurity competencies with two main competency areas, namely business and technical, and nine specialties, to which all the NIST NCWF work roles can be mapped. This simplicity makes it easier to use for cybersecurity management. As well, a reduced number of most

significant competency elements can be mapped to the areas and specialties to facilitate its use.

Once the ontology was completed, it was successfully validated and tested the usability of the cybersecurity competencies in the specific context of the target organization. A series of tests were performed that demonstrated how it can be used by financial organizations to help match individuals in work roles to competencies required to handle various cybersecurity situations, such as incidents. As indicated in Table 25, a good level of validity of the ontology in accurately representing the cybersecurity domain was observed (RQ-1.3). This supports hypothesis H1, as the prediction results obtained using the cybersecurity competency ontology are better than what could be expected in a random choice of 50%. This is shown using anF1-score, as seen in Section 8.4. To achieve this, the ontology was needed to be designed, constructed, and populated, which is described in Section 5. Thus, there is an underlying hypothesis. This study essentially demonstrates that a cybersecurity competency ontology does provide an effective tool for financial institutions to manage cybersecurity talents, supporting our research question (RQ-1). Further work would be needed, as is discussed in Section 4, to develop an operational solution or information system. However, at this point, the current results would indicate that this is feasible and further work in this direction would be justified. As mentioned in Section 9.4, a machine learning solution could use the ontology as a supervised learning tool to further enhance the value of a tool developed using the ontology.

9.2 Advances to dynamic capabilities theory

Improvements in cybersecurity management in financial organizations make the process more efficient and can contribute to creating or maintaining a competitive advantage. A core aspect of managing financial organizations today revolves around the information technologies required to provide the services that are expected by customers. As suggested by hypothesis H4, dynamic capabilities can contribute to effective cybersecurity by creating an organizational capability to rapidly adapt to new and emerging threats and vulnerabilities to an essential component of modern organizations,

its information technologies. Developing dynamic capabilities in cybersecurity, as proposed in this research project by increasing the preparedness of cybersecurity competencies, will help the organization deal with uncertainty. Here are a few examples of how improved competency management, as described here, contributes to improving dynamic capabilities:

- Competency management tools developed using the knowledge produced in developing the ontology and by using the ontology as a source of knowledge makes it possible to develop strategic tools and business processes that can contribute to organizational effectiveness and efficiency, which contribute to the creation of dynamic capabilities.
- Better management of competencies and the adequation of competency levels contribute to having more competent actors, in relation to the work role they play in the organization; actors will be better equipped to cope with the uncertainties of their business environment.
- The cybersecurity competency ontology contributes to creating the conditions to facilitate the continuous evaluation, selection, and implementation of emerging technological innovation and contributes to ensuring that the organization will have the competencies to benefit from innovation.
- With the introduction of machine learning in a future version of the application of the ontology, a capability to dynamically adjust the organizational structure of the organizations and realign competencies to cybersecurity can be envisioned.
- Automation, made possible by using the ontology, could be used to increase the competencies of actors and the firms' human resources in a more dynamic manner, becoming a catalyst of change. Further investigation of this link can also become a future contribution of dynamics capabilities theory.

This study contributes to advance dynamic capabilities by providing evidence of a link between cybersecurity and the creation and maintenance of competitive advantages. In addition, it contributes to creating a culture where a dynamic adaptation of human and material resources involved in cybersecurity can be facilitated. As mentioned previously,

dynamic capabilities contribute to intentionally creating an agile environment to foster adaptability and a strong long-term competitive advantage (Mirabeau & Maguire, 2014; Pfeffer & Sutton, 2006).

In the next section, we present some examples of how early results from this study have already started to be used by the organization by creating dynamic capabilities that can contribute to effective cybersecurity through a capability to rapidly adapt to new and emerging threats and vulnerabilities to an essential component of modern organizations – its information technologies.

9.3 Practical applications

Using the knowledge gained through this study and documented in this dissertation has led the target organization to develop some early tools. In this section, some of the practical applications that can be regarded as benefits of this study are presented. In particular, the following applications are presented:

- Defining cybersecurity work roles in Canadian financial organizations
- Competency evaluation tool
- Identifying training needs

These are presented in the following sections.

9.3.1 Defining cybersecurity work roles in Canadian financial organizations

As described in this dissertation, during the different stages of this study, the knowledge that forms the basic skills elements for all cybersecurity work roles in the organization was defined. This was used by the organization to rethink how these are used, as is presented in this section. At the same time, the level of competency required from each competency element that formed a work role was determined, in reference to the seven that are mentioned in Section 3.2. For example, it was determined that the minimum competency level required of all competency elements for all roles is “Understand” (2). This has led the organization to define this level of 2 for all the base knowledge for the

competency elements common to all cybersecurity work roles as a baseline requirement, which was used to define a common core of training for all new cybersecurity workers, as presented in Section 9.3.3. This base knowledge for all work roles includes the following areas:

- Problem-solving
- Computer network concepts and protocols
- Laws, regulations, policies, and ethics regarding cybersecurity and privacy protection
- Principles of cybersecurity and confidentiality
- Cyber threats and vulnerabilities management
- Cybersecurity standards, management framework, and methodologies
- Risk management
- Change management

In addition to the base knowledge, it was identified that all individuals in cybersecurity work roles must have minimal communication agility and interpersonal competencies at the “Apply” level (3). To ensure that the minimum level of competence is reached in this area, mandatory training in “Communication Agility” was put in place for all team members. From there, it was determined that the work roles could be grouped into two (2) main categories: business and technical, with nine (9) specialties that regroup the 52 NCWF work roles. This was believed to be more manageable in the organization than what was proposed in the NIST NCWF. These work role categories as used in the target organization are described further in the next sections.

9.3.1.1 Cybersecurity business work role category

The cybersecurity business work role category is made up of individuals who are experts in aspects of information security and information systems related to the organization's mission and business objectives. They combine an understanding of technologies with a strong strategic understanding of cybersecurity and understand how cybersecurity brings value to the organization and all its stakeholders. For example, a strategic advisor to a

manager in a business unit who can advise him or her on information security planning in a project is in this category. To ensure that the minimum level of competence is reached, the organization realized that the CISSP training would be the most opportune and cost-effective path, as this was offered by local universities. While developing the ontology, it was shown that there was a good fit between the required competencies and the CISSP training. Thus, it was made mandatory for all actors in the cybersecurity business work role category. Business Cybersecurity 1 and 2 training paths were also created on Udemy to complete the training of all workers.

The main tasks of the cybersecurity business work role category:

- Support and advise information security stakeholders and the organization in terms of logical and physical security to adequately protect information and ensure its confidentiality.
- Help protect the information of customers, employees, and suppliers by developing, maintaining, and monitoring security processes and frameworks. Ensure that it is linked to business processes and technologies to make them more secure.
- Ensure that security requirements are properly considered in all aspects of information management, including management frameworks, architectures, solutions, and systems necessary to achieve the mission of the organization and the delivery of products and services to customers.

The main responsibilities of the cybersecurity business work role category:

1. Determine information protection needs and translate them into requirements while respecting the culture, values, risk appetite, constraints, and business strategy of the organization.
2. Identify, analyze, quantify, prioritize, verify (audit), and document cybersecurity risks.
3. Recommend and follow risk treatment plans for unacceptable risks.
4. Evaluate and design appropriate security management mechanisms.

5. Monitor the evolution of security threats, vulnerabilities, and risks, and make recommendations.
6. Define, operationalize, monitor, verify, measure, and improve security management processes.
7. Define, implement, and monitor strategic, tactical, and operational level security frameworks as per the best practices and information security standards.
8. Provide advice, opinions, and recommendations on security requirements to be included in projects or to be integrated into IT operations.
9. Produce and monitor key performance indicators to measure the achievement of security objectives and ensure accountability to targeted stakeholders.
10. Identify, produce, disseminate, measure, and monitor continuing education and security awareness programs for customers, employees, and managers.

9.3.1.2 Specialties of the cybersecurity business work role

The two main specialties of the cybersecurity business work role category are the cybersecurity analyst and the cybersecurity advisor. It is appropriate to put them in a continuum in connection with the tasks and responsibilities mentioned above. For the skills of the cybersecurity business work roles, the minimum competency level required is “Apply” (level 3). An individual in an analyst role should be able to “Analyze” (level 4) and in the advisory roles should be able to “Evaluate” (level 5) or “Create” (level 6).

Another work was identified in this category, the specialist in cybersecurity awareness and training; although this specialty is more in support of cybersecurity activities of the organization, it is nonetheless critical, as customers, employees, and other stakeholders, form the weakest link in cybersecurity. The manager work role is also identified as a specialty identified in the cybersecurity business work role category. For each of these specialties, the organization identified the detailed competency elements requirements, which are not presented in detail here. We present an overview of these specialties of this cybersecurity business work roles.

Cybersecurity analyst: This is a work role at an intermediate level in the organization. The typical new actor in this work role is a recent graduate from a university undergraduate program in business technology management, computer science, commerce or management who would join a cybersecurity team as an analyst. Likewise, the holder of a vocational college degree in IT who has a few years (5+) of experience in the organization could occupy an analyst position. In addition to the competency elements already identified for all cybersecurity business work roles, the cybersecurity analyst specialist must have competencies at "Analyze" (level 4) in most of the following areas:

- Information systems and operating systems
- Network service management
- Information security technologies
- Personal information protection solutions
- Management of critical infrastructures
- Principles of data classification
- Encryption
- Identity and access management

The following NIST NCWF work roles fall into this specialty:

- Systems Security Analyst (OM-ANA-001)
- Cyber Intel Planner (CO-OPL-001)
- Cyber Ops Planner (CO-OPL-002)
- Authorizing Official/Designating Representative (SP-RSK-001)
- Security Control Assessor (SP-RSK-002)
- Research and Development Specialist (SP-TRD-001)
- Systems Requirements Planner (SP-SRP-001)
- Data Analyst (OM-DTA-002)
- Knowledge Manager (OM-KMG-001)
- IT Program Auditor (OV-PMA-005)

- All Source-Collection Manager (CO-CLO-001)
- All Source-Collection Requirements Manager (CO-CLO-002)

Cybersecurity advisor: The cybersecurity advisor is a senior role. Business information security officers (BISO) are an example of a very senior advisor. This is not a role for a new cybersecurity worker; however, a recent graduate of a university master's program in business technology management, computer science, engineering, cybersecurity, governance (MABM GASTI) or management (MBA) with a good knowledge of the financial sector could join the security teams as an advisor. Likewise, the holder of a baccalaureate or a DEC who has several years (8+) of experience in the financial sector could occupy a position of advisor. In addition to the elements already identified, the cybersecurity advisor specialist must have competencies at "Evaluate" (level 5) or "Create" (level 6) in several of the following areas:

- Strategic business process
- Identification and management of IT needs
- Human–machine Interface
- IT architecture
- Networking and connected services
- Virtualization, containers, and DevSecOps
- Cloud computing
- Emerging threats and vulnerabilities
- Emerging cybersecurity technologies

In addition to all the work roles from the cybersecurity analyst specialty, the following NIST NWCF work roles fall into this category:

- Enterprise Architect (SP-ARC-001)
- Security Architect (SP-ARC-002)
- Cyber Legal Advisor (OV-LGA-001)
- Privacy Officer/Privacy Compliance Manager (OV-LGA-002)
- Cyber Policy and Strategy Planner (OV-SPP-002)

- Program Manager (OV-PMA-001)
- IT Project Manager (OV-PMA-002)
- Product Support Manager (OV-PMA-003)
- IT Investment/Portfolio Manager (OV-PMA-004)

Cybersecurity awareness and training specialist: This is a role at different levels – from junior to senior. A recent graduate from a university certificate program in pedagogy or andragogy could fill this position. The holder of an undergraduate or graduate degree in psychology, sociology, criminology, or public security, ideally with an internship, a capstone project, or a dissertation on an aspect of cybersecurity would be an ideal candidate. Likewise, a cybersecurity analyst or advisor who has several years (5+) of experience as a post-secondary lecturer could fill this position. In this specialty, we also find specialists in training creation, content, and educational management software. The NIST NCWF work roles Cyber Instructional Curriculum Developer (OV-TEA-001) and Cyber Instructor (OV-TEA-002) fall into this category.

9.3.1.3 Cybersecurity technical work roles category

The cybersecurity technical work role category includes IT specialists, programmers, technicians, and engineers who are involved in the various technical aspects, hardware, software, and tools of cybersecurity. They combine strong technical expertise with an understanding of business issues. Some of the individuals in this work role category have IT support or maintenance roles, which include a cybersecurity component. For example, a network manager who configures Microsoft Active Directory services on servers, penetration test specialist, or vulnerability assessment analyst are in the cybersecurity technical work role category. To ensure that the minimum level of competence is reached, the organization determined that the CEH training would be the most opportune and cost-effective path as this was offered by local universities. While developing the ontology, it was shown that there was a good fit between the required competencies and the CEH training. Thus, it was made mandatory training for all actors in the cybersecurity technical work role category. Technical cybersecurity 1 and 2 training paths were also created on Udemy to complete the training of all workers.

The main tasks of the cybersecurity technical work role category:

- Install, configure, and use tools, techniques, and methods for the secure management of data and information systems.
- Identify, analyze, report, and resolve vulnerabilities, events, and incidents that occur or could occur within the network to adequately protect information, information systems, networks, and buildings.

The main responsibilities of the cybersecurity technical work role category:

1. Exploit security tools.
2. Determine the safe operation of a system.
3. Evaluate controls in accordance with the best cybersecurity practices.
4. Apply best practices to organizational requirements.
5. Evaluate the adequacy of cybersecurity measures in organizational requirements.
6. Recognize and classify vulnerabilities and associated attacks.
7. Analyze malicious activity to determine exploited vulnerabilities, methods of exploitation, and impacts.
8. Provide recommendations for threats and vulnerabilities.
9. Recommend vulnerability fixes.
10. Receive and analyze alerts.
11. Identify abnormal activities and potential threats.
12. Document and escalate incidents that may have an impact.
13. Perform event correlation to determine the effectiveness of an attack.
14. Perform cybersecurity research and analysis.
15. Monitor cybersecurity.

9.3.1.4 Specialties of the cybersecurity technical work role category

For all the core competencies of cybersecurity technical work, the minimum level of competency required is level 3. Ideally, all actors in this work role should be able to “Analyze” (level 4) and in the senior work roles “Evaluate” (level 5) and successfully “Create” (level 6). There are several specialties in the technical cybersecurity category,

such as IT exploitation, defensive security, offensive security, security architect, and physical security.

Information technology exploitation includes several management and IT support roles where one of the components of the work role of the actor is related to cybersecurity. Sometimes part of an IT department, outside the main cybersecurity department, this work role is often a gateway for a worker to eventually join the cybersecurity department. This plays a critical role in controlling and managing risk in the organization. A recent graduate of a high school vocational program, post-secondary technical program, or undergraduate program in business technology management, computer science, or IT support could be hired in an IT Exploitation work role. Subsequently, after a few years in the organization this individual could take on a defensive or offensive security position or an analyst role in a cybersecurity department.

The following NIST NCWF work roles fall into this specialty:

- Information Systems Security Developer (SP-SYS-001)
- Systems Developer (SP-SYS-002)
- Database Administrator (OM-DTA-001)
- Technical Support Specialist (OM-STS-001)
- Network Operations Specialist (OM-NET-001)
- System Administrator (OM-ADM-001)
- Exploitation Analyst (AN-EXP-001)
- All-Source Analyst (AN-ASA-001)
- Partner Integration Planner (CO-OPL-003)
- Cyber Operator (CO-OPS-001)

The **defensive cybersecurity specialist** intervenes to proactively protect information systems and data, for example, by identifying and mitigating vulnerabilities. This role is at several levels – from junior to senior roles. A recent graduate of a post-secondary program in business technology management, computer science, or engineering could enter defensive security teams in a junior role. Likewise, the holder of a vocational

college degree in IT who has a few years (5+) of experience in the organization could fill this role. In addition to the elements already identified, the defensive cybersecurity specialist must have a minimum of level 3 competencies in the following areas:

- Application security
- Network security architecture
- OWASP top 10
- MITER ATT&CK
- Microsoft server networks and services (AD, GPO, etc.)
- TCP-IP and derived protocols (DNS, DHCP, LDAP)
- Operationalization of security patches

The following NIST NCWF work roles fall into this specialty:

- Software Developer (SP-DEV-001)
- Secure Software Assessor (SP-DEV-002)
- System Testing and Evaluation Specialist (SP-TST-001)
- Cyber Defense Analyst (PR-CDA-001)
- Cyber Defense Infrastructure Support Specialist (PR-INF-001)
- Cyber Defense Incident Responder (PR-CIR-001)
- Vulnerability Assessment Analyst (PR-VAM-001)
- Threat/Warning Analyst (AN-TWA-001)
- All Source-Collection Manager (CO-CLO-001)
- Cyber Crime Investigator (IN-INV-001)
- Law Enforcement /Counterintelligence Forensics Analyst (IN-FOR-001)
- Cyber Defense Forensics Analyst (IN-FOR-002)

The **offensive cybersecurity specialist** plays the role of an attacker who seeks to test the limits of cybersecurity protections and processes to identify and mitigate unacceptable risks and avoid unwanted incidents. This role is highly know-how-oriented and focuses on the ability to ethically test the limits of our information systems and business processes. It exists on several levels, from N9 to N11. Thus, there are no specific degrees

identified. A high school or post-secondary graduate with plenty of demonstrated talent could find a place in offensive security teams. For example, a finalist in an international cybersecurity competition would be an ideal candidate for this role. In addition to the elements already identified, the defensive cybersecurity specialist must have a minimum of level 4 competencies in the following areas:

- Intrusion techniques
- Planning and execution of attacks

The following NIST NCWF work roles fall into this specialty:

- Mission Assessment Specialist (AN-ASA-002)
- Target Developer (AN-TGT-001)
- Target Network Analyst (AN-TGT-002)
- Multi-Disciplined Language Analyst (AN-LNG-001)

The **cybersecurity architect** consultant is an expert who collaborates in the design and implementation of IT solutions that respect the organization's risk appetite. This is a senior role. The holder of a bachelor’s degree in computer engineering or a vocational college degree in IT who has several years (8+) of experience in IT architecture, ideally in the financial sector, and demonstrated expertise in cybersecurity could occupy a position of Cybersecurity Architect. The Enterprise Architect (SP-ARC-001) and Security Architect (SP-ARC-002) NIST NCWF work roles are in this specialty.

The **physical security specialist** has technical expertise in video surveillance, access control, intrusion detection, incident management, and notification devices. It is a role at different levels, from junior to senior. A graduate in psychology, sociology, criminology, or public safety would be an ideal candidate for this position. Likewise, an individual with IT skills who has several years (5+) of experience in public security, law enforcement, or the military could fill this position. This specialty is not found in the NIST NCWF work roles.

9.3.2 Competency evaluation tool

One of the tools that was developed by the participating organization once the early results were shared with them is a questionnaire-based competency auto-evaluation. Done with Microsoft Forms, this evaluation used the competency elements for the different work roles that were identified in the development of the ontology and presented in this dissertation in the previous sections leading to revised cybersecurity work roles. On the basis of this, a series of questionnaires were created and tested in a small pilot project. An example of one of the questionnaire-based self-evaluation is presented in Appendix P. A first questionnaire captures information about the role, education, and socio-demographic data. Next is a questionnaire for the specific work-role tasks and competency elements. At present, three have been created, one for all the cybersecurity business work roles, one for the cybersecurity defense specialty and one for the cybersecurity offensive specialty. These are followed by questionnaires for the know-how, know-what, and know-how-to-be elements. Once the questionnaires are completed, the data was then made available in a Microsoft Excel format. From there, the data was used in management dashboards created in Microsoft Power BI that enable managers to view the competency coverage for their teams. Eventually, this is intended to be used for continuous education planning, conformity coverage, or other uses currently under investigation. An example of a Power BI dashboard is presented in Appendix Q.

9.3.3 Continuing education in cybersecurity

Another early practical outcome of this research project for the participating organization was the implementation of a cybersecurity continuing education strategy. The strategy focuses on three areas: formal training, active learning, and opportunism, in addition to aiming to increase the competency level of individual actors in their current work roles. It seeks to mitigate potential vulnerabilities caused by the lack of certain competency elements at the required level. It also seeks to increase organizational cybersecurity resilience. The training activities, per specialty and work role, are based on the information that was collected in the research project and integrated into the ontology. This strategy encompasses four levels:

- A common core for all cybersecurity work roles. An internal recognition was created, named the Security Certificate, and was used to recognize the achievement of this common core. The main goal of this training was to ensure a minimum competency level of 2 for all cybersecurity work roles for the essential competency elements, as mentioned in Section 9.3.1.
- A required learning path was then created for each of the two work role categories. The CISSP training was compulsory for all the actors in the cybersecurity business category. CEH training was compulsory for all in the cybersecurity technical category. In addition, targeted learning paths in Udemy and Pluralsight were used.
- A selection of targeted courses was made available, adapted by specialties, for each of the nine specialties that have been presented; these are varied. For example, for the defensive cybersecurity specialty, the organization is in the process of implementing the Immersive Labs Cyber-range training system hands-on training system. Lists of exercises have been targeted by work role, as the vendor of this tool has mapped all the exercises to the NIST NCWF.
- Individual courses have been offered to actors as per an individualized training plan created for each individual actor, at his or her request or at the request of his manager. In addition to considering the current and future needs of the organization, the individualized plan considers the current and future interests of the employee.

9.4 Limitations of this study

There are limitations to the effectiveness of qualitative methodologies in general and cybersecurity matters in particular (Richter & Koch, 2004). Cybersecurity is a sensitive topic for any organization and introduces limits to what it can allow to be published and made public. At the same time, the organization and researchers recognize that as this is an important subject that has national and strategic interests, there are good reasons for the research to be conducted and published. This study was able to ensure the scientific merit of the endeavor. Mechanisms to control biases and adherence to a scientific

research methodology, ADR, were part of the effort to maximize validity. Notwithstanding, to obtain permission from the participating organization, some information was only shared with the research team and was not published in the results, the dissertation, or subsequent articles in scientific journals at the request of the organization. This has introduced limits to this study that will be documented but not necessarily published. Without this, however, the research could not take place.

Moreover, this study had certain limitations inherent to formal validation and testing processes (Whittemore et al., 2001). The researchers were confronted with a difficult challenge in part because of the difficult balance to maintain between rigor, subjectivity, and creativity. The ongoing peer review process involving the research team, including the academic dissertation supervisors, the subject matter experts within the organization, and the informants, contributes to ensuring the integrity, authenticity, and credibility of the results, thereby allowing descriptive and interpretative quality. At the same time, the choice of ADR allowed us to maximize the congruence and thoroughness of the results. These all contribute to the internal validity of the study, faithfully presenting and interpreting the reality of the financial institution where this study was conducted and enabling the researchers to develop an understanding of the situation that led to successfully developing and testing the proposed solution. As to the notion of external validity, because of the implication of industry players, there are some aspects of the solution that could be generalized, and this applied to other similar financial organizations in Canada. However, this would require further studies and additional empirical research.

9.5 Ethical considerations

This study did not include experimentation on human beings, which does not indicate the absence of ethical concerns. The qualitative approach in research being adopted in organizations requires the establishment of trust between the research team and the participants in the study. This trust cannot exist without respect for the individuals involved. The researchers in the study needed to remain attentive and sensitive to the values and culture of the participants and the organization. Trust made it possible to

acquire access to the data of the organization and the internal perspective of participants that formed the raw knowledge that was required to successfully execute the study.

This study was submitted to and approved by the UQO Research Ethics Committee at the start of the project. In addition, informed consent was obtained in writing from the participants when they were enrolled in the study before the data collection process could start. This was done using the consent form that had been prepared and approved. The form clarified to the participants that their participation in this study was absolutely free and informed. Participation in this study was voluntary and that the participants had the option to withdraw at any time. A few participants participated only in some phases of the study. In particular, a few participants in the initial interviews moved on to another organization during the study. However, their contribution had ended at that time, and they had no objection to the use of the information they provided be included in the study. Risk management measures were in place in accordance with UQO regulations and processes to protect the integrity of the participants and adequately protect their personal data. The participants of the study were identified by a code, and their names will not appear in any document. If the results of the study are published, then no name, code, and initial will be released. Furthermore, as per the confidentiality agreement to perform the study, no information that can make it possible to identify the organization is included in this dissertation or in any subsequent publications.

Only the researchers and persons mandated by the ethics committee will have access to the files for adequately monitoring this study, and this is in accordance with a strict confidentiality policy. The documents used for the study will be kept for five years and destroyed thereafter.

9.6 Future work

This study resulted in a completed cybersecurity competency ontology that describes the work roles of cybersecurity workers in a Canadian financial institution. This allows the institution to improve information security by reducing the vulnerabilities that can result from competency gaps and other benefits mentioned. Once the study was completed, a

future direction could involve the production of competency metrics, indicators that would contribute to limit the subjectivity and biases in allocating human and economic resources to mitigate the risks created by cybersecurity competency gaps and vulnerabilities.

Future work with this study would be to explore the integration of machine learning or other components to automate HR systems, create management information systems that can help match potential candidates for work roles using big data sources, such as LinkedIn or other data sources. Using the ontology as the basis of the solution could help human resources departments in organizations narrow down the lists to identify individuals who could be potential recruits, including non-traditional candidates, such as minorities or other underrepresented groups. Furthermore, there is a potential to identify competent individuals who can emerge from other sources than the usual academic and degree-granting profiles and have acquired competencies that could be recognized and proven.

The ontology can also provide a tool for organizations to assist with legal and regulatory compliance issues. Informal discussions and presentations to industry interest groups, such as the Canadian Bankers Association, have generated interest in pursuing this avenue in the future. This can also lead to the design of similar cybersecurity competency ontologies in many other fields with strong compliance requirements, such as public utilities.

10 Conclusion

In this research, as presented in this research proposal, the “action design research methodology,” ADR, was used to develop and test an ontology of cybersecurity professional skills for the financial services industry in Canada. With the help of a panel of experts, the study successfully combined renowned frameworks, such as the NIST NICE, bodies of knowledge (Newhouse, Keith, Scribner, & Witte, 2017b; NIST, 2021; Petrella, 2017), and current best practices with the actual in vivo experiences of the cybersecurity practitioners of a world-class Canadian financial institution working with a team of academic researchers.

This allowed the researchers to design, develop, populate, and test a cybersecurity competency ontology representing the actual need for competencies of financial organizations that are required to fulfill its mission successfully. How successfully this ontology can assist them is one of the elements that were tested in the field. Nonetheless, the cyclical iterative nature of ADR should allow us to emerge from the study with a useful tool that can be further improved when it is implemented. The reflective nature of ADR and the implication of members of the organization provided additional benefits by helping create a culture of security and life-long learners.

As IT is such an important component of creating a competitive advantage for financial institutions, cybersecurity has become a crucial topic. In addition, various challenges faced in today’s world, such as pandemics, increase cybercrime cases, the reduction of the number of available, competent talents, and the number of many other issues increase the importance of cybersecurity and the need for this study. This is a critical issue for which organizations need solutions.

11 Bibliography

Abawajy, J. 2014. User preference of cyber security awareness delivery methods. ***Behaviour & Information Technology***, 33(3): 237–248.

Agrawal, A., Finnie, G., & Krishnan, P. 2010. A General Framework to Measure Organizational Risk during Information Systems Evolution and its Customization. ***Journal of Research & Practice in Information Technology***, 42(1): 37–60.

Ahmed, N., & Abraham, A. 2013. Modeling Security Risk Factors in a Cloud Computing Environment. ***Journal of Information Assurance and Security***, 8(2013): 279–289.

Ahmed-Kristensen, S., Kim, S., & Wallace, K. 2007. A Methodology for Creating Ontologies for Engineering Design. ***Transactions of the ASME of Journal of Computing in Information Science in Engineering***, 7(June): 132–140.

Allodi, L., & Massacci, F. 2017. Security Events and Vulnerability Data for Cybersecurity Risk Estimation. ***Risk Analysis***, 37(8): 1606–1627.

Alter, S., & Sherer, S. A. 2004. A General, But Readily Adaptable Model of Information System Risk. ***Communications of the Association for Information Systems***, 14. https://doi.org/10.17705/1CAIS.01401.

Amarachi, A. A., Okolie, S. O., & Ajaegbu, C. 2013. Information security management system: Emerging issues and prospect. ***IOSR Journal of Computer Engineering***, 12(3): 96–102.

Apache Lucene Search. 2021. ***Apache Lucene Search (Lucene 7.4.0 API)***. https://lucene.apache.org/core/7_4_0/core/org/apache/lucene/search/package-summary.html#scoringBasics.

Armstrong, M. E., Jones, K. S., Namin, A. S., & Newton, D. C. 2018. The Knowledge, Skills, and Abilities Used by Penetration Testers: Results of Interviews with Cybersecurity Professionals in Vulnerability Assessment and Management. ***Proceedings of the Human Factors and Ergonomics Society Annual Meeting***, 62(1): 709.

Asim, M. N., Wasim, M., Khan, M. U. G., Mahmood, W., & Abbasi, H. M. 2018. A survey of ontology learning techniques and applications. ***Database***, 2018(bay101). https://doi.org/10.1093/database/bay101.

Asosheh, A., Hajinazari, P., & Khodkari, H. 2013. A practical implementation of ISMS. ***7th International Conference on e-Commerce in Developing Countries:with focus on e-Security***, 1–17. Presented at the 7th International Conference on e-Commerce in Developing Countries:with focus on e-Security.

Babb, S. 2014. Managing the Risk Portfolio Using COBIT 5. ***COBIT Focus***, 1–2.

Bacigalupo, M., Kampylis, P., McCallum, E., & Punie, Y. 2016. ***Promoting the entrepreneurship competence of young adults in Europe: Towards a self-assessment tool***, 611–621. Presented at the International Technology, Education and Development Conference, Seville, Spain.

Bahli, B., & Rivard, S. 2003. The information technology outsourcing risk: A transaction

cost and agency theory-based perspective. ***Journal of Information Technology (Routledge, Ltd.)***, 18(3): 211–221.

Bakshi, S. 2012. Risk IT Framework for IT Risk Management: A Case Study of National Stock Exchange of India Limited. ***COBIT Focus***, 1: 5–10.

Banham, R. 2017. Cybersecurity: A new engagement opportunity. ***Journal of Accountancy***. https://www.journalofaccountancy.com/issues/2017/oct/cybersecurity-engagement-for-cpas.html.

Bannerman, P. L. 2008. Risk and risk management in software projects: A reassessment. ***Journal of Systems and Software***, 81(12): 2118–2133.

Beck, J. B., & Wiersema, M. F. 2013. Executive Decision Making: Linking Dynamic Managerial Capabilities to the Resource Portfolio and Strategic Outcomes. ***Journal of Leadership & Organizational Studies***, 20(4): 408–419.

Bell, M. R. S., University, K. S., Vasserman, E. Y., University, K. S., Sayre, E. C., et al. 2014. A Longitudinal Study of Students in an Introductory Cybersecurity Course. ***2014 ASEE Annual Conference & Exposition***, 11.

Benaroch, M., Chernobai, A., & Goldstein, J. 2012. An internal control perspective on the market value consequences of IT operational risk events. ***International Journal of Accounting Information Systems***, 13(4): 357–381.

Bernabé-Moreno, J., Tejeda-Lorente, Á., Herce-Zelaya, J., Porcel, C., & Herrera-Viedma, E. 2019. An automatic skills standardization method based on subject expert knowledge extraction and semantic matching. ***Procedia Computer Science***, 162: 857–864.

Beucher, S., Veyret, Y., & Reghezza, M. 2004. ***Les risques***. Editions Bréal.

Bevilacqua, M., & Ciarapica, F. E. 2018. Human factor risk management in the process industry: A case study. ***Reliability Engineering & System Safety***, 169: 149–159.

Bonollo, M., & Massimiliano, N. 2012. Data quality in banking: Regulatory requirements and best practices. ***Journal of Risk Management in Financial Institutions***, 5(2): 146–161.

Bouras, A. A., & Zainal, A. A. 2016a. Education Ontology Modeling for Competency Gap Analysis. ***3rd IEEE International Conference on Computational Science and Computational Intelligence (CSCI 2016)***. Las Vegas, United States. https://hal.archives-ouvertes.fr/hal-01483305.

Bouras, A. A., & Zainal, A. A. 2016b. Education Ontology Modeling for Competency Gap Analysis. ***3rd IEEE International Conference on Computational Science and Computational Intelligence (CSCI 2016)***. Las Vegas, United States. https://hal.archives-ouvertes.fr/hal-01483305.

Boyatzis, R. E. 2008. Competencies in the 21st century. (R. Boyatzis, Ed.)***Journal of Management Development***, 27(1): 5–12.

Bradfield, R., Wright, G., Burt, G., Cairns, G., & Van Der Heijden, K. 2005. The origins and evolution of scenario techniques in long range business planning. ***Futures***, 37(8): 795–812.

Cable, D. M., & Parsons, C. K. 2001. Socialization tactics and person-organization fit. ***Personnel Psychology***, 54(1): 1–23.

Callen-Naviglia, J., & James, J. 2018. Fintech, regtech and the importance of cybersecurity. ***Issues in Information Systems***, 19(3): 220–225.

Camillo, M. 2017. Cybersecurity: Risks and management of risks for global banks and financial institutions. ***Journal of Risk Management in Financial Institutions***, 10(2): 196.

Champion, M. A., Rajivan, P., Cooke, N. J., & Jariwala, S. 2012. Team-based cyber defense analysis. ***2012 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support***, 218–221. Presented at the 2012 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA 2012), New Orleans, LA, USA: IEEE.

Chicco, D., & Jurman, G. 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. ***BMC Genomics***, 21. https://doi.org/10.1186/s12864-019-6413-7.

Chicco, D., Tötsch, N., & Jurman, G. 2021. The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. ***BioData Mining***, 14(1): 13.

Chornous, G., & Ursulenko, G. 2013. Risk Management in Banks: New Approaches to Risk Assessment and Information Support. ***Ekonomika / Economics***, 92(1): 120.

Cleveland, M., & Spangler, T. 2018. Toward a Model for Ethical Cybersecurity Leadership. ***International Journal of Smart Education and Urban Society (IJSEUS)***, 9(4): 29–36.

Cleveland, S., & Cleveland, M. 2018. Toward cybersecurity leadership framework. ***Proceedingsof the 13th Midwest Association for Information Systems Conference, St. Louis, MO, May***, 17–18.

Common Vulnerabilities and Exposures. 2020. ***CVE***. https://cve.mitre.org/.

Cooper, M. D. 2000. Towards a model of safety culture. ***Safety Science***, 36(2): 111–136.

Cox S & Flin R. 1998. Safety culture: Philosopher's stone or man of straw? ***Work & Stress***, 12(3): 189.

Crichton, D. 2002. UK and Global insurance responses to flood hazard. ***Water International***, 27: 119–131.

CyberSeek. 2021. ***Cybersecurity Supply And Demand Heat Map***. https://www.cyberseek.org/heatmap.html.

Dawson, J., & Thomson, R. 2018. The Future Cybersecurity Workforce: Going Beyond Technical Skills for Successful Cyber Performance. ***Frontiers in Psychology***, 9. https://doi.org/10.3389/fpsyg.2018.00744.

Dawson, M. E. J., Crespo, M., & Brewster, S. 2013. DoD cyber technology policies to secure automated information systems. ***International Journal of Business Continuity***

and Risk Management, 4(1): 1.

De Haes, S., & Van Grembergen, W. 2009. An Exploratory Study into IT Governance Implementations and its Impact on Business/IT Alignment. ***Information Systems Management***, 26(2): 123–137.

De Haes, S., Van Grembergen, W., & Debreceny, R. S. 2013a. COBIT 5 and Enterprise Governance of Information Technology: Building Blocks and Research Opportunities. ***Journal of Information Systems***, 27(1): 307–324.

De Haes, S., Van Grembergen, W., & Debreceny, R. S. 2013b. COBIT 5 and Enterprise Governance of Information Technology: Building Blocks and Research Opportunities. ***Journal of Information Systems***, 27(1): 307–324.

Disterer, G. 2013. ISO/IEC 27000, 27001 and 27002 for Information Security Management. ***Journal of Information Security***, 04(02): 92–100.

Douglas, M., & Wildavsky, A. 1983. ***Risk and culture: An essay on the selection of technological and environmental dangers***. Univ of California Press.

Draganidis, F., & Mentzas, G. 2006. Competency based management: A review of systems and approaches. ***Information Management & Computer Security***.

Draksler, T. Z., & Širec, K. 2018a. Conceptual Research Model for Studying Students' Entrepreneurial Competencies. ***Konceptualni Raziskovalni Model Za Raziskovanje Podjetniških Kompetenc Študentov.***, 64(4): 23–33.

Draksler, T. Z., & Širec, K. 2018b. Conceptual Research Model for Studying Students' Entrepreneurial Competencies. ***Konceptualni Raziskovalni Model Za Raziskovanje Podjetniških Kompetenc Študentov.***, 64(4): 23–33.

Edgar, T. W., & Manz, D. O. 2017. ***Research methods for cyber security***. Syngress.

Eisenhardt, K. M., & Martin, J. A. 2000. Dynamic capabilities: What are they? ***Strategic Management Journal***, 21(10–11): 1105–1121.

Elahi, G., Yu, E., & Zannone, N. 2010. A vulnerability-centric requirements engineering framework: Analyzing security attacks, countermeasures, and requirements based on vulnerabilities. ***Requirements Engineering***, 15(1): 41–62.

Eloumri. 2019. ***Integrating semantic reasoning in a multi-agent system in erlang***. Université du Québec en Outaouais.

Elshahat, A., Parhizgari, A., & Hong, L. 2012. The information content of the banking regulatory agencies and the depository credit intermediation institutions. ***Journal of Economics and Business***, 64(1): 90–104.

Ergashev, B. 2012. A Theoretical Framework for Incorporating Scenarios into Operational Risk Modeling. ***Journal of Financial Services Research***, 41(3): 145–161.

Ericson, M. 2014. On the dynamics of fluidity and open-endedness of strategy process toward a strategy-as-practicing conceptualization. ***Scandinavian Journal of Management***, 30(1): 1–15.

Fenz, S., Goluch, G., Ekelhar, A., Riedl, B., & Weippl, E. 2007a. Information Security

Fortification by Ontological Mapping of the ISO/IEC 27001 Standard. ***13th Pacific Rim International Symposium on Dependable Computing (PRDC 2007)***, 381–388. Presented at the 13th Pacific Rim International Symposium on Dependable Computing (PRDC 2007), Melbourne, Australia: IEEE.

Fenz, S., Goluch, G., Ekelhar, A., Riedl, B., & Weippl, E. 2007b. Information Security Fortification by Ontological Mapping of the ISO/IEC 27001 Standard. ***13th Pacific Rim International Symposium on Dependable Computing (PRDC 2007)***, 381–388. Presented at the 13th Pacific Rim International Symposium on Dependable Computing (PRDC 2007), Melbourne, Australia: IEEE.

Fernandes, C., Ferreira, J. J., Raposo, M. L., Estevão, C., Peris-Ortiz, M., et al. 2017. The dynamic capabilities perspective of strategic management: A co-citation analysis. ***Scientometrics***, 112(1): 529–555.

Fernandez-Reyes, F. C. 2017, October 16. Using Machine Learning to Retrieve Relevant CVs Based on Job Description. ***Dzone.com***. https://dzone.com/articles/cv-r-cvs-retrieval-system-based-on-job-description.

Fontenele, M. P. 2017, May 31. ***Designing a method for discovering expertise in cyber security communities: An ontological approach***. phd, University of Reading. http://centaur.reading.ac.uk/71325/.

Fontenele, M., & Sun, L. 2016. Knowledge management of cyber security expertise: An ontological approach to talent discovery. ***2016 International Conference On Cyber Security And Protection Of Digital Services (Cyber Security)***, 1–13. Presented at the 2016 International Conference On Cyber Security And Protection Of Digital Services (Cyber Security), London, United Kingdom: IEEE.

Frelinger, B. 2012. Building Acceptance and Adoption of Governance of Enterprise IT. ***COBIT Focus***, 1: 1–3.

Furnell, S., & Bishop, M. 2020. Addressing cyber security skills: The spectrum, not the silo. ***Computer Fraud & Security***, 2020(2): 6–11.

Galliano, J. S. 2017. ***Improved Matching of Cybersecurity Professionals' Skills to Job-Related Competencies: An Exploratory Study***. SSRN Scholarly Paper no. ID 3076897, Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=3076897.

Gavrilova, T., Leshcheva, I., & Strakhovich, E. 2015. Gestalt principles of creating learning business ontologies for knowledge codification. ***Knowledge Management Research & Practice***, 13(4): 418–428.

GDPR. 2021. ***General Data Protection Regulation (GDPR)***. https://gdpr-info.eu/.

Gilbert, A. L. 2000. Using multiple scenario analysis to map the competitive futurescape: A practice-based perspective. ***Competitive Intelligence Review***, 11(2): 12–19.

Gomez-Perez, A. 1998. Ontological engineering: A state of the art. ***ResearchGate***. https://www.researchgate.net/publication/50600362_Ontological_Engineering_A_state_of_the_Art.

Government of Canada, S. C. 2019, March 20. ***The labour force in Canada and its regions: Projections to 2036***. https://www150.statcan.gc.ca/n1/pub/75-006-x/2019001/article/00004-eng.htm.

Grimm, S., Abecker, A., Völker, J., & Studer, R. 2011. Ontologies and the Semantic Web. In J. Domingue, D. Fensel, & J. A. Hendler (Eds.), ***Handbook of Semantic Web Technologies***: 507–579. Berlin, Heidelberg: Springer.

Gruber, T. R. 1993. A translation approach to portable ontology specifications. ***Knowledge Acquisition***, 5(2): 199–220.

Gruber, T. R. 1995. Toward principles for the design of ontologies used for knowledge sharing? ***International Journal of Human-Computer Studies***, 43(5–6): 907–928.

Guldenmund, F. W. 2000. The nature of safety culture: A review of theory and research. ***Safety Science***, 34(1–3): 215–257.

Hand, D. J., Christen, P., & Kirielle, N. 2021. F*: An interpretable transformation of the F-measure. ***Machine Learning***, 110(3): 451–456.

Hannah, S. T., Jennings, P. L., Bluhm, D., Peng, A. C., & Schaubroeck, J. M. 2014. Duty orientation: Theoretical development and preliminary construct testing. ***Organizational Behavior and Human Decision Processes***, 123(2): 220–238.

Herjavec. 2020. The 2019/2020 Official Annual Cybersecurity Jobs Report. ***Herjavec Group***. https://www.herjavecgroup.com/2019-cybersecurity-jobs-report-cybersecurity-ventures/.

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

Hiller, J. S., & Russell, R. S. 2013. The challenge and imperative of private sector cybersecurity: An international comparison. ***Computer Law & Security Review***, 29(3): 236–245.

Huang, L., & Zhu, Q. 2019. Strategic Learning for Active, Adaptive, and Autonomous Cyber Defense. ***Adaptive Autonomous Secure Cyber Systems***. https://doi.org/10.1007/978-3-030-33432-1_10.

Humphreys, T. 2006. State-of-the-art information security management systems with ISO/IEC 27001:2005. ***ISO Management Systems***, 4.

Hunton, J. E. 1, Wright, A. M. 2, & Wright, S. 2004. Are Financial Auditors Overconfident in Their Ability to Assess Risks Associated with Enterprise Resource Planning Systems? ***Journal of Information Systems***, 18(2): 7–28.

Infosec. 2021. ***2021 Cybersecurity Role & Career Path Clarity Study***. Infosec.

Ioannidis, C., Pym, D., & Williams, J. 2012. Information security trade-offs and optimal patching policies. ***European Journal of Operational Research***, 216(2): 434–444.

Isaca. 2009. ***The Risk IT Framework***. ISACA. https://books.google.ca/books?hl=fr&lr=lang_en|lang_fr&id=tG7VMihmwtsC&oi=fnd&pg=PA7&dq=%22risk+scenario%22+%22IT+risk%22+%22risk+evaluation%22&ots=TI

p9XS7Itl&sig=i6yx4d9Bh4VUIfKzjw6gnhnXti8.

ISC2. 2020a. ***2019 Cybersecurity Workforce Study***. https://www.isc2.org:443/Research/2019-Cybersecurity-Workforce-Study.

ISC2. 2020b. ***Strategies for Building and Growing Strong Cybersecurity Teams, (ISC)2 CYBERSECURITY WORKFORCE STUDY, 2019***. https://www.isc2.org/-/media/ISC2/Research/2019-Cybersecurity-Workforce-Study/ISC2-Cybersecurity-Workforce-Study-2019.ashx.

Iszatt-White, M. 2010. Strategic leadership: The accomplishment of strategy as a 'perennially unfinished project.' ***Leadership***, 6(4): 409–424.

Jarzabkowski, P., & Kaplan, S. 2015. Strategy tools-in-use: A framework for understanding "technologies of rationality" in practice. ***Strategic Management Journal***, 36(4): 537–558.

Jarzabkowski, P., & Paul Spee, A. 2009. Strategy-as-practice: A review and future directions for the field. ***International Journal of Management Reviews***, 11(1): 69–95.

Jirasek, V. 2012. Practical application of information security models. ***Information Security Technical Report***, 17(1–2): 1–8.

Joshi, A., Bollen, L., & Hassink, H. 2013. An Empirical Assessment of IT Governance Transparency: Evidence from Commercial Banking. ***Information Systems Management***, 30(2): 116–136.

Keet, C. M. 2020. ***An Introduction to Ontology Engineering***.

Keijzer-Broers, W., & de Reuver, M. 2016. Action Design Research for Social Innovation: Lessons from Designing a Health and Wellbeing Platform. ***Proceedings of International Conference on Information Systems (ICIS)***, 1–20.

Krathwohl, D. R. 2002. A revision of Bloom's taxonomy: An overview. ***Theory into Practice***, 41(4): 212–218.

Křemen, P., Mička, P., Blaško, M., & Šmíd, M. 2012. Ontology-driven mindmapping. ***Proceedings of the 8th International Conference on Semantic Systems - I-SEMANTICS***, 39(12). https://www.academia.edu/7752095/Ontology-driven_mindmapping.

Kukulies, J., Falk, B., & Schmitt, R. H. 2016. Development of Optimized Test Planning Procedures for Stabilizing Ramp-up Processes by Means of Design Science Research. ***Procedia CIRP***, 51: 93–98.

Léger, M.-A. 2001. ***Introduction à la gestion de risque informationnel***. Montréal: Fondation de recherche Léger.

Léger, M.-A. 2003. ***Un processus d'analyse des vulnérabilités technologiques comme mesure de protection contre les cyber-attaques, Rapport d'activité de synthèse***. UQAM, Montréal.

Leitner, A., & Schaumuller-Bichl, I. 2009. ARIMA-A new approach to implement ISO/IEC 27005. ***Logistics and Industrial Informatics, 2009. LINDI 2009. 2nd***

International, 1–6. IEEE.

Leung, R. 2018. Cybersecurity regulation in the banking sector: Global emerging themes. ***ResearchGate***. https://www.researchgate.net/publication/328419728_Cybersecurity_regulation_in_the_banking_sector_global_emerging_themes.

Maisey, M. 2014. Moving to analysis-led cyber-security. ***Network Security***, 2014(5): 5–12.

Man, T. W. Y., Lau, T., & Chan, K. F. 2002a. The competitiveness of small and medium enterprises A conceptualization with focus on entrepreneurial competencies$. ***Journal of Business Venturing***, 20.

Man, T. W. Y., Lau, T., & Chan, K. F. 2002b. The competitiveness of small and medium enterprises A conceptualization with focus on entrepreneurial competencies$. ***Journal of Business Venturing***, 20.

Marcelo, A., Rodríguez, A., & Trucharte, C. 2008. Stress tests and their contribution to financial stability. ***Journal of Banking Regulation***, 9(2): 65–81.

McClelland, D. C. 1973. Testing for competence rather than for" intelligence.". ***American Psychologist***, 28(1): 1.

McCurdy, N., Dykes, J., & Meyer, M. 2016. Action Design Research and Visualization Design. ***Proceedings of the Beyond Time and Errors on Novel Evaluation Methods for Visualization - BELIV '16***, 10–18.

Mcube, U. U. 2017. ***A scenario-based ICT risk assessment in local government***. Nelson Mandela Metropolitan University. http://ir.nrf.ac.za/handle/10907/1302.

Mirabeau, L., & Maguire, S. 2014. From autonomous strategic behavior to emergent strategy: From Autonomous Strategic Behavior to Emergent Strategy. ***Strategic Management Journal***, 35(8): 1202–1229.

Mitchelmore, S., & Rowley, J. 2010. Entrepreneurial competencies: A literature review and development agenda. ***International Journal of Entrepreneurial Behavior & Research***, 16(2): 92–111.

MITRE. 2021. ***MITRE ATT&CK***. https://attack.mitre.org/.

Mullarkey, M. T., & Hevner, A. R. 2019. An elaborated action design research process model. (P. Ågerfalk, Ed.)***European Journal of Information Systems***, 28(1): 6–20.

Musman, S. 2016. Assessing prescriptive improvements to a system's cyber security and resilience. ***2016 Annual IEEE Systems Conference (SysCon)***, 1–6. Presented at the 2016 Annual IEEE Systems Conference (SysCon).

Newhouse, W., Keith, S., Scribner, B., & Witte, G. 2017a. ***National Initiative for Cybersecurity Education (NICE) Cybersecurity Workforce Framework***. no. NIST SP 800-181, Gaithersburg, MD: National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-181.

Newhouse, W., Keith, S., Scribner, B., & Witte, G. 2017b. ***National Initiative for***

Cybersecurity Education (NICE) Cybersecurity Workforce Framework. no. NIST SP 800-181, Gaithersburg, MD: National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-181.

Niederman, F., & March, S. T. 2012. Design science and the accumulation of knowledge in the information systems discipline. ***ACM Transactions on Management Information Systems***, 3(1): 1:1-1:15.

NIST. 2021. ***NICE Cybersecurity Workforce Framework***. https://niccs.us-cert.gov/workforce-development/cyber-security-workforce-framework.

Obrst, L. 2010. Ontological architectures. ***Theory and applications of ontology: Computer applications***: 27–66. Springer.

Ochuko, R. E. 2013. ***E-banking operational risk assessment. A soft computing approach in the context of the Nigerian banking industry.*** University of Bradford. https://bradscholars.brad.ac.uk/handle/10454/5733.

Oltsik, J. 2019. The cybersecurity skills shortage is getting worse. ***CSO Online***. https://www.csoonline.com/article/3331983/the-cybersecurity-skills-shortage-is-getting-worse.html.

Opdahl, A. L., & Henderson-Sellers, B. 2002. Ontological Evaluation of the UML Using the Bunge–Wand–Weber Model. ***Software and Systems Modeling***, 1(1): 43–67.

Partida, A., & Andina, D. 2010a. Vulnerabilities, Threats and Risks in IT. ***IT Security Management***: 1–21. Springer.

Partida, A., & Andina, D. 2010b. Vulnerabilities, Threats and Risks in IT. ***IT Security Management***: 1–21. Springer.

PCI Council. 2021. PCI Security Standards Council Site. ***PCI***. https://www.pcisecuritystandards.org/document_library?category=pcidss&document=pci_dss.

Petrella, E. 2017, January 10. NICE Cybersecurity Workforce Framework. ***NIST***. https://www.nist.gov/itl/applied-cybersecurity/nice/resources/nice-cybersecurity-workforce-framework.

Pfeffer, J., & Sutton, R. I. 2006, January. Evidence-based management. ***Harvard Bussiness Review***, 62–74.

Pidgeon, N. 1998. Safety culture: Key theoretical issues. ***Work & Stress***, 12(3): 202.

Prescott, R. K. 2012. ***The Encyclopedia of Human Resource Management, Volume 1: Short Entries***, vol. 1. John Wiley & Sons.

PricewaterhouseCoopers. 2020. 24th Annual Global CEO Survey. ***PwC***. https://www.pwc.com/gx/en/ceo-agenda/ceosurvey/2021.html.

Radevski, V., & Trichet, F. 2006. ***Ontology-Based Systems Dedicated to Human Resources Management: An Application in e-Recruitment***, 4278: 1068–1077.

Rajivan, P., & Cooke, N. J. 2018. Information-Pooling Bias in Collaborative Security Incident Correlation Analysis: ***Human Factors***.

https://doi.org/10.1177/0018720818769249.

Richter, A., & Koch, C. 2004. Integration, differentiation and ambiguity in safety cultures. ***Safety Science***, 42(8): 703–722.

Rigby, D., & Bilodeau, B. 2007. A Growing Focus on Preparedness. ***Harvard Business Review***, 85(7/8): 21–22.

Rippel, M., & Teply, P. 2010. Operational Risk–Scenario Analysis. ***World Academy of Science, Engineering and Technology, International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering***, 4(6): 723–730.

Rohrbeck, R., & Schwarz, J. O. 2013. The value contribution of strategic foresight: Insights from an empirical study of large European companies. ***Technological Forecasting and Social Change***, 80(8): 1593–1606.

Ross, R., McEvilley, M., & Oren, J. 2018. ***Systems Security Engineering: Considerations for a Multidisciplinary Approach in the Engineering of Trustworthy Secure Systems***. no. NIST Special Publication (SP) 800-160 Vol. 1, National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-160v1.

Sanchez, R. 2004. Understanding competence-based management. ***Journal of Business Research***, 57(5): 518–532.

Savoie-Zajc, L. 2009. L'entrevue semi-dirigée. ***Recherche Sociale de la problématique à la collecte des données*** (5th ed.). Presses de l'Université du Québec.

Scarfone, K. A., Jansen, W., & Tracy, M. 2008. ***SP 800-123. Guide to general server security***. National Institute of Standards & Technology.

Schatz, D., Bashroush, R., & Wall, J. 2017. Towards a more representative definition of cyber security. ***Journal of Digital Forensics, Security and Law***, 12(2): 53–74.

Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. 2011. Action design research. ***MIS Quarterly***, 37–56.

Seong, J. Y., Kristof-Brown, A. L., Park, W.-W., Hong, D.-S., & Shin, Y. 2015. Person-Group Fit: Diversity Antecedents, Proximal Outcomes, and Performance at the Group Level. ***Journal of Management***, 41(4): 1184–1213.

SERENE-RISC. 2020. ***Cybersecurity course directory—SERENE-RISC***. https://www.serene-risc.ca/en/cybersecurity-course-directory.

Sheikhpour, R., & Modiri, N. 2012. A best practice approach for integration of ITIL and ISO/IEC 27001 services for information security management. ***Indian Journal of Science and Technology***, 5(2): 2170–2176.

Shillair, R., Cotten, S. R., Tsai, H.-Y. S., Alhabash, S., LaRose, R., et al. 2015. Online safety begins with you and me: Convincing Internet users to protect themselves. ***Computers in Human Behavior***, 48: 199–207.

Sokolova, M., & Lapalme, G. 2009. A systematic analysis of performance measures for classification tasks. ***Information Processing & Management***, 45(4): 427–437.

Stadlhofer, B., Salhofer, P., & Durlacher, A. 2013. ***An Overview of Ontology***

Engineering Methodologies in the Context of Public Administration, 36–42. Presented at the SEMAPRO 2013, The Seventh International Conference on Advances in Semantic Processing.

StarDog. 2021, January 1. Stardog 7 Manual. ***Stardog 7 Manual***. https://www.stardog.com/docs.

Subramaniam, T., Nizam, I., & Kamil Eissa, A. M. 2019. The Impact of Core Competencies of IT Professionals on Business Success in Malaysia. ***International Journal of Management, Accounting & Economics***, 6(7): 496–520.

Sure-Vetter, Y., Staab, S., & Studer, R. 2009. Ontology Engineering Methodology. ***Handbook on Ontologies***: 135–152.

Taubenberger, S. 2014. ***Vulnerability Identification Errors in Security Risk Assessments***. The Open University. http://oro.open.ac.uk/39626/.

Teece, D. J. 2007. Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. ***Strategic Management Journal***, 28(13): 1319–1350.

Teece, D. J. 2018. Business models and dynamic capabilities. ***Long Range Planning***, 51(1): 40–49.

Teece, D. J., Pisano, G., & Shuen, A. 1997. Dynamic capabilities and strategic management. ***Strategic Management Journal***, 18(7): 509–533.

Tharwat, A. 2020. Classification assessment methods. ***Applied Computing and Informatics***, 17(1): 168–192.

Thomsen, M., Sørensen, P. B., Fauser, P., & Porragas, G. E. 2006. Risk scenario analysis. ***Epidemiology***, 17(6): S486–S487.

Tilakaratna, P., & Rajapakse, J. 2017. Evaluation of the Ontological Completeness and Clarity of Object-Oriented Conceptual Modelling Grammars. ***J. Database Manag.*** https://doi.org/10.4018/JDM.2017040101.

Twum, F., & Ahenkora, K. 2012. Internet Banking Security Strategy: Securing Customer Trust. ***Journal of Management and Strategy***, 3(4): p78.

Ula, M., Ismail, Z. bt, & Sidek, Z. M. 2011. A Framework for the Governance of Information Security in Banking System. ***Journal of Information Assurance & Cybersecurity***, 12.

USC. 2002. [USC02] 15 USC Ch. 98: PUBLIC COMPANY ACCOUNTING REFORM AND CORPORATE RESPONSIBILITY. ***15 USC Ch. 98***, vol. 15. https://uscode.house.gov/view.xhtml?path=/prelim@title15/chapter98&edition=prelim.

Uschold, M., & Gruninger, M. 1996. Ontologies: Principles, methods and applications. ***The Knowledge Engineering Review***, 11(2): 93–136.

van Aken, J. E. 2004. Management research based on the paradigm of the design sciences: The quest for field-tested and grounded technological rules. ***Journal of Management Studies***, 41(2): 219–246.

van Kessel, P. 2018. ***Is cybersecurity about more than protection?*** https://www.ey.com/en_kw/advisory/global-information-security-survey-2018-2019.

Van Rijsbergen, C. J. 1974. Foundation of evaluation. ***Journal of Documentation***, 30(4): 365–373.

Velasco, D., & Rodriguez, G. 2017. Ontologies for Network Security and Future Challenges. ***ArXiv:1704.02441 [Cs]***. http://arxiv.org/abs/1704.02441.

Verdonck, M., & Gailly, F. 2016. Insights on the Use and Application of Ontology and Conceptual Modeling Languages in Ontology-Driven Conceptual Modeling. In I. Comyn-Wattiau, K. Tanaka, I.-Y. Song, S. Yamamoto, & M. Saeki (Eds.), ***Conceptual Modeling***, 83–97. Cham: Springer International Publishing.

Vyšniauskas, E., Nemuraite, L., & Paradauskas, B. 2012. Preserving Semantics of Owl 2 Ontologies in Relational Databases Using Hybrid Approach. ***Information Technology And Control***, 41: 103–115.

Westley, F., & Mintzberg, H. 1989. Visionary leadership and strategic management. ***Strategic Management Journal***, 10(S1): 17–32.

White House. 2016, February 9. FACT SHEET: Cybersecurity National Action Plan. ***Whitehouse.gov***. https://obamawhitehouse.archives.gov/the-press-office/2016/02/09/fact-sheet-cybersecurity-national-action-plan.

Whittemore, R., Chase, S. K., & Mandle, C. L. 2001. Validity in Qualitative Research. ***QUALITATIVE HEALTH RESEARCH***.

Wilkinson, A., & Kupers, R. 2013. Living in the Futures. ***Harvard Business Review***, 91(5): 118–127.

Yoe, C. 2011. ***Principles of risk analysis: Decision making under uncertainty***. CRC press. https://books.google.ca/books?hl=fr&lr=lang_en|lang_fr&id=Bx_qmjDHD0IC&oi=fnd&pg=PP1&dq=%22risk+scenario%22+%22IT+risk%22+%22risk+evaluation%22&ots=-AfxQZkFlJ&sig=wPvxk380DP-4olqWk3NMWGVI69E.

Zainal, A. 2017. ***ICT industry integrated curricula: Towards an ontology based competency model***. http://qspace.qu.edu.qa/handle/10576/5353.

12 Appendix A: Detailed query results and analysis table

Document	Cybersecurity Analyst			Security Analyst			Red Team Analyst			Blue Team Analyst			Risk Scenarios	
List	Q5	Q9	Q16	Q6	Q10	Q17	Q7	Q11	Q18	Q8	Q12	Q19	SC1	SC2
0	TP	TP	TN	TP	FN	FN	FP	FP	TN	FP	FP	TN	TN	TN
1	TN	TN	TN	TP	FN	FN	FP	FP	TN	FP	FP	TN	TN	TN
2	TP	FN	FN	TP	FN	FN	FP	FP	TN	FP	FP	TN	TN	TN
3	TP	TP	TP	TP	FN	TP	FP	FP	FP	FP	FP	FP	TP	TN
4	TP	TP	TP	TP	FN	TP	FP	FP	TN	TP	TP	FN	TP	TP
5	TP	TP	TP	TP	FN	TP	TP	TP	TP	TP	TP	TP	TP	TP
6	TP	FN	TP	TP	FN	TP	TP	TP	TP	FP	FP	FP	TP	TP
7	TP	FN	FN	TP	FN	FN	FP	FP	TN	TP	TP	FN	TN	FN
8	FP	TN	TN	FP	TN	FP	TN	TN	TN	FP	TN	TN	TN	TN
9	FP	FP	TN	TP	FN	FN	FP	FP	FN	FP	FP	TN	TN	TN
10	FN	FN	FN	TP	FN	TP	FP	FP	FP	FP	FP	FP	TP	TN
11	FP	TN	FP	FP	TN	FP	FP	FP	FP	FP	FP	FP	TN	TN
12	FP	FP	TN	FP	TN	TN	FP	FP	TN	TP	TP	FN	TN	TN

13		FP	FP	FP		FP	TN	FP		FP	FP	FP		FP	FP	FP		TP	FP
14		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
15		FP	FP	FP		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
16		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
17		TN	TN	TN		FP	TN	FP		TN	TN	TN		TN	TN	TN		TN	TN
18		TN	TN	TN		FP	TN	FP		TN	TN	TN		TN	TN	TN		TN	TN
19		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
20		TN	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN
21		TP	TN	TN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
22		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TP
23		TN	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN
24		FP	FP	TN		FP	TN	TN		FP	FP	TN		FP	FP	TN		TN	TN
25		FP	FP	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TP	TN
26		TP	FN	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
27		TN	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
28		TN	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TP	FP

29		TP	FN	FN		TP	FN	TP		TP	TP	TP		TP	TP	TP		TP	TP
30		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
31		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TN
32		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
33		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
34		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
35		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TN
36		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
37		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
38		FP	TN	TN		FP	TN	FP		TN	FN	FP		FP	TN	FP		TP	TN
39		FP	FN	FN		FP	TN	FP		FP	FP	FN		FP	FP	FN		TP	TP
40		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	FP
41		TP	TP	FN		TP	FN	FN		FP	FP	TN		FP	FP	FP		TN	TN
42		TN	TN	TN		FP	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN
43		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TP	TN
44		FN	FN	FN		TP	FN	TP		TP	TP	TP		TP	TP	TP		TN	TN

45		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
46		TP	TN	TN		TP	FN	TP		TN	TN	TN		TP	TN	TN		TP	TN
47		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
48		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
49		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	FP
50		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TN	TN
51		TN	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN
52		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
53		TN	TN	TN		FP	TN	TN		FP	FP	FN		FP	FP	TN		TN	TN
54		TP	TP	TP		TP	FN	TP		FP	FP	TN		FP	FP	TN		TP	FP
55		FP	FP	FP		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
56		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
57		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
58		TN	TN	TN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
59		TN	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TP	TN
60		TP	TP	FN		TP	FN	FN		TN	TN	TN		FP	FP	TN		TN	TN

61		TP	TP	TP		TP	FN	TP		TN	TN	FP		FP	FP	FP		TN	TN
62		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
63		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
64		TP	TP	FN		TP	FN	FN		FP	FP	FN		TP	TP	FN		TN	FN
65		TP	TP	TP		TP	FN	TP		FP	FP	TP		TP	TP	TP		TP	TN
66		TP	TP	TP		TP	FN	TP		FP	FP	TP		TP	TP	TP		TP	TP
67		TP	TP	TP		TP	FN	TP		FP	FP	FN		FP	FP	TN		TP	TN
68		TP	TP	TP		TP	FN	TP		FP	FP	TN		FP	FP	TN		TP	TP
69		FN	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
70		TP	TP	FN		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
71		TP	FN	FN		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
72		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
73		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
74		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	FN
75		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
76		TP	TP	TP		TP	FN	TP		FP	FP	TN		TP	TP	FN		TN	TN

77		TN	TN	TN		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TN
78		TN	TN	TN		FN	FN	FN		TN	TN	TN		TN	TN	TN		TN	TN
79		TN	TN	TN		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TN
80		TN	TN	TN		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TN
81		TN	TN	TN		TP	FN	TP		FP	FP	FP		TP	TP	TP		TN	TP
82		TP	TP	FN		TP	FN	TP		FP	FP	TN		FP	FP	TN		TP	TN
83		TP	TP	TP		TP	FN	FN		FP	FP	FP		FP	FP	FP		TN	TN
84		TP	TP	FN		FP	TN	TN		FP	FP	TN		FP	FP	TN		TN	TN
85		TP	TP	TP		TP	FN	TP		FP	FP	TN		TP	TP	FN		TN	TP
86		TP	TP	TP		TP	FN	TP		FP	FP	TN		TP	TP	FN		TP	TP
87		TP	TN	TN		TP	FN	TP		TN	TN	TN		FP	TN	TN		TN	TN
88		TP	TN	TN		TP	FN	TP		FP	FP	TN		FP	FP	FP		TP	TN
89		TP	TP	TP		TP	FN	TP		FP	FP	TN		TP	TP	FN		TN	FN
90		FP	TN	TN		FP	TN	FP		TN	TN	TN		TN	TN	TN		TN	TN
91		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TP	TP
92		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN

93		FP	FP	FP		FP	TN	FP		FP	FP	FP		FP	FP	FP		TP	FP
94		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
95		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
96		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
97		TP	TP	TP		TP	FN	TP		TN	TN	TN		FP	TN	TN		TN	TN
98		TP	FN	FN		TP	FN	TP		FP	FP	TN		FP	FP	FN		TP	TP
99		FN	FN	FN		TP	FN	TN		TN	TN	TN		FP	TN	TN		TN	TN
100		FN	FN	FN		TP	FN	FN		TN	TN	TN		TN	TN	TN		TN	TN
101		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
102		FN	FN	FN		TP	FN	TP		FP	FP	FP		TP	TP	TP		TN	TP
103		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
104		FP	FP	FP		FP	TN	FP		FP	FP	FP		FP	FP	FP		TP	FP
105		TN	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
106		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TP	TN
107		TP	TP	TP		TP	FN	TP		TP	TP	TP		TP	TP	TP		TP	TN
108		TN	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN

109	TP	TP	TP	TP	FN	TP	FP	FP	FP	TP	TP	TP	TP	FP
110	TP	TP	TP	TP	FN	TP	FP	FP	FP	TP	TP	TP	TP	TN
111	FN	FN	FN	TP	FN	TP	TN	TN	TN	TN	TN	TN	TP	TN
112	FP	FP	FP	FP	FN	FP	TN	TN	TN	FN	FN	FN	TN	TN
113	TP	TP	TP	TP	FN	TP	FP	FP	FP	TP	TP	TP	TP	FN
114	TP	TP	TP	TP	FN	TP	FP	FP	FP	TP	TP	TP	TP	TN
115	TP	TP	TP	TP	FN	TP	FP	FP	FP	TP	TP	TP	TP	TP
116	FN	FN	FN	TP	FN	TP	FP	FP	FP	FP	FP	FP	TP	TN
117	TP	FN	FN	TP	FN	TP	FP	FP	FP	FP	FP	FP	TP	TN
118	TN	TN	TN	FP	TN	FP	FP	FP	FP	TP	TP	TP	TP	FN
119	TP	TP	TP	TP	FN	TP	FP	FP	FP	TP	TP	TP	TN	TP
120	TP	TP	TP	TP	FN	TP	FP	FP	FP	FP	FP	FP	TN	TP
121	TP	TP	FN	TP	FN	FN	FP	FP	FP	FP	FP	FP	TN	TN
122	TP	FN	FN	TP	FN	TP	FP	FP	FP	TP	TP	TP	TN	TN
123	FP	TN	TN	FP	TN	FP	FP	FP	FP	TP	TP	TP	TP	TP
124	TP	FN	FN	TP	FN	FN	FP	FP	FP	TP	TP	TP	TN	TN

125		FP	TN	TN		FP	TN	FP		FP	FP	FP		TP	TP	TP		TP	TP
126		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
127		FP	FP	TN		FP	TN	TN		FP	FP	FP		TP	TP	FN		TN	TN
128		FN	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
129		TN	TN	TN		FP	TN	FP		FP	FP	FP		TP	TP	TP		TP	TN
130		FP	FP	FP		FP	TN	FP		FP	FP	FP		TP	TP	TP		TP	TP
131		FP	TN	TN		FP	TN	FP		TN	TN	TN		TP	FN	FN		TN	TN
132		TP	TP	FN		TP	FN	TP		FP	FP	FP		TP	TP	TP		TN	TN
133		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TN
134		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
135		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TP
136		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	FP
137		FP	TN	TN		TP	FN	FN		FP	FP	FP		FP	FP	FP		TP	TP
138		FP	TN	TN		FP	TN	TN		FP	FP	FP		FP	FP	FP		TP	TP
139		FP	FP	TN		FP	TN	TN		TP	TP	FN		FP	FP	TN		TN	TN
140		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TN	FP

141	TP	FN	FN	TP	FN	TP	TN	TN	TN	141	TN	TN	TN	TN
142	TN	TN	TN	TP	FN	FN	FP	FP	TN	TP	TP	TP	TN	TN
143	TP	TP	TP	TP	FN	TP	FP	FP	TN	FP	FP	TN	TN	TN
144	TP	TP	TP	TP	FN	TP	FP	FP	FP	FP	FP	FP	TP	FP
145	FP	FP	FP	FP	FN	FP	FP	FP	FP	FP	FP	FP	TP	TN
146	TN	TN	TN	FP	FN	FP	FP	FP	FP	FP	FP	FP	TN	TN
147	FN	FN	FN	TP	FN	TP	TN	TN	TN	TP	TP	TP	TP	TN
148	TP	TP	TP	TP	FN	TP	FP	FP	FP	FP	FP	FP	TP	TP
149	TP	TP	TP	TP	FN	TP	FP	FP	FP	FP	FP	FP	TP	TN
150	TP	TP	TP	TP	FN	TP	FP	FP	FP	FP	FP	FP	TP	TN
151	TP	TP	FN	TP	FN	FN	FP	FP	TN	TP	TP	FN	TN	TN
152	TP	FN	FN	TP	FN	TP	FP	FP	FP	FP	FP	FP	TP	TP
153	TP	TP	TP	TP	FN	TP	FP	FP	FP	FP	FP	FP	TP	TP
154	TP	TP	FN	TP	FN	FN	FP	FP	FP	FP	FP	TN	TN	TN
155	FN	FN	FN	TP	FN	TP	TN	TN	TN	FN	FN	FN	TP	FN
156	TP	TP	FN	TP	TP	FN	FP	FP	TN	FP	FP	TN	TN	TN

157		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
158		TP	TP	TP		TP	FN	TP		FP	FP	TN		FP	FP	TN		TP	FP
159		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TP
160		TP	TP	TP		TP	FN	TP		FP	FP	TN		TP	TP	FN		TN	TN
161		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TN	TP
162		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TN	TN
163		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TP
164		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TN
165		TP	TP	FN		TP	FN	FN		FP	FP	TN		TP	TP	FN		TN	TN
166		TP	FN	FN		TP	FN	TP		TN	TN	TN		FP	TN	TN		TP	TN
167		FN	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
168		FN	FN	FN		FN	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
169		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
170		FN	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
171		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
172		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP

173		FN	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TP
174		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
175		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
176		TP	TP	FN		TP	TN	TN		FP	FP	FP		FP	FP	TN		TN	TN
177		TN	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
178		TN	TN	TN		FP	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN
179		TP	TN	TN		TP	TN	TN		FP	FP	TN		TP	TP	FN		TN	TN
180		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TP
181		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TN	TN
182		FN	FN	FN		TP	TN	TP		FP	FP	FP		FP	FP	FP		TP	TP
183		TP	TP	TP		TP	TP	TP		FP	FP	FP		TP	TP	TP		TP	TP
184		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TP
185		TN	TN	TN		FP	TN	FP		FP	FP	FP		TP	TP	TP		TN	FN
186		TN	TN	TN		FP	TN	FP		FN	FN	FN		TN	TN	TN		TN	TN
187		TP	TP	FN		TP	FN	FN		FP	FP	FP		FP	FP	TN		TN	TN
188		FP	FP	FP		FP	TN	FP		FP	FP	FP		FP	FP	FP		TP	TN

189		FN	FN	FN		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
190		TP	FN	FN		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
191		FN	FN	FN		TP	FN	TP		TN	TN	TN		TN	TN	TN		TP	TN
192		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TN	TN
193		FN	FN	FN		TP	FN	TP		TN	TN	TN		FN	FN	FN		TP	TN
194		FN	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
195		TP	TP	TP		TP	FN	TP		TN	TN	TN		FN	FN	FN		TN	FN
196		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
197		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	FP
198		TP	TP	TP		TP	FN	TP		FP	FP	TN		TP	TP	FN		TN	TN
199		FN	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
200		FN	FN	FN		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TP
201		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
202		TP	FN	FN		TP	FN	TP		FP	FP	FP		TP	TP	TP		FN	FN
203		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
204		FP	FP	FP		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	FP

205		FP	FP	FP		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TP
206		FP	FP	TN		FP	TN	TN		TP	TP	TP		FP	FP	TN		TN	TN
207		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
208		FP	FP	FP		FP	TN	FP		FP	FP	FP		FP	FP	FP		TP	TP
209		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
210		FP	TN	TN		FP	TN	TN		FP	FP	FP		FP	FP	TN		TN	TN
211		FN	FN	FN		TP	FN	TP		FP	FP	FP		TP	TP	TP		FN	TP
212		TP	TP	TP		TP	FN	TP		TP	TP	TP		FP	FP	FP		TP	TP
213		FN	FN	FN		TP	FN	TP		TN	TN	FP		TN	TN	TN		TP	TN
214		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TP
215		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
216		FN	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
217		FN	FN	FN		TP	FN	TP		TN	TN	FP		TN	TN	TN		TN	TN
218		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
219		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TN	TN
220		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TN

221		TP	TP	TP		TP	TP	TP		FP	FP	FP		FP	FP	FP		TP	TP
222		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
223		TP	TP	TP		FP	TN	FP		FP	FP	FP		TP	TP	TP		FP	TP
224		FP	FP	FP		FP	TN	FP		FP	FP	TN		FP	FP	FP		TN	TN
225		TP	FN	FN		TP	FN	TP		TP	TP	FN		TP	TP	TP		TN	TN
226		TP	FN	FN		TP	FN	TP		FP	FP	TN		FP	FP	FP		TN	FP
227		TP	TP	TP		TP	FN	TP		FP	FP	TN		FP	FP	TN		TN	TN
228		TP	FN	FN		TP	FN	TP		FP	FP	TN		FP	FP	FP		TP	TN
229		TP	TP	FN		TP	FN	FN		FP	FP	TN		TP	TP	FN		FN	FN
230		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
231		TN	TN	TN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
232		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	FN
233		TP	FN	FN		TP	FN	TP		FP	FP	FP		TP	TP	FN		TN	TN
234		TP	FN	FN		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TN
235		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
236		FP	FP	FP		FP	TN	FP		TN	TN	TN		FN	FN	FN		FP	TN

237	FP	FP	FP	FP	TN	FP	FP	FP	FP	TP	TP	FN	FP	FN
238	TP	FN	FN	TP	FN	FN	FP	FP	FP	FP	FP	TN	TN	TN
239	TP	FN	FN	TP	FN	FN	FP	FP	FP	FP	FP	FP	TP	TP
240	TP	FN	FN	TP	FN	TP	FP	FP	FP	TP	TP	TP	TN	TP
241	TP	FN	FN	TP	FN	TP	FP	FP	FP	FP	FP	FP	TN	TP
242	TP	TP	TP	TP	FN	TP	FP	FP	FP	TP	TP	TP	TP	TP
243	TP	TP	TP	TP	FN	TP	TP	TP	FN	FP	FP	TN	FN	TN
244	TP	FN	FN	TP	FN	TP	FP	FP	FP	TP	TP	TP	TN	TN
245	TP	FN	FN	TP	FN	TP	FP	FP	TN	FP	FP	TN	TN	TN
246	TP	FN	FN	TP	FN	TP	FP	FP	FP	FP	FP	FP	TP	TN
247	TN	TN	TN	TP	FN	TP	FP	FP	FP	FP	FP	FP	TN	TN
248	TN	TN	TN	TP	FN	FN	FP	FP	FP	FP	FP	FP	TP	TN
249	TP	FN	FN	TP	FN	FN	FP	FP	FP	FP	FP	TN	TN	TN
250	TN	TN	TN	TP	FN	TP	TN	TN	TN	TN	TN	TN	TP	TN
251	TP	FN	FN	TP	FN	TP	FP	FP	FP	FP	FP	FP	TP	TN
252	TP	FN	FN	TP	FN	TP	FP	FP	FP	FP	FP	FP	TP	TN

253		FN	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	FP
254		TP	TP	TP		TP	FN	TP		FP	FP	TN		FP	FP	TN		TP	TN
255		TP	TP	FN		TP	FN	TP		FP	FP	TN		FP	FP	TN		TN	TN
256		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
257		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
258		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
259		TN	TN	TN		TP	FN	FN		FP	FP	FP		FP	FP	FP		TN	TN
260		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	TN		TN	TN
261		TP	FN	FN		TP	FN	TP		TP	TP	TP		TP	TP	TP		TN	TN
262		TN	TN	TN		FN	FN	TP		TN	TN	TN		FN	FN	FN		FN	TN
263		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		FN	TN
264		TP	TP	FP		TP	FN	TP		FP	FP	TN		TP	TP	FN		FN	TN
265		TN	TN	TN		FP	TN	FP		TN	TN	TN		TN	TN	TN		FP	TN
266		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
267		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TN	TN
268		TP	TP	TP		TP	FN	FN		FP	FP	FP		FP	FP	FP		TN	TN

269		TP	TP	TP		TP	FN	FN		FP	FP	FP		FP	FP	FP		TN	TN
270		TP	TP	TP		TP	FN	TP		FP	FP	FN		TP	TP	FN		TN	FN
271		FP	TN	TN		FP	TP	FP		TP	TP	FN		FP	FP	TN		TN	TN
272		FP	TN	TN		FP	TN	FP		TN	TN	TN		TN	TN	TN		FP	TN
273		FN	FN	FN		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TP
274		TN	TN	TN		FP	TN	FP		FP	FP	FP		TP	TP	TP		FP	TP
275		TN	TN	TN		FP	TN	FP		FP	FP	TN		FP	FP	TN		TN	TN
276		TP	TP	TP		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
277		TN	TN	TN		FP	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN
278		TP	TP	FN		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
279		FP	FP	TN		FP	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN
280		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
281		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
282		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		FP	TN
283		FP	TN	TN		TN	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
284		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN

285		TP	TP	TP		TP	TP	TP		FP	FP	FP		TP	TP	TP		TP	TP
286		FP	FP	FP		FP	TN	FP		FP	FP	TN		FP	FP	TN		TN	TN
287		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	FN
288		TN	TN	TN		TP	FN	TP		TN	TN	TN		FN	FN	FN		TP	TN
289		TP	FN	FN		TP	FN	TP		TP	TP	TP		TP	TP	TP		TN	TN
290		FP	FP	FP		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
291		FP	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TP	TN
292		TP	FN	FN		FP	TN	FP		FP	FP	TN		FP	FP	TN		TN	TN
293		TN	TN	TN		FP	TN	FP		FP	FP	FP		TP	TP	TP		TN	FN
294		TP	TP	FN		TP	FN	TP		FP	FP	TN		FP	FP	TN		TN	TN
295		TP	TP	TP		TP	FN	TP		FP	FP	TN		FP	FP	TN		TN	TN
296		TN	TN	TN		TP	FN	TP		TN	TN	TN		FN	FN	FN		TN	TN
297		TN	TN	TN		TP	FN	TP		TN	TN	TN		TN	TN	TN		TP	TN
298		TP	FN	FN		FP	TN	TN		FP	FP	FP		FP	FP	FP		FP	TN
299		TN	TN	TN		FP	TN	TN		FP	FP	FP		FP	FP	FP		FP	TP
300		TP	TP	TP		TP	FN	TP		TN	TN	TN		TN	TN	TN		TP	TN

301		FN	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
302		FN	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		FP	TP
303		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
304		TP	FN	FN		TP	FN	TP		FP	FP	FP		TP	TP	TP		FN	TP
305		FP	FP	FP		FP	TN	FP		FP	FP	FP		FP	FP	FP		FP	FP
306		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	FN
307		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TP
308		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TP
309		TP	TP	FN		TP	FN	FN		FP	FP	FP		TP	TP	FN		TN	TN
310		TN	TN	TN		FN	FN	FN		FP	FP	FP		FP	FP	FP		TN	TN
311		TN	TN	TN		TP	FN	TP		TN	TN	TN		TN	TN	TN		TN	TN
312		TP	FN	FN		FP	TN	FP		FP	FP	FP		FP	FP	FP		TN	TN
313		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
314		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TP
315		TP	FN	FN		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
316		TN	TN	TN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN

317		TN	TN	TN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
318		TP	FN	FN		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
319		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
320		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
321		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	FP
322		TP	TP	TP		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
323		TN	TN	TN		FP	TN	TN		FP	FP	TN		FP	FP	TN		TN	TN
324		TN	TN	TN		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TN
325		TP	FN	FN		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
326		TP	TP	FN		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
327		TP	TP	FN		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
328		TN	TN	TN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
329		TN	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN
330		TN	TN	TN		TP	FN	TP		TN	TN	TN		TN	TN	TN		TP	TN
331		TN	TN	TN		FN	FN	FN		TN	TN	TN		TN	TN	TN		TN	TN
332		TP	TP	FN		TP	FN	TP		FP	FP	TN		FP	FP	TN		TN	TN

333		TN	TN	TN		FN	FN	FN		TN	TN	TN		TN	TN	TN		TN	TN
334		TN	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN
335		TN	TN	TN		TP	FN	TP		TN	TN	TN		TN	TN	TN		TN	TN
336		TN	TN	TN		FN	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN
337		TN	TN	TN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
338		TN	TN	TN		FN	FN	FN		TN	TN	TN		TN	TN	TN		TN	TN
339		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
340		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
341		TN	TN	TN		TP	FN	TP		TN	TN	TN		TN	TN	TN		TP	TN
342		TN	TN	TN		FP	TN	FP		FP	FP	FP		TP	TP	TP		FP	TP
343		TN	TN	TN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
344		TN	TN	TN		FP	TN	TN		TN	TN	TN		TN	TN	TN		TN	TN
345		FP	FP	FP		FP	TN	FP		FP	FP	TN		TP	TP	FN		FP	TP
346		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
347		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TP
348		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN

349		TP	FN	FN		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
350		TP	FN	FN		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
351		TP	TP	TP		TP	FN	TP		FP	FP	FP		TP	TP	TP		TP	TP
352		TN	TN	TN		TP	FN	TP		TN	TN	TN		TN	TN	TN		TN	TN
353		TN	TN	TN		TP	FN	FN		TN	TN	TN		TN	TN	TN		TN	TN
354		FP	FP	FP		FP	TN	FP		FP	FP	FP		TP	TP	TP		FN	TP
355		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TP
356		TP	TP	FN		TP	FN	FN		FP	FP	TN		FP	FP	TN		TN	TN
357		TP	TP	TP		TP	FN	TP		TN	TN	TN		FP	TN	TN		TP	TN
358		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
359		TP	TP	TP		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
360		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TN	TN
361		TN	TN	TN		FP	TN	FP		FP	FP	FP		FP	FP	FP		FP	FP
362		TP	FN	FN		TP	FN	TP		FP	FP	FP		FP	FP	FP		TP	TN
		Cybersecurity Analyst				**Security Analyst**				**Red Team Analyst**				**Blue Team Analyst**				**Risk Scenarios**	
		Q5	**Q9**	**Q16**		**Q6**	**Q10**	**Q17**		**Q7**	**Q11**	**Q18**		**Q8**	**Q12**	**Q19**		**SC1**	**SC2**

# docs		363	363	363		363	363	363		363	363	363		363	363	363		363	363
TP		211	146	123		262	5	217		13	13	11		100	98	77		165	86
TN		65	100	108		7	93	27		56	55	111		36	45	96		175	241
FP		57	28	22		86	0	70		293	293	230		218	211	158		14	19
FN		30	89	110		8	249	49		1	2	11		8	9	32		9	17
SME		270	270	270		270	270	270		16	16	16		109	109	109			
Precision		0.79	0.84	0.85		0.75	1.00	0.76		0.04	0.04	0.05		0.31	0.32	0.33		0.92	0.82
Recall		0.8755	0.6213	0.5279		0.9704	0.0197	0.8158		0.9286	0.8667	0.5		0.9259	0.9159	0.7064		0.9483	0.835
F1-score		0.83	0.71	0.65		0.85	0.04	0.78		0.08	0.08	0.08		0.47	0.47	0.45		0.93	0.83

13 Appendix B: Evaluation of tools

13.1 Solution 1: Stardog

The Stardog product (https://www.stardog.com/) is described as an enterprise knowledge graph platform. Knowledge graphs have become a popular IT tool in recent years to manage unstructured data, such as RDF triple stores and heterogeneous data. Most of the solutions that were investigated describe themselves as being in this category. Being the only commercial product investigated, the researchers were initially reticent to select this solution even considering it came highly recommended. However, following the initial analysis, it was selected.

Pros

- A good and complete solution that could do the work using a module called Bites that has the text analytics and NLP plug-ins that could be used
- It is easy to load the ontology and the supporting material as it has native support for OWL and CSV data formats
- It is possible to load job descriptions and other resources as PDF to analyze documents based on the ontology
- It is available to use at the university for free; there is also a 30-day trial license available which could be sufficient for the tests
- SPARQL support
- Client-server mode with Stardog server used as the backend database and Stardog Studio installed on the local workstation
- It is used by many Canadian financial institutions, in the fraud prevention sector, an activity often closely tied to cybersecurity in this sector.

Cons

- It is an extremely expensive solution to consider after the study for a permanent solution for an organization (200 000$/year for a license)

- Possible lack of support for French characters when imported into Stardog
- Requires support from a local resource on campus to access some of the functions, and to load large files locally, this could be an impediment during the current health crisis
- Predictive analysis functions would require pre-processing and NLP to implement but would not be required as text analysis could be sufficient

13.2 Solution 2: GraphDB

GraphDB is a graph database solution proposed by Ontotext (https://www.ontotext.com/). Initial research indications are that it is a popular product in this category. Search results often mention this tool in similar applications to what was being considered.

Pros

- Open-source community edition for development can be installed and used at no charge
- It seems easy to use and would be the simplest of the solutions analyzed
- It can run as a standalone on my home computer
- It runs in a browser (http://localhost:7200/)
- There is a large availability of support and assistance for problems in development via forums and multiple websites
- Allows importing data with OnToRefine (Google)
- SPARQL support

Cons

- Needs to figure out how to load PDF and run discovery, but it should be doable
- No support at the university
- No native results scoring to speed up the test

13.3 Solution 3: Neo4j

Neo4j is also a graph database product that is often mentioned in the business intelligence community forum (https://neo4j.com/).

Pros

- Open-source community edition for development
- Easy to use
- SPARQL support
- Standalone on my home computer
- Large community of users in data analytics
- Significant availability of support and assistance for problems in development for this solution

Cons

- One of the researchers involved in the study has used this solution and claims it is sub-optimal for applications such as ours
- There are no native results scoring to speed up the test

13.4 Solution 4: TerminusDB

TerminusDB is an open-source knowledge graph database that was identified and recommended by the community (https://terminusdb.com/).

Pros

- Open-source community edition for development
- Standalone on my home computer and runs as an app
- Has tools for facilitating importing CSV into an RDF – triples format
- Easy to work as a group should it be determined to develop future studies based on the results of this study
- Uses GIT for version control and group work

Cons

- Lack of documentation and tutorials to understand how to use
- Lack of availability of support and assistance for problems in development
- Requires learning WOQL rather than using SPARQL
- No native results scoring to speed up the test

13.5 Solution 5: OWLready2

Owlready2 is a module for ontology-oriented programming in Python 3, including an optimized RDF quadstore (https://pypi.org/project/Owlready2/)

Pros

- Open-source community edition for development
- Standalone on my home computer

Cons

- More difficult to use in initial tests
- Requires Python programming to use
- Loads OWL ontology as a Python object
- Lack of availability of support and assistance for problems in development
- No native results scoring to speed up the test

14 Appendix C: Stardog documentation Augmenting Search

Jess Balint

Jul 17, 2018

https://www.stardog.com/blog/augmenting-search/

NOTE to the reader of this study: The classification results without the full-text component are presented in Appendix L of this dissertation. It demonstrates that the document classification schema does work. What is explained in this appendix, from the Stardog online documentation, is how full-text searches are combined with the graphical database capabilities of Stardog to generate usable results from unstructured data. The F1-scores and the other statistical tests that have been performed are what allows the determination that the results are statistically significant. In addition, it must be understood that this is being done using unstructured data, which is being classified against the OWL ontology and the RDF information exported from the ontology into the Stardog database and queried using SPARQL. This is very different from using structured data, such as with a relational database, which could be queried using SQL. Implementing this augmented search capability in this study was done by following the guidelines provided in Appendix C.

Give your Knowledge Graph search results a makeover.

You've mapped and loaded a few sets of data into Stardog. Now what? Depending on your use case, you may be building reports based on SPARQL queries or a search-oriented front-end to a unified view of some **unstructured data**. Stardog provides a capable **full-text index** (FTS) to support searching as well as several other features which can significantly add value to search results. This post explores some of these feature combinations to inspire some ideas for your own applications.

14.1 Document Indexing with BITES

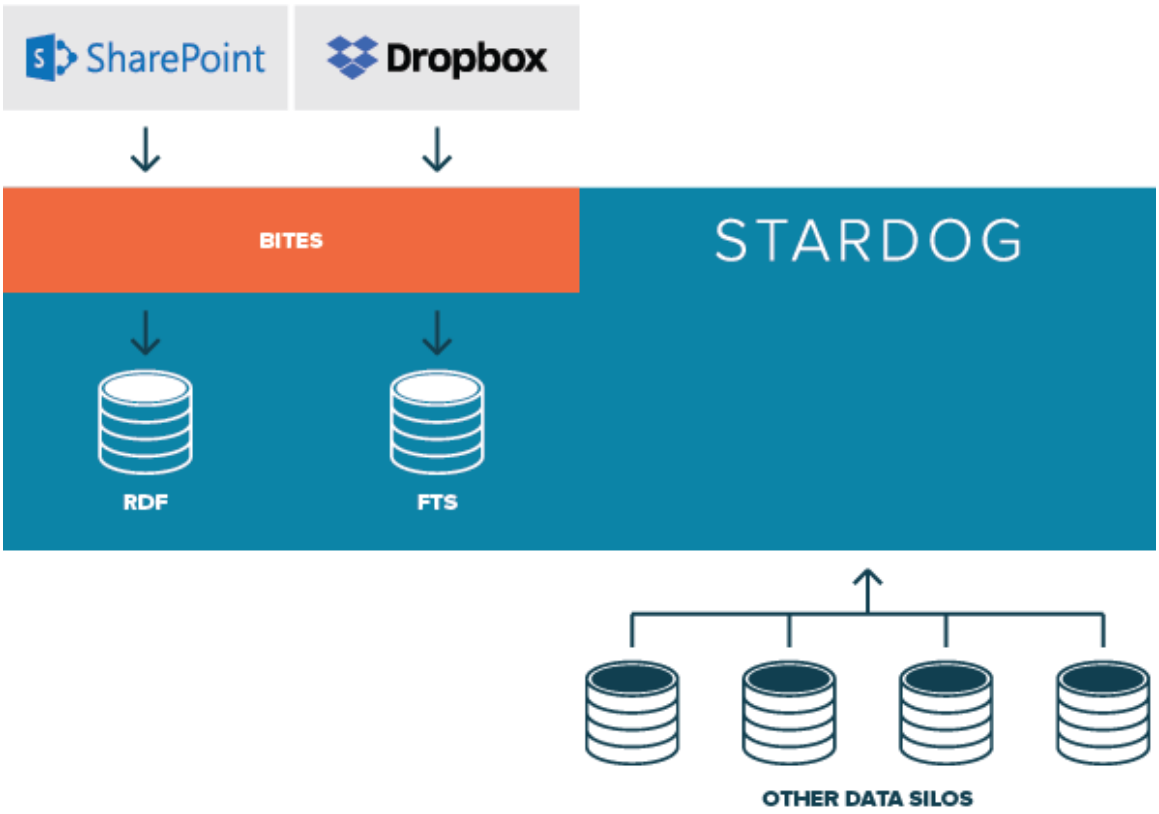
Let us sketch out an example scenario based around a corpus of documents loaded into Stardog's BITES system. BITES provides document storage and indexing as well as some general NLP services. It is not intended to replace any other document management systems such as SharePoint, although it is certainly capable of functioning as the backend of such an application. BITES shines when employed as a document search and processing system used to connect document contents to the rest of your Knowledge Graph.

BITES can index and process documents from your current document storage solution, including SharePoint, Dropbox, Confluence, etc. It is completely general and includes pluggable extension points to configure ingest of any type of file. Additionally, it allows customizable extraction processing and ships with several NLP modules including **entity extraction**.

So let us assume that you have loaded some documents into BITES, potentially from several different parts of your organization. You are now equipped with a searchable view of these documents as well as structured data extracted from the corpus.

The other ingredient is an existing Knowledge Graph, whether materialized into Stardog, or federated as a set of **virtual graphs**—or some combination of these access patterns. Remember: a key value proposition of a Knowledge Graph is *data location does not matter*. Data is invariably linked; hence, creating a unified view over disparate sources is the challenge that Stardog addresses.

Here is what we are working with in terms of data:



14.2 Searching the Document Store

Stardog's built-in **full-text index** provides search capabilities over the graph and the BITES document set. SPARQL queries can use the <tag:stardog:api:property:textMatch> predicate to perform these search queries.

If we extract entities with BITES, we can augment search results with other entities found in documents matching the search. This is where Knowledge Graph unification shines. What if we searched for "George Clooney" and found a review of Ocean's Eleven mentioning other actors in the film? These can be shown alongside the search results, correlated with each document.

A similar approach can be used to add relevant product results to a recipe search. A dictionary-based linker provides recognition of entities in the graph. Product details such as price and availability can be retrieved from external sources. Another possibility is extracting publisher and publication dates from documents. Combined with a source of

publisher locations, we can improve search relevance by prioritizing recent and nearby results. A user in New York searching for “events” likely wouldn’t have much interest in results from a local Mexican newspaper.

We could even pass the search query through the **entity extraction service**. This would provide us with the entities used in the query allowing us to combine the text search result with a query over entity mentions in the BITES index. A search for “Will Smith” might also match documents containing the words “Will” and “Smith” individually. If we discover that “Will Smith” is a named entity, we can filter out results that do not explicitly mention “Will Smith.”

14.3 Extending Search Results with Entity Extraction

Using the built-in entity linker, we extract a set of RDF triples from each document. These triples represent “mentions” in the document. A mention is a reference to a known entity in the graph. The entity linking process is completely independent of use case and searches the graph for known entities. A movie review mentioning George Clooney and Bernie Mac might add the follow triple to the BITES document named graph:

```
review:Oceans11Review.pdf {
        entity:0d25b4ed rdfs:label "George Clooney" ;
                dc:references name:nm0000123 .
        entity:9811ac8c rdfs:label "Bernie Mac" ;
                dc:references name:nm0005170 .
}
```

The IRIs name:nm0000123, name:nm0005170 here identify George Clooney and Bernie Mac, respectively, as nodes in the graph. Using the dc:references predicate, we can query

the graph for documents referring to named entities. Combining this with a search query, we can retrieve a list of named entities for each document in the search result:

```
select ?doc ?mention ?type ?label where {
 # Full-text query
 ?doc <tag:stardog:api:property:textMatch> "George Clooney"
 # Mentions in matched docs
 graph ?doc {
  ?doc dc:references ?mention
 }
 # Class of mentioned entities
 ?mention a ?type ; rdfs:label ?label
}
```

Executing this query would return a result including matching documents, their mentions (IRIs), and classes and labels of the mentions. It might look like so:

```
+-------------------------|---------------|----------|---------------+
| doc                     | mention       | type     | label         |
+-------------------------|---------------|----------|---------------+
| review:Oceans11Review.pdf | name:nm0000123 | :Director | George Clooney |
| review:Oceans11Review.pdf | name:nm0005170 | :Actor    | Bernie Mac     |
```

```
| review:Oceans11Review.pdf | name:nm0005170 | :Comedian | Bernie Mac     |
+-------------------------|---------------|----------|---------------+
```

In addition to the matched documents, we can use mentions, including their type and label, to augment individual search results. Search results become significantly more useful when linked with relevant data. This type of linking is trivial when data is unified in a Knowledge Graph. We can adjust the SPARQL query in many ways to make use of the connected nature of the graph.

14.4 Extending Search Results with External Data Sources

As demonstrated, we can combine our text queries with arbitrary SPARQL queries over the unified graph. The recipes example can be expressed in SPARQL like so:

```
select ?recipe ?product ?productName ?productPrice {
 # Full-text query
 ?recipe <tag:stardog:api:property:textMatch> "potato salad"

 # Product mentions in matched recipes
 graph ?recipe {
  ?recipe dc:references ?product
 }

 # Virtual graph with product details and availability
 graph <virtual://product> {
```

```
    ?product a :Product ;
      :name ?productName ;
      :price ?productPrice ;
      :availableQty ?productQty
    filter(?productQty > 0)
  }
}
```

Entity references to products are stored for each document. This data is combined with an external data source mapped into the graph providing product details and availability.

In the same vein, given a set of documents pertaining to local events, we could combine it with publisher addresses stored in the graph to increase result relevancy. The text search query is over a set of documents for which we extracted the publisher and publication date (using BITES but not the entity extractor). The publisher is then linked to the graph to find its location. A [geospatial query] (https://www.stardog.com/blog/geospatial-a-primer/) allows us to compute the distance between two points and order results by relevance:

```
select ?event ?pubDate ?publisher ?dist ?age {
  # Full-text query
  ?event <tag:stardog:api:property:textMatch> "concert"

  # Document graph with extracted details
```

```
  graph ?event {
    ?event :publishedOn ?pubDate ;
      :publishedBy ?publisher
  }

  # Graph (potentially virtual) with publisher data
  graph <publishers> {
    ?publisher geo:hasGeometry ?publisherLocation
  }

  # Compute the distance between the publisher and the location of the user
  bind(geof:distance(?publisherLocation, :UserLocation, unit:MileUSStatute) as ?dist)
  # Compute the amount of time since the article was published
  bind(now() - ?pubDate as ?age)
}
order by desc(?dist) ?age
```

This query finds concerts using the text search and then orders them first by the shortest distance from the user location and then by the age of the publication date (more recent entries first).

14.5 Use Your Data in Searches

This post contains a glimpse of the ways that search a Knowledge Graph is awesome. It is possible to do significantly more than otherwise possible with a simple full-text index. Feel free to use these ideas directly or experiment using other Stardog features such as machine learning and path queries to improve search results.

15 Appendix D: MITRE ATT&CK validation scenarios

Q	Category	Attack Technique	Mitigation measures used in a query
1	Initial access	Exploit Public-Facing Application https://attack.mitre.org/techniques/T1190/	• Application isolation will limit what other processes and system features the exploited target can access. • Web Application Firewalls may be used to limit exposure of applications to prevent exploit traffic from reaching the application. • Segment externally facing servers and services from the rest of the network with a DMZ or on separate hosting infrastructure. • Use least privilege for service accounts will limit what permissions the exploited process gets on the rest of the system. • Regularly scan externally facing systems for vulnerabilities and establish procedures to rapidly patch systems when critical vulnerabilities are discovered through scanning and through public disclosure. • Regularly scan externally facing systems for vulnerabilities and establish procedures to rapidly patch systems when critical vulnerabilities are discovered through scanning and through public disclosure.
2		External Remote Services https://attack.mitre.org/techniques/T1133/	• Disable or block remotely available services that may be unnecessary. • Limit access to remote services through centrally managed concentrators such as VPNs and other managed remote access systems. • Use strong two-factor or multi-factor authentication for remote service accounts to mitigate an adversary's ability to leverage stolen credentials, but be aware of Two-Factor Authentication Interception techniques for some two-factor authentication implementations.

			• Deny direct remote access to internal systems through the use of network proxies, gateways, and firewalls.
3		Trusted Relationship https://attack.mitre.org/techniques/T1199/	• Network segmentation can be used to isolate infrastructure components that do not require broad network access. • Properly manage accounts and permissions used by parties in trusted relationships to minimize potential abuse by the party and if the party is compromised by an adversary.
4		Valid Accounts https://attack.mitre.org/techniques/T1078/	• Ensure that applications do not store sensitive data or credentials insecurely. • Applications and appliances that utilize default username and password should be changed immediately after the installation, and before deployment to a production environment. When possible, applications that use SSH keys should be updated periodically and properly secured. • Audit domain and local accounts as well as their permission levels routinely to look for situations that could allow an adversary to gain wide access by obtaining credentials of a privileged account. These audits should also include if default accounts have been enabled, or if new local accounts are created that have not to be authorized. Follow best practices for design and administration of an enterprise network to limit privileged account use across administrative tiers.
5	Execution	Command and Scripting Interpreter https://attack.mitre.org/techniques/T1059/	• Anti-virus can be used to automatically quarantine suspicious files. • Where possible, only permit execution of signed scripts. • Disable or remove any unnecessary or unused shells or interpreters. • Use application control where appropriate. • When PowerShell is necessary, restrict PowerShell execution policy to administrators. Be aware that there are methods of bypassing the PowerShell execution policy, depending on environment configuration.

			• Script blocking extensions can help prevent the execution of scripts and HTA files that may commonly be used during the exploitation process. For malicious code served up through ads, adblockers can help prevent that code from executing in the first place.
6		Exploitation for Client Execution https://attack.mitre.org/techniques/T1203/	• Browser sandboxes can be used to mitigate some of the impacts of exploitation, but sandbox escapes may still exist. • Other types of virtualizations and application microsegmentations may also mitigate the impact of client-side exploitation. Risks of additional exploits and weaknesses in those systems may still exist. • Security applications that look for behavior used during exploitation such as Windows Defender Exploit Guard and the Enhanced Mitigation Experience Toolkit can be used to mitigate some exploitation behavior. Control flow integrity checking is another way to potentially identify and stop a software exploit from occurring. Many of these protections depend on the architecture and target application binary for compatibility.
7		Software Deployment Tools https://attack.mitre.org/techniques/T1072/	• Ensure proper system and access isolation for critical network systems through use of group policy. • Ensure proper system and access isolation for critical network systems through use of multi-factor authentication. • Ensure proper system isolation for critical network systems through use of a firewall. • Verify that account credentials that may be used to access deployment systems are unique and not used throughout the enterprise network. • Grant access to application deployment systems only to a limited number of authorized administrators. • If the application deployment system can be configured to deploy only signed binaries, then ensure that the trusted signing certificates are not co-located with the application

			deployment system and are instead located on a system that cannot be accessed remotely or to which remote access is tightly controlled. • Patch deployment systems regularly to prevent potential remote access through Exploitation for Privilege Escalation. • Ensure that any accounts used by third-party providers to access these systems are traceable to the third-party and are not used throughout the network or used by other third-party providers in the same environment. Ensure there are regular reviews of accounts provisioned to these systems to verify continued business need, and ensure there is governance to trace de-provisioning of access that is no longer required. Ensure proper system and access isolation for critical network systems through use of account privilege separation. • Have a strict approval policy for the use of deployment systems.
8		System Services https://attack.mitre.org/techniques/T1569/	• Ensure that permissions disallow services that run at a higher permissions level from being created or interacted with by a user with a lower permission level. • Ensure that high permission level service binaries cannot be replaced or modified by users with a lower permission level. • Prevent users from installing their own launch agents or launch daemons.
9	Persistence	Account manipulation https://attack.mitre.org/techniques/T1098/	• Use multi-factor authentication for user and privileged accounts. • Configure access controls and firewalls to limit access to critical systems and domain controllers. Most cloud environments support separate virtual private cloud instances that enable further segmentation of cloud systems. • Protect domain controllers by ensuring proper security configuration for critical servers to limit access by potentially unnecessary protocols and services, such as SMB file sharing.

			• Do not allow domain administrator accounts to be used for day-to-day operations that may expose them to potential adversaries on unprivileged systems.
10		BITS Jobs https://attack.mitre.org/techniques/T1197/	• Modify network and/or host firewall rules, as well as other network controls, to only allow legitimate BITS traffic. • Consider reducing the default BITS job lifetime in Group Policy or by editing the JobInactivityTimeout and MaxDownloadTime Registry values • Consider limiting access to the BITS interface to specific users or groups.
11		Implant Container Image https://attack.mitre.org/techniques/T1525/	• Periodically check the integrity of images and containers used in cloud deployments to ensure they have not been modified to include malicious software. • Several cloud service providers support content trust models that require container images to be signed by trusted sources. • Limit permissions associated with creating and modifying platform images or containers based on the principle of least privilege.
12		Compromise Client Software Binary https://attack.mitre.org/techniques/T1554/	• Ensure all application component binaries are signed by the correct application developers.
13		Server Software Component: SQL Stored Procedures https://attack.mitre.org/techniques/T1505/001/	• Regularly check component software on critical services that adversaries may target for persistence to verify the integrity of the systems and identify if unexpected changes have been made. • Ensure all application component binaries are signed by the correct application developers. • Do not allow administrator accounts that have permissions to add component software on these services to be used for day-

			to-day operations that may expose them to potential adversaries on unprivileged systems.
14	Privilege escalation	Abuse Elevation Control Mechanism: Bypass User Account Control https://attack.mitre.org/techniques/T1548/002/	• Check for common UAC bypass weaknesses on Windows systems to be aware of the risk posture and address issues where appropriate. • Remove users from the local administrator group on systems. • Consider updating Windows to the latest version and patch level to utilize the latest protective measures against UAC bypass. • Although UAC bypass techniques exist, it is still prudent to use the highest enforcement level for UAC when possible and mitigate bypass opportunities that exist with techniques such as DLL Search Order Hijacking.
15		Abuse Elevation Control Mechanism: Elevated Execution with Prompt https://attack.mitre.org/techniques/T1548/004/	• System settings can prevent applications from running that haven't been downloaded through the Apple Store which may help mitigate some of these issues. Not allowing unsigned applications from being run may also mitigate some risk.
16		Create or Modify System Process: Windows Service https://attack.mitre.org/techniques/T1543/003/	• Use auditing tools capable of detecting privilege and service abuse opportunities on systems within an enterprise and correct them. • Limit privileges of user accounts and groups so that only authorized administrators can interact with service changes and service configurations.
17		Create or Modify System Process: Systemd Service https://attack.mitre.org/techniques/T1543/002/	• Restrict software installation to trusted repositories only and be cautious of orphaned software packages. • The creation and modification of systemd service unit files are generally reserved for administrators such as the Linux root user and other users with superuser privileges.

			• Restrict read write access to systemd unit files to only select privileged users who have a legitimate need to manage system services. • Limit user access to system utilities such as 'systemctl' to only users who have a legitimate need.
18		Exploitation for Privilege Escalation https://attack.mitre.org/techniques/T1068/	• Make it difficult for adversaries to advance their operation through exploitation of undiscovered or unpatched vulnerabilities by using sandboxing. Other types of virtualizations and application microsegmentations may also mitigate the impact of some types of exploitation. Risks of additional exploits and weaknesses in these systems may still exist. • Security applications that look for behavior used during exploitation such as Windows Defender Exploit Guard (WDEG) and the Enhanced Mitigation Experience Toolkit (EMET) can be used to mitigate some exploitation behavior. Control flow integrity checking is another way to potentially identify and stop a software exploit from occurring. Many of these protections depend on the architecture and target application binary for compatibility and may not work for software components targeted for privilege escalation. • Develop a robust cyber threat intelligence capability to determine what types and levels of threat may use software exploits and 0-days against a particular organization. • Update software regularly by employing patch management for internal enterprise endpoints and servers.
19		Process Injection https://attack.mitre.org/techniques/T1055/	• Some endpoint security solutions can be configured to block some types of process injection based on common sequences of behavior that occur during the injection process. • Utilize Yama to mitigate ptrace-based process injection by restricting the use of ptrace to privileged users only. Other mitigation controls involve the deployment of security kernel

			modules that provide advanced access control and process restrictions such as SELinux, grsecurity, and AppArmor.
20		Valid Accounts https://attack.mitre.org/techniques/T1078/	• Ensure that applications do not store sensitive data or credentials insecurely. • Applications and appliances that utilize default username and password should be changed immediately after the installation, and before deployment to a production environment. When possible, applications that use SSH keys should be updated periodically and properly secured. • Audit domain and local accounts as well as their permission levels routinely to look for situations that could allow an adversary to gain wide access by obtaining credentials of a privileged account. These audits should also include if default accounts have been enabled, or if new local accounts are created that have not be authorized. Follow best practices for design and administration of an enterprise network to limit privileged account use across administrative tiers.
21	Defense Evasion	Access Token Manipulation: Token Impersonation/Theft https://attack.mitre.org/techniques/T1134/001/	• Limit permissions so that users and user groups cannot create tokens. This setting should be defined for the local system account only. Also, define who can create a process level token to only the local and network service. • Administrators should log in as a standard user but run their tools with administrator privileges using the built-in access token manipulation command *runas*. • An adversary must already have administrator level access on the local system to make full use of this technique; be sure to restrict users and accounts to the least privileges they require.
22		Domain Policy Modification: Group Policy Modification https://attack.mitre.org/techniques/T1484/001/	• Identify and correct GPO permissions abuse opportunities using auditing tools such as BloodHound. • Consider implementing WMI and security filtering to further tailor which users and computers a GPO will apply to.

23		Domain Policy Modification: Domain Trust Modification https://attack.mitre.org/techniques/T1484/002/	• Use the principal of least privilege and protect administrative access to domain trusts.
24		File and Directory Permissions Modification: Windows File and Directory Permissions Modification https://attack.mitre.org/techniques/T1222/001/	• Ensure critical system files as well as those known to be abused by adversaries have restrictive permissions and are owned by an appropriately privileged account, especially if access is not required by users nor will inhibit system functionality. • Applying more restrictive permissions to files and directories could prevent adversaries from modifying the access control lists.
25		Impair Defenses: Disable Windows Event Logging https://attack.mitre.org/techniques/T1562/002/	• Ensure proper process and file permissions are in place to prevent adversaries from disabling or interfering logging. • Ensure proper Registry permissions are in place to prevent adversaries from disabling or interfering logging. • Ensure proper user permissions are in place to prevent adversaries from disabling or interfering with logging.
26		Impair Defenses: Disable or Modify Cloud Firewall https://attack.mitre.org/techniques/T1562/007/	• Routinely check account role permissions to ensure only expected users and roles have permission to modify cloud firewalls. • Ensure least privilege principles are applied to Identity and Access Management (IAM) security policies.
27		Modify Authentication Process: Domain Controller Authentication https://attack.mitre.org/techniques/T1556/001/	• Integrating multi-factor authentication as part of organizational policy can greatly reduce the risk of an adversary gaining control of valid credentials that may be used for additional tactics such as initial access, lateral movement, and collecting information. MFA can also be used to restrict access to cloud resources and APIs.

			• Audit domain and local accounts as well as their permission levels routinely to look for situations that could allow an adversary to gain wide access by obtaining credentials of a privileged account. These audits should also include if default accounts have been enabled, or if new local accounts are created that have not been authorized. Follow best practices for design and administration of an enterprise network to limit privileged account use across administrative tiers. • Enabled features, such as Protected Process Light, for LSA.
28		Modify System Image: Patch System Image https://attack.mitre.org/techniques/T1601/001/	• Some vendors of embedded network devices provide cryptographic signing to ensure the integrity of operating system images at boot time. Implement where available, following vendor guidelines. • Many vendors provide digitally signed operating system images to validate the integrity of the software used on their platform. Make use of this feature where possible in order to prevent and or detect attempts by adversaries to compromise the system image. • Some embedded network devices are capable of storing passwords for local accounts in either plain-text or encrypted formats. Ensure that, where available, local passwords are always encrypted, per vendor recommendations. • Use multi-factor authentication for user and privileged accounts. Most embedded network devices support TACACS+ and or RADIUS. Follow vendor prescribed best practices for hardening access control. • Refer to NIST guidelines when creating password policies. • Restrict administrator accounts to as few individuals as possible, following least privilege principles. Prevent credential overlap across systems of administrator and privileged accounts, particularly between network and non-network platforms, such as servers or endpoints.

29	Credential access	Credentials from Password Stores: Credentials from Web Browsers https://attack.mitre.org/techniques/T1555/003/	• Organizations may consider weighing the risk of storing credentials in web browsers. If web browser credential disclosure is a significant concern, technical controls, policy, and user training may be used to prevent storage of credentials in web browsers.
30		Exploitation for Credential Access https://attack.mitre.org/techniques/T1212/	• Make it difficult for adversaries to advance their operation through exploitation of undiscovered or unpatched vulnerabilities by using sandboxing. Other types of virtualizations and application microsegmentations may also mitigate the impact of some types of exploitation. Risks of additional exploits and weaknesses in these systems may still exist. • Security applications that look for behavior used during exploitation such as Windows Defender Exploit Guard and the Enhanced Mitigation Experience Toolkit can be used to mitigate some exploitation behavior. Control flow integrity checking is another way to potentially identify and stop a software exploit from occurring. Many of these protections depend on the architecture and target application binary for compatibility and may not work for software targeted for defense evasion. • Develop a robust cyber threat intelligence capability to determine what types and levels of threat may use software exploits and 0-days against a particular organization. • Update software regularly by employing patch management for internal enterprise endpoints and servers.
31		Man-in-the-Middle: ARP Cache Poisoning https://attack.mitre.org/techniques/T1557/002/	• A physical second factor key that uses the target login domain as part of the negotiation protocol will prevent session cookie theft through proxy methods.[7] • Configure browsers or tasks to regularly delete persistent cookies. • Train users to identify aspects of phishing attempts where they're asked to enter credentials into a site that

			has the incorrect domain for the application they are logging into.
32		Man in the middle https://attack.mitre.org/techniques/T1557/	• Disable legacy network protocols that may be used for MiTM if applicable and they are not needed within an environment. • Ensure that all wired and/or wireless traffic is encrypted appropriately. Use best practices for authentication protocols, such as Kerberos, and ensure web traffic that may contain credentials is protected by SSL/TLS. • Use network appliances and host-based security software to block network traffic that is not necessary within the environment, such as legacy protocols that may be leveraged for MiTM. • Limit access to network infrastructure and resources that can be used to reshape traffic or otherwise produce MiTM conditions. • Network intrusion detection and prevention systems that can identify traffic patterns indicative of MiTM activity can be used to mitigate activity at the network level. • Network segmentation can be used to isolate infrastructure components that do not require broad network access. This may mitigate, or at least alleviate, the scope of MiTM activity. • Train users to be suspicious about certificate errors. Adversaries may use their own certificates in an attempt to MiTM HTTPS traffic. Certificate errors may arise when the application's certificate does not match the one expected by the host.
33		Steal Web Session Cookie https://attack.mitre.org/techniques/T1539/	• A physical second factor key that uses the target login domain as part of the negotiation protocol will prevent session cookie theft through proxy methods. • Configure browsers or tasks to regularly delete persistent cookies.

			• Train users to identify aspects of phishing attempts where they're asked to enter credentials into a site that has the incorrect domain for the application they are logging into.
34	Discovery	Network Service Scanning https://attack.mitre.org/techniques/T1046/	• Ensure that unnecessary ports and services are closed to prevent risk of discovery and potential exploitation. • Use network intrusion detection prevention systems to detect and prevent remote service scans. • Ensure proper network segmentation is followed to protect critical servers and devices.
35	Lateral movement	Exploitation of Remote Services https://attack.mitre.org/techniques/T1210/	• Make it difficult for adversaries to advance their operation through exploitation of undiscovered or unpatched vulnerabilities by using sandboxing. Other types of virtualizations and application microsegmentations may also mitigate the impact of some types of exploitation. Risks of additional exploits and weaknesses in these systems may still exist. • Minimize available services to only those that are necessary. • Security applications that look for behavior used during exploitation such as Windows Defender Exploit Guard (WDEG) and the Enhanced Mitigation Experience Toolkit (EMET) can be used to mitigate some exploitation behavior. Control flow integrity checking is another way to potentially identify and stop a software exploit from occurring. Many of these protections depend on the architecture and target application binary for compatibility and may not work for all software or services targeted. • Segment networks and systems appropriately to reduce access to critical systems and services to controlled methods. • Minimize permissions and access for service accounts to limit impact of exploitation. • Develop a robust cyber threat intelligence capability to determine what types and levels of threat may use software exploits and 0-days against a particular organization.

			• Update software regularly by employing patch management for internal enterprise endpoints and servers. • Regularly scan the internal network for available services to identify new and potentially vulnerable services.
36		Remote Service Session Hijacking https://attack.mitre.org/techniques/T1563/	• Disable the remote service if it is unnecessary. • Enable firewall rules to block unnecessary traffic between network security zones within a network. • Do not allow remote access to services as a privileged account unless necessary. • Limit remote user permissions if remote access is necessary.
37		Software Deployment Tools https://attack.mitre.org/techniques/T1072/	• Ensure proper system and access isolation for critical network systems through use of group policy. • Ensure proper system and access isolation for critical network systems through use of multi-factor authentication. • Ensure proper system isolation for critical network systems through use of firewalls. • Verify that account credentials that may be used to access deployment systems are unique and not used throughout the enterprise network. • Grant access to application deployment systems only to a limited number of authorized administrators. • If the application deployment system can be configured to deploy only signed binaries, then ensure that the trusted signing certificates are not co-located with the application deployment system and are instead located on a system that cannot be accessed remotely or to which remote access is tightly controlled. • Patch deployment systems regularly to prevent potential remote access through Exploitation for Privilege Escalation. • Ensure that any accounts used by third-party providers to access these systems are traceable to the third-party and are not used throughout the network or used by other third-party providers in the same environment. Ensure there are regular

			reviews of accounts provisioned to these systems to verify continued business need, and ensure there is governance to trace de-provisioning of access that is no longer required. Ensure proper system and access isolation for critical network systems through use of account privilege separation. • Have a strict approval policy for use of deployment systems.
38	Collection	Automated Collection https://attack.mitre.org/techniques/T1119/	• Encryption and off-system storage of sensitive information may be one way to mitigate collection of files, but may not stop an adversary from acquiring the information if an intrusion persists over a long period of time and the adversary is able to discover and access the data through other means. Strong passwords should be used on certain encrypted documents that use them to prevent offline cracking through Brute Force techniques.
39		Data from Information Repositories: Sharepoint https://attack.mitre.org/techniques/T1213/002/	• Consider periodic review of accounts and privileges for critical and sensitive SharePoint repositories. • Enforce the principle of least-privilege. Consider implementing access control mechanisms that include both authentication and authorization. • Develop and publish policies that define acceptable information to be stored in SharePoint repositories.
40		Data from Cloud Storage Object https://attack.mitre.org/techniques/T1530/	• Frequently check permissions on cloud storage to ensure proper permissions are set to deny open or unprivileged access to resources. • Encrypt data stored at rest in cloud storage. Managed encryption keys can be rotated by most providers. At a minimum, ensure an incident response plan to storage breach includes rotating the keys and test for impact on client applications. • Cloud service providers support IP-based restrictions when accessing cloud resources. Consider using IP allow listing along with user account management to ensure that data

			access is restricted not only to valid users but only from expected IP ranges to mitigate the use of stolen credentials to access data. • Consider using multi-factor authentication to restrict access to resources and cloud storage APIs. • Use access control lists on storage systems and objects. • Configure user permissions groups and roles for access to cloud storage. • Implement strict Identity and Access Management (IAM) controls to prevent access to storage solutions except for the applications, users, and services that require access. Ensure that temporary access tokens are issued rather than permanent credentials, especially when access is being granted to entities outside of the internal security boundary.
41	Command and control	Data encoding https://attack.mitre.org/techniques/T1132/001/	• Network intrusion detection and prevention systems that use network signatures to identify traffic for specific adversary malware can be used to mitigate activity at the network level. Signatures are often for unique indicators within protocols and may be based on the specific obfuscation technique used by a particular adversary or tool, and will likely be different across various malware families and versions. Adversaries will likely change tool C2 signatures over time or construct protocols in such a way as to avoid detection by common defensive tools.
42		Junk data https://attack.mitre.org/techniques/T1001/001/	• Network intrusion detection and prevention systems that use network signatures to identify traffic for specific adversary malware can be used to mitigate some obfuscation activity at the network level.
43		Dynamic Resolution: Domain Generation Algorithms https://attack.mitre.org/techniques/T1568/002/	• Network intrusion detection and prevention systems that use network signatures to identify traffic for specific adversary malware can be used to mitigate activity at the network level. Malware researchers can reverse engineer malware variants

			that use DGAs and determine future domains that the malware will attempt to contact, but this is a time and resource intensive effort.[1][15] Malware is also increasingly incorporating seed values that can be unique for each instance, which would then need to be determined to extract future generated domains. In some cases, the seed that a particular sample uses can be extracted from DNS traffic.[5] Even so, there can be thousands of possible domains generated per day; this makes it impractical for defenders to preemptively register all possible C2 domains due to the cost. • In some cases a local DNS sinkhole may be used to help prevent DGA-based command and control at a reduced cost.
44		Non-Standard port https://attack.mitre.org/techniques/T1571/	• Network intrusion detection and prevention systems that use network signatures to identify traffic for specific adversary malware can be used to mitigate activity at the network level. • Properly configure firewalls and proxies to limit outgoing traffic to only necessary ports for that particular network segment.
45	Exfiltration	Data Transfer Size Limits https://attack.mitre.org/techniques/T1030/	• Network intrusion detection and prevention systems that use network signatures to identify traffic for specific adversary command and control infrastructure and malware can be used to mitigate activity at the network level.
46		Exfiltration Over Alternative Protocol https://attack.mitre.org/techniques/T1048/	• Enforce proxies and use dedicated servers for services such as DNS and only allow those systems to communicate over respective ports protocols, instead of all systems within a network. • Network intrusion detection and prevention systems that use network signatures to identify traffic for specific adversary command and control infrastructure and malware can be used to mitigate activity at the network level.

			• Follow best practices for network firewall configurations to allow only necessary ports and traffic to enter and exit the network.
47		Transfer data to cloud account https://attack.mitre.org/techniques/T1537/	• Implement network-based filtering restrictions to prohibit data transfers to untrusted VPCs. • Consider rotating access keys within a certain number of days to reduce the effectiveness of stolen credentials. • Limit user account and IAM policies to the least privileges required. Consider using temporary credentials for accounts that are only valid for a certain period of time to reduce the effectiveness of compromised accounts.
48	Impact	Abuse Elevation Control Mechanism https://attack.mitre.org/techniques/T1548/	• Check for common UAC bypass weaknesses on Windows systems to be aware of the risk posture and address issues where appropriate. • System settings can prevent applications from running that haven't been downloaded from legitimate repositories which may help mitigate some of these issues. Not allowing unsigned applications from being run may also mitigate some risk. • • Applications with known vulnerabilities or known shell escapes should not have the setuid or setgid bits set to reduce potential damage if an application is compromised. Additionally, the number of programs with setuid or setgid bits set should be minimized across a system. Ensuring that the sudo tty_tickets setting is enabled will prevent this leakage across tty sessions. • Remove users from the local administrator group on systems. • By requiring a password, even if an adversary can get terminal access, they must know the password to run anything in the sudoers file. Setting the timestamp_timeout to 0 will require the user to input their password every time sudo is executed.

			• The sudoers file should be strictly edited such that passwords are always required and that users can't spawn risky processes as users with higher privilege. • Although UAC bypass techniques exist, it is still prudent to use the highest enforcement level for UAC when possible and mitigate bypass opportunities that exist with techniques such as DLL Search Order Hijacking.
49		Network denial of service https://attack.mitre.org/techniques/T1498/	• When flood volumes exceed the capacity of the network connection being targeted, it is typically necessary to intercept the incoming traffic upstream to filter out the attack traffic from the legitimate traffic. Such defenses can be provided by the hosting Internet Service Provider (ISP) or by a 3rd party such as a Content Delivery Network (CDN) or providers specializing in DoS mitigations. • Depending on flood volume, on-premises filtering may be possible by blocking source addresses sourcing the attack, blocking ports that are being targeted, or blocking protocols being used for transport. • As immediate response may require rapid engagement of 3rd parties, analyze the risk associated with critical resources being affected by Network DoS attacks and create a disaster recovery plan business continuity plan to respond to incidents.
50		Data manipulation https://attack.mitre.org/techniques/T1565/	• Consider encrypting important information to reduce an adversary's ability to perform tailored data modifications. • Identify critical business and system processes that may be targeted by adversaries and work to isolate and secure those systems against unauthorized access and tampering. • Consider implementing IT disaster recovery plans that contain procedures for taking regular data backups that can be used to restore organizational data. Ensure backups are stored off system and are protected from common methods adversaries may use to gain access and manipulate backups.

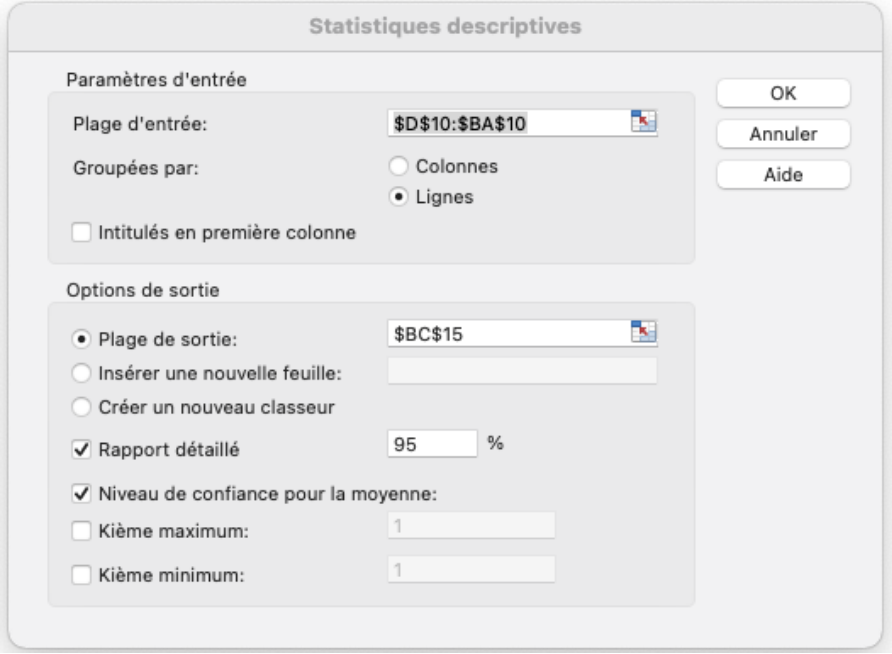
			• Ensure least privilege principles are applied to important information resources to reduce exposure to data manipulation risk.

16 Appendix E: MITRE ATT&CK scenarios query results

Doc	Nb	TP	TN	FP	FN	Precision	Recall	F1-score	MCC
Q1	363	178	38	55	92	0.76	0.66	**0.71**	0.06
Q2	363	191	41	52	79	0.79	0.71	**0.74**	0.14
Q3	363	167	38	55	103	0.75	0.62	**0.68**	0.02
Q4	363	179	34	59	91	0.75	0.66	**0.70**	0.03
Q5	363	35	86	7	235	0.83	0.13	**0.22**	0.07
Q6	363	166	46	47	104	0.78	0.61	**0.69**	0.10
Q7	363	183	37	56	87	0.77	0.68	**0.72**	0.07
Q8	363	170	40	53	100	0.76	0.63	**0.69**	0.05
Q9	363	159	49	44	111	0.78	0.59	**0.67**	0.10
Q10	363	127	52	41	133	0.77	0.51	**0.61**	0.04
Q11	363	140	47	46	130	0.75	0.52	**0.61**	0.02
Q12	363	199	31	62	71	0.76	0.74	**0.75**	0.07
Q13	363	199	32	61	71	0.77	0.74	**0.75**	0.08
Q14	363	151	46	47	119	0.76	0.56	**0.65**	0.05
Q15	363	217	24	69	53	0.76	0.80	**0.78**	0.07
Q16	363	195	29	64	75	0.75	0.72	**0.74**	0.03
Q17	363	122	60	33	148	0.79	0.45	**0.57**	0.09
Q18	363	205	30	63	65	0.76	0.76	**0.76**	0.08
Q19	363	213	24	69	57	0.76	0.79	**0.77**	0.05
Q20	363	179	34	59	91	0.75	0.66	**0.70**	0.03
Q21	363	181	39	54	89	0.77	0.67	**0.72**	0.08
Q22	363	170	36	57	100	0.75	0.63	**0.68**	0.02
Q23	363	171	43	50	99	0.77	0.63	**0.70**	0.09
Q24	363	197	26	67	73	0.75	0.73	**0.74**	0.01
Q25	363	200	26	67	70	0.75	0.74	**0.74**	0.02
Q26	363	201	28	65	69	0.76	0.74	**0.75**	0.04
Q27	363	106	62	31	164	0.77	0.39	**0.52**	0.05
Q28	363	157	51	42	113	0.79	0.58	**0.67**	0.11
Q29	363	207	29	64	63	0.76	0.77	**0.77**	0.08
Q30	363	205	30	63	65	0.76	0.76	**0.76**	0.08
Q31	363	7	91	2	263	0.78	0.03	**0.05**	0.01
Q32	363	114	52	41	156	0.74	0.42	**0.54**	−0.02
Q33	363	114	52	41	156	0.74	0.42	**0.54**	−0.02
Q34	363	184	38	55	86	0.77	0.68	**0.72**	0.08
Q35	363	166	46	47	104	0.78	0.61	**0.69**	0.10

Q36	363	86	73	20	184	0.81	0.32	**0.46**	0.10
Q37	363	183	37	56	87	0.77	0.68	**0.72**	0.07
Q38	363	215	24	69	55	0.76	0.80	**0.78**	0.06
Q39	363	145	56	37	125	0.80	0.54	**0.64**	0.12
Q40	363	144	48	45	126	0.76	0.53	**0.63**	0.04
Q41	363	215	24	69	55	0.76	0.80	0.78	0.06
Q42	363	215	28	65	55	0.77	0.80	0.78	0.10
Q43	363	204	28	65	66	0.76	0.76	0.76	0.06
Q44	363	158	43	50	112	0.76	0.59	0.66	0.04
Q45	363	217	26	67	53	0.76	0.80	0.78	0.09
Q46	363	186	31	62	84	0.75	0.69	0.72	0.02
Q47	363	162	46	47	108	0.78	0.60	0.68	0.08
Q48	363	160	48	45	110	0.78	0.59	0.67	0.10
Q49	363	164	45	48	106	0.77	0.61	0.68	0.08
Q50	363	222	22	71	48	0.76	0.82	0.79	0.07
					Mean	**0.77**	**0.63**	**0.67**	**0.06**
					SD	0.02	0.16	0.13	0.03
					Variance	0.000	0.027	0.018	0.001
					Number of observations	50	50	50	50

17 Appendix F: Student test results from Excel



Confidence interval of 95%	
Mean	0.67261045
Standard Error	0.01895631
Median	0.70472441
Mode	0.70472441
Standard Deviation	0.13404138
Sample Variance	0.01796709
Kurtosis	10.7677638
Skewness	−2.9693163
Range	0.73845312
Minimum	0.05017921
Maximum	0.78863233
Sum	33.6305226
Count	50
Confidence Level (95.0%)	0.03809414

18 Appendix G: Interview invitation

Dans les dernières semaines, vous avez vu passer un sondage initial que nous avons réalisé sur les compétences et les connaissances en sécurité de l'information dans notre organisation. Selon les données que nous avons recueillies dans le sondage, il semble que l'ensemble des domaines du cadre de référence en cybersécurité NICE Framework sont couverts dans nos équipes de cybersécurité. De façon générale, les technologies en place sont maîtrisées. De plus, nous avons observé que nous avons beaucoup d'individus diplômés qui détiennent les principales certifications en sécurité. Nos collaborateurs nous disent qu'ils estiment avoir de bonnes connaissances de l'ensemble de nos principaux domaines de sécurité et des connaissances approfondies dans leur domaine d'activité spécifique.

Plus spécifiquement, voici le portrait de ce que nous avons vu dans votre secteur.

(Ajouter un tableau pour chacun des secteurs)

Afin de compléter notre analyse de la situation actuelle, pour nous aider à planifier nos besoins de formation et à optimiser la gestion des talents, nous aimerions réaliser une entrevue avec vous. Nous sollicitons une rencontre de 45 à 60 minutes. Lors de ces rencontres, nous souhaitons approfondir les résultats dans votre secteur.

19 Appendix H: semi-structured interview guide

Guide d'entrevue avec les participants en sécurité de l'information sur les compétences professionnelles en cybersécurité

Nom : _____

Local : _____

Date : _____ Fin :_____

Rôle :

Chef d'équipe:

Direction :

Direction Principale :

Vous avez vu passer un sondage que nous avons réalisé sur les compétences et les connaissances en sécurité de l'information dans notre organisation. Afin de compléter notre analyse de la situation actuelle, nous souhaitons approfondir les résultats dans votre secteur.

(Avoir en mail les résultats pour leur secteur, qui leur a été envoyé avec l'invitation)

1. Quand on regarde les données du sondage, que j'ai ici, mais que je vous ai déjà envoyé avec l'invitation, quel est votre impression?

a. Quels sont vos principaux enjeux en matière de gestion des compétences dans votre secteur ?

b. En matière de compétences, connaissances, habiletés et tâches requises dans votre secteur, quels sont vos besoins à court terme?

c. À moyen terme ?

2. Parmi vos ressources humaines actuelles,

a. Avez-vous les rôles (acteurs ou ressources) et les expertises requises?

b. Sinon, quelles sont les lacunes que vous identifiez?

3. Est-ce que vous anticipez des pénuries de ressources humaines compétentes ?

a. Si oui, avez-vous un plan ou une stratégie en place pour combler les lacunes et les pénuries de ressources?

b. Avez-vous identifié, dans votre secteur, des ressources ou des compétences qui sont particulièrement difficiles à trouver ?

c. Est-ce que la formation des ressources en place est une de vos stratégies?

4. Quels sont vos besoins de formation?

a. Avez-vous besoin de certification des membres de vos équipes? Lesquelles?

b. Comment pouvons-nous vous aider?

(E.g. : formations, vidéos, conférences, conférenciers, ateliers, etc.)

c. Quelles formations pouvons-nous mettre en place?

20 Appendix I: Discussion group invitation

Groupe 1 : métiers techniques en cybersécurité

Bonjour,

Suite aux entrevues et à l'atelier réalisés avec les gestionnaires de (*confidentiel*) dans le cadre du projet d'évolution des métiers, nous souhaitons faire un nouvel atelier afin de valider nos résultats sur les métiers de la cybersécurité. Vous avez été identifiés comme participant pour les métiers de la famille de métiers techniques de la Cybersécurité.

L'atelier aura lieu le 9 avril 2020 au local (*confidentiel*) de 13h00 à 16h00. Vous n'avez rien à préparer pour la rencontre, il s'agit de discuter des compétences et savoirs qui sont nécessaires pour occuper un poste comme le vôtre dans nos équipes.

La participation à la rencontre est volontaire. Nous demanderons aux participants de signer un formulaire de consentement, car certaines données, anonymisées, pourront être utilisées dans le cadre d'un projet de recherche sur les métiers de la cybersécurité que nous réalisons en collaboration avec CyberÉco et un groupe de chercheurs.

Liste des participants et leur équipe invités au groupe 1:

- Participant 1, Évolution opérationnelle
- Participant 2, Évolution opérationnelle
- Participant 3, Évolution opérationnelle
- Participant 4, Tactique
- Participant 5, Tactique
- Participant 6, Sécurité offenvise
- Participant 7, Sécurité offenvise
- Participant 8, Sécurité offenvise
- Participant 9, Vulnérabilités
- Participant 10, Vulnérabilités
- Participant 11, Vulnérabilités
- Participant 12, Vulnérabilités
- Participant 13, Investigation
- Participant 14, Investigation
- Participant 15, Investigation
- Participant 16, Recherche
- Participant 17, Renseignements

Groupe 2 : métiers Affaires en cybersécurité

Bonjour,

Suite aux entrevues et à l'atelier réalisés avec les gestionnaires de (confidentiel) dans le cadre du projet d'évolution des métiers, nous souhaitons faire un nouvel atelier afin de valider nos résultats sur les métiers de la cybersécurité. Vous avez été identifiés comme participant pour les métiers de la famille de métiers d'affaires de la Cybersécurité.

L'atelier aura lieu le 10 avril 2020 au local (confidentiel) de 13h00 à 16h00. Vous n'avez rien à préparer pour la rencontre, il s'agit de discuter des compétences et savoirs qui sont nécessaires pour occuper un poste comme le vôtre dans nos équipes.

La participation à la rencontre est volontaire. Nous demanderons aux participants de signer un formulaire de consentement, car certaines données, anonymisées, pourront être utilisées dans le cadre d’un projet de recherche sur les métiers de la cybersécurité que nous réalisons en collaboration avec CyberÉco et un groupe de chercheurs.

Liste des participants et leur équipe invités au groupe 2:

Participant 1, RSI

Participant 2, RSI

Participant 3, RSI

Participant 4, Services-Conseil en sécurité

Participant 5, Services-Conseil en sécurité

Participant 6, Services-Conseil en sécurité

Participant 7, Gouvernance

Participant 8, Gouvernance

Participant 9, Gouvernance

Participant 10, Sensibilisation

Participant 11, Sensibilisation

Participant 12, Sensibilisation

Participant 13, Réalisation de projets

Participant 14, Réalisation de projets

Participant 15, Réalisation de projets

Participant 16, Gestion des Identités et des Accès

Participant 17, Gestion des Identités et des Accès

Participant 18, Gestion des Identités et des Accès

21 Appendix J: Study mind map

22 Appendix K: Workshop material

Atelier : Imaginez l'évolution du métier cybersécurité

La livraison sécuritaire de services aux membres et clients demande à maintenir le fragile équilibre entre risque et opportunité, entre transformation numérique innovante et prudence.

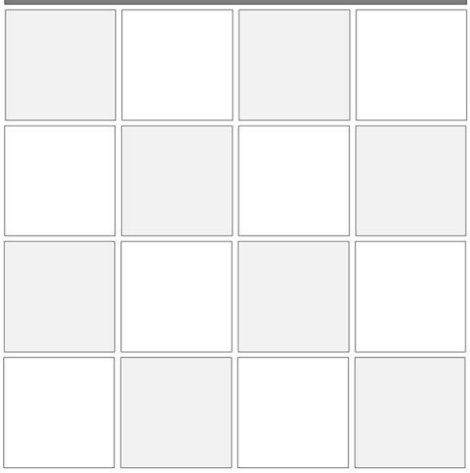
Assurer la confidentialité, l'intégrité et la disponibilité des informations est un défi majeur pour toutes les entreprises. Afin de faire face aux nombreux défis, nous devons identifier les compétences, connaissances et habiletés requises des membres de nos équipes. Ainsi, nous allons réaliser cet atelier de co-création afin d'alimenter notre réflexion.

Dans l'atelier d'aujourd'hui, nous allons imaginer les différents métiers de la cybersécurité qui s'offrent à un futur employé, Philippe, le fils de vos voisins. Philippe s'est tourné vers vous car il sait que vous travaillez en cybersécurité dans une grande entreprise. Il sollicite des conseils à un moment charnière, alors qu'il doit faire des choix pour son entrée à l'université, l'an prochain.

C'est l'histoire de Philippe…

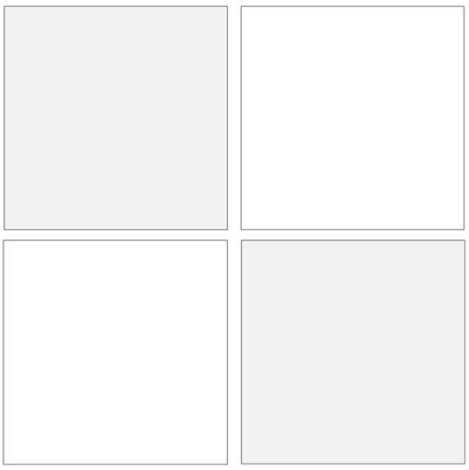
Philippe est né au Québec, en Montérégie, en 2000. Son père est originaire du Vietnam et sa mère de Chicoutimi. Il termine son CÉGEP en avril prochain avec une moyenne de plus de 80%. Il pense étudier à HEC Montréal ou à Sherbrooke en Cybersécurité. Passionné d'informatique, il est assez Geek, amateur de jeux vidéos, de voyages, de vélo et de ski.

1: Les postes



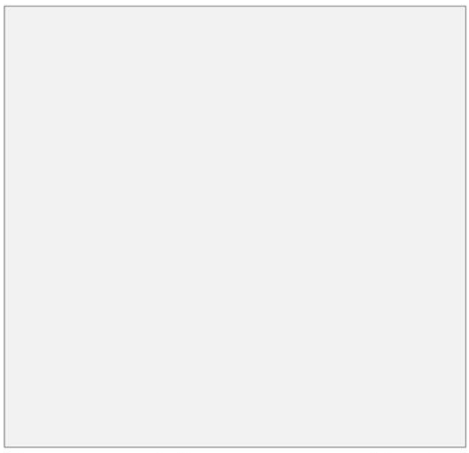
Individuellement: imaginez des postes qui pourraient intéresser Philippe dans nos équipes de sécurité.

2: Les métiers



En groupe: regrouper ces postes en catégories ou en familles de métier de la cybersécurité.

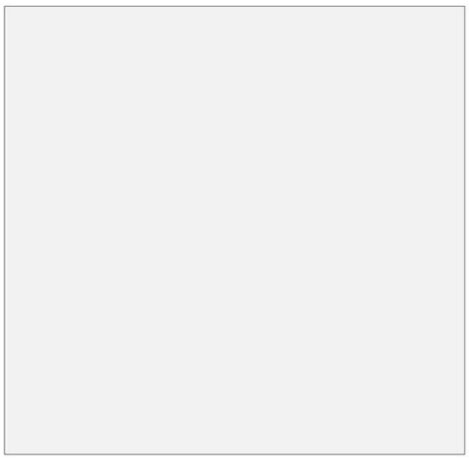
3: Cas d'utilisation



En groupe: réalisez une courte histoire du futur de Philippe dans nos équipes cybersécurité.

Il était une fois Philippe…

4: Compétences



Discussion ouverte avec tous les participants: pour chacun des métiers, identifiez les compétences que Phillipe doit développer.

23 Appendix L: Graph ?doc query results

This appendix presents in the column analyst the results of the expert classification of the cybersecurity system analyst and in the column GRAPH ?doc the results of the ontology classification with the following query:

```
prefix stardogapi: <tag:stardog:api:>

select ?entity ?doc ?mention ?type ?label where

{

 graph ?doc

 {

     ?doc stardog:docs:hasEntity ?entity .

     ?entity <http://purl.org/dc/terms/references> ?mention

  }

  ?mention a ?type ; rdfs:label ?label

}

Order by ?doc
```

Doc	Analyst	GRAPH ?doc
0	0	
1	1	
2	2	
3	3	3
4	4	4
5	5	5
6	6	6
7	7	

8		8
9	9	
10	10	10
11		11
12		
13		13
14	14	14
15		15
16		16
17		17
18		18
19	19	19
20		20
21	21	21
22	22	22
23		
24		
25		25
26	26	26
27		27
28		28
29	29	29
30	30	30
31	31	31
32		32
33		33
34		34
35	35	35
36		36
37	37	37
38		38
39		39
40		40
41	41	
42		
43		43
44	44	44
45	45	45

46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51		
52	52	52
53		
54	54	54
55		55
56	56	56
57	57	57
58	58	58
59		59
60	60	
61	61	61
62	62	62
63	63	63
64	64	
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	
71	71	
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	
79	79	79
80	80	80
81	81	81
82	82	
83	83	83

84		
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90		90
91		91
92		92
93		93
94	94	94
95	95	95
96	96	96
97	97	
98	98	98
99	99	99
100	100	
101	101	101
102	102	102
103		103
104		104
105		105
106		106
107	107	107
108		
109	109	109
110	110	110
111	111	111
112		112
113	113	113
114	114	114
115	115	115
116	116	116
117	117	117
118		118
119	119	119
120	120	120

121	121	
122	122	122
123		123
124	124	
125		125
126	126	126
127		
128	128	128
129		129
130		130
131		
132	132	
133	133	133
134	134	134
135	135	135
136	136	136
137	137	137
138		138
139		
140	140	140
141	141	141
142	142	
143	143	143
144	144	144
145		145
146		146
147	147	147
148	148	148
149	149	149
150	150	150
151	151	
152	152	152
153	153	153
154	154	
155	155	155
156	156	
157	157	157
158	158	158

159	159	159
160	160	160
161	161	161
162	162	162
163	163	163
164	164	164
165	165	
166	166	166
167	167	167
168	168	168
169	169	169
170	170	170
171	171	171
172	172	172
173	173	173
174	174	174
175	175	175
176	176	
177		177
178		
179	179	
180	180	180
181	181	181
182	182	182
183	183	183
184	184	184
185		185
186		186
187	187	
188		188
189	189	
190	190	
191	191	191
192	192	192
193	193	193
194	194	194
195	195	195
196	196	196

197	197	197
198	198	198
199	199	199
200	200	200
201	201	201
202	202	202
203	203	203
204		204
205		205
206		
207	207	207
208		208
209	209	209
210		
211	211	211
212	212	212
213	213	213
214	214	214
215	215	215
216	216	216
217	217	217
218	218	218
219	219	219
220	220	220
221	221	221
222	222	222
223		223
224		224
225	225	225
226	226	226
227	227	227
228	228	228
229	229	
230	230	230
231	231	231
232	232	232
233	233	

234	234	234
235	235	235
236		236
237		237
238	238	
239	239	239
240	240	240
241	241	241
242	242	242
243	243	243
244	244	244
245	245	
246	246	246
247	247	247
248	248	248
249	249	
250	250	250
251	251	251
252	252	252
253	253	253
254	254	254
255	255	
256	256	256
257	257	257
258	258	258
259	259	259
260	260	260
261	261	261
262	262	262
263	263	263
264	264	
265		265
266	266	266
267	267	267
268	268	268
269	269	269
270	270	270
271		

272		272
273	273	273
274		274
275		
276	276	276
277		277
278	278	
279		
280		280
281	281	281
282		282
283		283
284		284
285	285	285
286		286
287	287	287
288	288	288
289	289	289
290		290
291		291
292		
293		293
294	294	
295	295	295
296	296	296
297	297	297
298		298
299		299
300	300	300
301		301
302		302
303	303	303
304	304	304
305		305
306	306	306
307	307	307
308	308	308

309	309	
310	310	310
311	311	311
312		312
313	313	313
314	314	314
315	315	
316	316	316
317	317	317
318	318	
319	319	319
320	320	320
321	321	321
322	322	
323		
324	324	324
325	325	
326	326	
327	327	
328	328	328
329		
330	330	330
331	331	331
332	332	332
333	333	
334		
335	335	335
336	336	
337	337	337
338	338	338
339	339	339
340	340	340
341	341	341
342		342
343	343	343
344		
345		345
346	346	346

347	347	347
348	348	348
349	349	
350	350	
351	351	351
352	352	352
353	353	
354		354
355	355	355
356	356	
357	357	357
358	358	358
359	359	359
360	360	360
361		361
362	362	362
Count	**270**	**292**

24 Appendix M: Test 5 query results

tactic	mitigation	Doc	score	mentions
"T1199"	"User Account Control"	stardog:docs:CyberSec005:Profile_%2810%29.pdf	19.468135833740234	"CISM,CISM,NCWF,NCWF,NICE,NICE,NIST,NIST,CISO,CISO,Chief Information Security Officer,Chief Information Security Officer"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting152.pdf	15.166747093200684	"Bachelor,Bachelor"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting153.pdf	14.950447082519531	"Bachelor,Bachelor"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting47.pdf	12.87369155883789	"DIGITAL FORENSICS,DIGITAL FORENSICS,Bachelor,Bachelor,Digital Forensics,Digital Forensics,FOR,FOR,GIAC,GIAC,Global Information Assurance Certification,Global Information Assurance Certification,NCWF,NCWF,NICE,NICE,NIST,NIST,Information Security,Information Security"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting47.pdf	12.87369155883789	"DIGITAL FORENSICS,DIGITAL FORENSICS,Bachelor,Bachelor,Digital Forensics,Digital Forensics,FOR,FOR,GIAC,GIAC,Global Information Assurance Certification,Global Information Assurance Certification,NCWF,NCWF,NICE,NICE,NIST,NIST,Information Security,Information Security"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting48.pdf	12.804733276367188	"DIGITAL FORENSICS,DIGITAL FORENSICS,Digital Forensics,Digital Forensics,FOR,FOR,GIAC,GIAC,Global Information Assurance Certification,Global Information Assurance Certification,NCWF,NCWF,NICE,NICE,NIST,NIST,Information Security,Information Security"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting48.pdf	12.804733276367188	"DIGITAL FORENSICS,DIGITAL FORENSICS,Digital Forensics,Digital Forensics,FOR,FOR,GIAC,GIAC,Global Information Assurance Certification,Global Information Assurance Certification,NCWF,NCWF,NICE,NICE,NIST,NIST,Information Security,Information Security"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting218.pdf	12.804733276367188	"DIGITAL FORENSICS,DIGITAL FORENSICS,Digital Forensics,Digital Forensics,FOR,FOR,GIAC,GIAC,Global Information Assurance Certification,Global Information Assurance Certification,NCWF,NCWF,NICE,NICE,NIST,NIST,Information Security,Information Security,CISO,CISO,Chief Information Security Officer,Chief Information Security Officer"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting133.pdf	12.804733276367188	"Bachelor,Bachelor"

"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting218.pdf	12.804733276367188	"DIGITAL FORENSICS,DIGITAL FORENSICS,Digital Forensics,Digital Forensics,FOR,FOR,GIAC,GIAC,Global Information Assurance Certification,Global Information Assurance Certification,NCWF,NCWF,NICE,NICE,NIST,NIST,Information Security,Information Security,CISO,CISO,Chief Information Security Officer,Chief Information Security Officer"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting133.pdf	12.804733276367188	"Bachelor,Bachelor"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting173.pdf	12.786844253540039	"Windows,Windows,NCWF,NCWF,NICE,NICE,NIST,NIST,Director,Director"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting173.pdf	12.786844253540039	"Windows,Windows,NCWF,NCWF,NICE,NICE,NIST,NIST,Director,Director"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting222.pdf	11.69290542602539	"DIGITAL FORENSICS,DIGITAL FORENSICS,Bachelor,Bachelor,Digital Forensics,Digital Forensics,FOR,FOR,NCWF,NCWF,NICE,NICE,NIST,NIST,Information Security,Information Security,CISO,CISO,Chief Information Security Officer,Chief Information Security Officer"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting222.pdf	11.69290542602539	"DIGITAL FORENSICS,DIGITAL FORENSICS,Bachelor,Bachelor,Digital Forensics,Digital Forensics,FOR,FOR,NCWF,NCWF,NICE,NICE,NIST,NIST,Information Security,Information Security,CISO,CISO,Chief Information Security Officer,Chief Information Security Officer"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting13.pdf	10.929532051086426	"Execution,Execution,TA0002,TA0002,Organizational Crime,Organizational Crime,Director,Director"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting219.pdf	10.772171974182129	"Director,Director,Director,Director"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting17.pdf	10.772171974182129	"GIAC,GIAC,Global Information Assurance Certification,Global Information Assurance Certification,NCWF,NCWF,NICE,NICE,NIST,NIST,Director,Director"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting219.pdf	10.772171974182129	"Director,Director,Director,Director"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting31.pdf	10.772171974182129	"Director,Director,Director,Director"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting31.pdf	10.772171974182129	"Director,Director,Director,Director"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting17.pdf	10.772171974182129	"GIAC,GIAC,Global Information Assurance Certification,Global Information Assurance Certification,NCWF,NCWF,NICE,NICE,NIST,NIST,Director,Director"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:Profile_%2838%29.pdf	10.288320541381836	"CISA,CISA,Director,Director"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting30.pdf	10.097443580627441	"Director,Director,GSEC,GSEC"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting1.pdf	10.097443580627441	"AKW0005,AKW0005,Bilingualism spoken,Bilingualism spoken,Bilinguisme parlé,Bilinguisme parlé"

"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting73.pdf	10.097443580627441	"Bachelor,Bachelor,Bachelor,Bachelor,NCWF,NCWF,NICE,NICE,NIST,NIST,Information Security,Information Security"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting107.pdf	9.998512268066406	"Bachelor,Bachelor,ANA,ANA,Systems Analysis,Systems Analysis"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting259.pdf	9.998512268066406	"Bachelor,Bachelor,Execution,Execution,TA0002,TA0002"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting177.pdf	9.391603469848633	"Bachelor,Bachelor"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting277.pdf	9.391603469848633	"Bachelor,Bachelor"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting21.pdf	9.293309211730957	"Director,Director"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting362.pdf	9.127788543701172	"Bachelor,Bachelor,Information Security,Information Security"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting150.pdf	9.127788543701172	"Bachelor,Bachelor"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting161.pdf	9.127788543701172	"Cybersecurity Roles,Cybersecurity Roles,Execution,Execution,TA0002,TA0002,NCWF,NCWF,NICE,NICE,NIST,NIST,CISA,CISA"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting74.pdf	9.127788543701172	"Bachelor,Bachelor,CISM,CISM,Information Security,Information Security"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting3.pdf	9.127788543701172	"Bachelor,Bachelor,Information Security,Information Security"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting104.pdf	8.461455345153809	"Bachelor,Bachelor,Execution,Execution,TA0002,TA0002,ANA,ANA,ANA,ANA,Systems Analysis,Systems Analysis,Systems Analysis,Systems Analysis"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting340.pdf	8.02006721496582	"Bachelor,Bachelor"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting227.pdf	8.02006721496582	"CISM,CISM,NCWF,NCWF,NICE,NICE,NIST,NIST,CISA,CISA,PMP,PMP,Director,Director"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting26.pdf	7.756251811981201	"CISM,CISM,Information Security,Information Security,Director,Director"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting181.pdf	7.756251811981201	"NCWF,NCWF,NICE,NICE,NIST,NIST"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:Profile_%2825%29.pdf	7.756251811981201	"CISSP,CISSP,CISM,CISM,Blue Team,Blue Team,Cybersécurité Défensive,Cybersécurité Défensive,Defensive,Defensive,Fraud,Fraud,Fraud,CISA,CISA"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting201.pdf	7.756251811981201	"CISM,CISM,CISA,CISA,Information Security,Information Security,Director,Director"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting136.pdf	7.756251811981201	"Bachelor,Bachelor"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting221.pdf	7.756251811981201	"Fraud,Fraud,Information Security,Information Security"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting214.pdf	7.756251811981201	"NCWF,NCWF,NICE,NICE,NIST,NIST,CISA,CISA,Unix/Linux,Unix/Linux,GSEC,GSEC"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting159.pdf	7.756251811981201	"Bachelor,Bachelor"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting59.pdf	7.756251811981201	"Certifications,Certifications"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting303.pdf	6.705291748046875	"Bachelor,Bachelor,CISM,CISM,Internal Controls,Internal Controls,CRISC,CRISC"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting96.pdf	6.705291748046875	"CISM,CISM,Internal Controls,Internal Controls,CRISC,CRISC"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting52.pdf	6.488991737365723	"Internal Auditors,Internal Auditors,CISM,CISM,Fraud,Fraud,NCWF,NCWF,NICE,NICE,NIST,NIST,CISA,CISA,Director,Director"

"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting283.pdf	6.488991737365723	"Internal Auditors,Internal Auditors,Internal Auditors,Internal Auditors,Bachelor,Bachelor,Fraud,Fraud"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting339.pdf	6.488991737365723	"NCWF,NCWF,NICE,NICE,NIST,NIST"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting37.pdf	6.488991737365723	"Masters,Masters,CISSP,CISSP,CISM,CISM,Execution,Execution,TA0002,TA0002,NCWF,NCWF,NICE,NICE,NIST,NIST,CISA,CISA,Information Security,Information Security,Director,Director"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:Profile_%282%29.pdf	6.441003322601318	"Execution,Execution,TA0002,TA0002,Director,Director"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting40.pdf	6.441003322601318	"Enterprise Architect,Enterprise Architect,SP-ARC-001,SP-ARC-001"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting300.pdf	6.189499855041504	"Internal Auditors,Internal Auditors,Bachelor,Bachelor,Execution,Execution,TA0002,TA0002"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting284.pdf	6.189499855041504	"Fraud,Fraud"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting134.pdf	6.189499855041504	"Bachelor,Bachelor,CISM,CISM,GIAC,GIAC,Global Information Assurance Certification,Global Information Assurance Certification,ISACA,ISACA,Execution,Execution,TA0002,TA0002,NCWF,NCWF,NICE,NICE,NIST,NIST,CISA,CISA,Information Security,Information Security,CRISC,CRISC"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting124.pdf	6.004515171051025	"Execution,Execution,TA0002,TA0002"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting124.pdf	6.004515171051025	"Execution,Execution,TA0002,TA0002"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting36.pdf	5.998623371124268	"Information Security,Information Security,Director,Director,GSEC,GSEC"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting228.pdf	5.962841033935547	"NET,NET,Network Services,Network Services,OM-ADM-001,OM-ADM-001,System Administrator,System Administrator,Unix/Linux,Unix/Linux"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting228.pdf	5.962841033935547	"NET,NET,Network Services,Network Services,OM-ADM-001,OM-ADM-001,System Administrator,System Administrator,Unix/Linux,Unix/Linux"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting125.pdf	5.879767417907715	"Bachelor,Bachelor"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting125.pdf	5.879767417907715	"Bachelor,Bachelor"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting128.pdf	5.803407192230225	"Windows,Windows,OM-ADM-001,OM-ADM-001,System Administrator,System Administrator"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting128.pdf	5.803407192230225	"Windows,Windows,OM-ADM-001,OM-ADM-001,System Administrator,System Administrator"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting126.pdf	5.755837917327881	"Bachelor,Bachelor"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting130.pdf	5.755837917327881	"Windows,Windows,PMP,PMP"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting126.pdf	5.755837917327881	"Bachelor,Bachelor"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting130.pdf	5.755837917327881	"Windows,Windows,PMP,PMP"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting225.pdf	5.7473931312561035	"CISM,CISM,AKW0005,AKW0005,Bilingualism spoken,Bilingualism spoken,Bilinguisme parlé,Bilinguisme parlé,NCWF,NCWF,NICE,NICE,NIST,NIST"

"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting252.pdf	5.7473931312561035	"Bachelor,Bachelor,Enterprise Architect,Enterprise Architect,SP-ARC-001,SP-ARC-001"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting330.pdf	5.7473931312561035	"Bachelor,Bachelor,Execution,Execution,TA0002,TA0002,NCWF,NCWF,NICE,NICE,NIST,NIST"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting199.pdf	5.7473931312561035	"Bachelor,Bachelor,CISM,CISM,NCWF,NCWF,NICE,NICE,NIST,NIST,CRISC,CRISC"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting222.pdf	5.7473931312561035	"DIGITAL FORENSICS,DIGITAL FORENSICS,Bachelor,Bachelor,Digital Forensics,Digital Forensics,FOR,FOR,NCWF,NCWF,NICE,NICE,NIST,NIST,Information Security,Information Security,CISO,CISO,Chief Information Security Officer,Chief Information Security Officer"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting311.pdf	5.7473931312561035	"Bachelor,Bachelor,CISM,CISM,NCWF,NCWF,NICE,NICE,NIST,NIST,CISA,CISA,Information Security,Information Security,CRISC,CRISC"
"T1199"	"User Account Control"	stardog:docs:CyberSec005:JobPosting226.pdf	5.7473931312561035	"Bachelor,Bachelor,CISM,CISM,NCWF,NCWF,NICE,NICE,NIST,NIST,PMP,PMP"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting242.pdf	5.699774742126465	"AN,AN,Analyse,Analyse,Bachelor,Bachelor,Analyse,Analyse,Bloom level 4,Bloom level 4,AKW0005,AKW0005,Bilingualism spoken,Bilingualism spoken,Bilinguisme parlé,Bilinguisme parlé,NCWF,NCWF,NICE,NICE,NIST,NIST,Director,Director"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting242.pdf	5.699774742126465	"AN,AN,Analyse,Analyse,Bachelor,Bachelor,Analyse,Analyse,Bloom level 4,Bloom level 4,AKW0005,AKW0005,Bilingualism spoken,Bilingualism spoken,Bilinguisme parlé,Bilinguisme parlé,NCWF,NCWF,NICE,NICE,NIST,NIST,Director,Director"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting339.pdf	5.551084995269775	"NCWF,NCWF,NICE,NICE,NIST,NIST"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting339.pdf	5.551084995269775	"NCWF,NCWF,NICE,NICE,NIST,NIST"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting65.pdf	5.551084995269775	"Bachelor,Bachelor,Bachelor,Bachelor,OM-ADM-001,OM-ADM-001,System Administrator,System Administrator"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting345.pdf	5.551084995269775	"Malware,Malware"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting345.pdf	5.551084995269775	"Malware,Malware"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting65.pdf	5.551084995269775	"Bachelor,Bachelor,Bachelor,Bachelor,OM-ADM-001,OM-ADM-001,System Administrator,System Administrator"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting184.pdf	5.449540615081787	"Windows,Windows,Audit,Audit,M1047,M1047,Intrusion Detection Systems,Intrusion Detection Systems"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting224.pdf	5.449540615081787	"Execution,Execution,TA0002,TA0002,AKW0005,AKW0005,Bilingualism spoken,Bilingualism spoken,Bilinguisme parlé,Bilinguisme parlé,Director,Director"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting129.pdf	5.449540615081787	"Malware,Malware"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting183.pdf	5.449540615081787	"Windows,Windows,Intrusion Detection Systems,Intrusion Detection Systems"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting129.pdf	5.449540615081787	"Malware,Malware"

"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting115.pdf	5.449540615081787	"Bachelor,Bachelor,CISM,CISM,SSH,SSH,Director,Director"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting184.pdf	5.449540615081787	"Windows,Windows,Audit,Audit,M1047,M1047,Intrusion Detection Systems,Intrusion Detection Systems"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting224.pdf	5.449540615081787	"Execution,Execution,TA0002,TA0002,AKW0005,AKW0005,Bilingualism spoken,Bilingualism spoken,Bilinguisme parlé,Bilinguisme parlé,Director,Director"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting183.pdf	5.449540615081787	"Windows,Windows,Intrusion Detection Systems,Intrusion Detection Systems"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting115.pdf	5.449540615081787	"Bachelor,Bachelor,CISM,CISM,SSH,SSH,Director,Director"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting40.pdf	5.319789886474609	"Enterprise Architect,Enterprise Architect,SP-ARC-001,SP-ARC-001"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting50.pdf	5.319789886474609	"Bachelor,Bachelor"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting171.pdf	5.319789886474609	"CISM,CISM,Execution,Execution,TA0002,TA0002,NCWF,NCWF,NICE,NICE,NIST,NIST,CISA,CISA,Information Security,Information Security"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting256.pdf	5.319789886474609	"Bachelor,Bachelor"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting171.pdf	5.319789886474609	"CISM,CISM,Execution,Execution,TA0002,TA0002,NCWF,NCWF,NICE,NICE,NIST,NIST,CISA,CISA,Information Security,Information Security"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting256.pdf	5.319789886474609	"Bachelor,Bachelor"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting50.pdf	5.319789886474609	"Bachelor,Bachelor"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting40.pdf	5.319789886474609	"Enterprise Architect,Enterprise Architect,SP-ARC-001,SP-ARC-001"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting180.pdf	5.148183822631836	"Information Security,Information Security,RSA,RSA"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting29.pdf	5.148183822631836	"Windows,Windows,GIAC,GIAC,Global Information Assurance Certification,Global Information Assurance Certification,NCWF,NCWF,NICE,NICE,NIST,NIST,Director,Director"
"T1199"	"Network Segmentation"@en	stardog:docs:CyberSec005:JobPosting240.pdf	5.148183822631836	"AKW0005,AKW0005,Bilingualism spoken,Bilingualism spoken,Bilinguisme parlé,Bilinguisme parlé,PMP,PMP"
"T1199"	"Network Segmentation"	stardog:docs:CyberSec005:JobPosting240.pdf	5.148183822631836	"AKW0005,AKW0005,Bilingualism spoken,Bilingualism spoken,Bilinguisme parlé,Bilinguisme parlé,PMP,PMP"

25 Appendix N: Work roles by specialty area

Cybersecurity Business roles

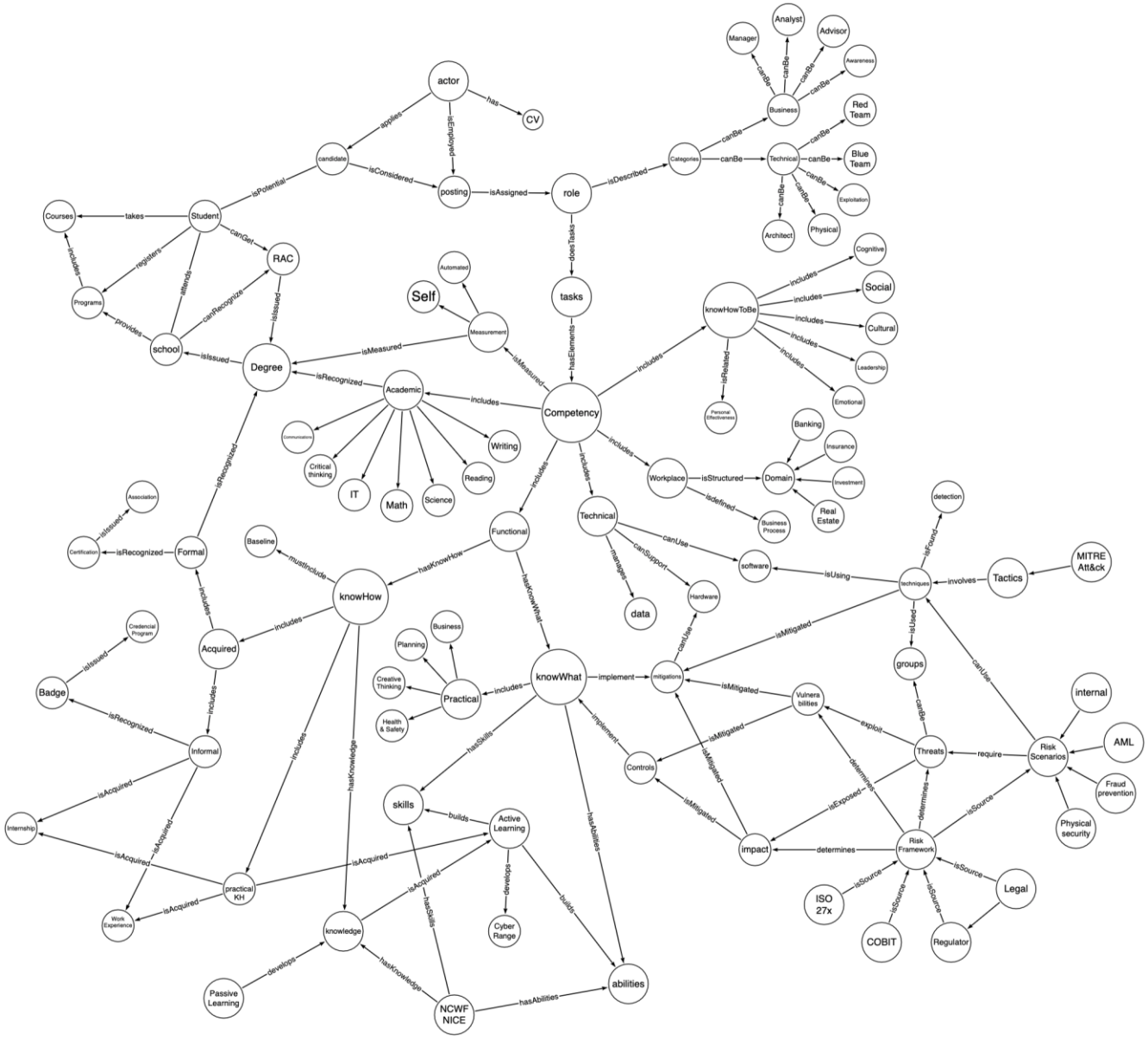
Speciality	NCWF Role	NCWF ID
Manager	Executive Cyber Leadership	OV-EXL-001
	Information Systems Security Manager	OV-MGT-001
	Communications Security (COMSEC) Manager	OV-MGT-002
	Cyber Workforce Developer and Manager	OV-SPP-001
Analyst	Systems Security Analyst	OM-ANA-001
	Cyber Intel Planner	CO-OPL-001
	Cyber Ops Planner	CO-OPL-002
	Authorizing Official/Designating Representative	SP-RSK-001
	Security Control Assessor	SP-RSK-002
	Research & Development Specialist	SP-TRD-001
	Systems Requirements Planner	SP-SRP-001
	Data Analyst	OM-DTA-002
	Knowledge Manager	OM-KMG-001
	IT Program Auditor	OV-PMA-005
	All Source-Collection Manager	CO-CLO-001
	All Source-Collection Requirements Manager	CO-CLO-002
Advisor	Enterprise Architect	SP-ARC-001
	Security Architect	SP-ARC-002
	Cyber Legal Advisor	OV-LGA-001
	Privacy Officer/Privacy Compliance Manager	OV-LGA-002
	Cyber Policy and Strategy Planner	OV-SPP-002
	Program Manager	OV-PMA-001
	IT Project Manager	OV-PMA-002
	Product Support Manager	OV-PMA-003
	IT Investment/Portfolio Manager	OV-PMA-004
Awareness	Cyber Instructional Curriculum Developer	OV-TEA-001
	Cyber Instructor	OV-TEA-002

Cybersecurity Technical roles

Specialty	NCWF Role	NCWF ID
RED Team	Mission Assessment Specialist	AN-ASA-002
	Target Developer	AN-TGT-001
	Target Network Analyst	AN-TGT-002
	Multi-Disciplined Language Analyst	AN-LNG-001
BLUE Team	Software Developer	SP-DEV-001
	Secure Software Assessor	SP-DEV-002
	System Testing and Evaluation Specialist	SP-TST-001
	Cyber Defense Analyst	PR-CDA-001
	Cyber Defense Infrastructure Support Specialist	PR-INF-001
	Cyber Defense Incident Responder	PR-CIR-001
	Vulnerability Assessment Analyst	PR-VAM-001
	Threat/Warning Analyst	AN-TWA-001
	All Source-Collection Manager	CO-CLO-001
	Cyber Crime Investigator	IN-INV-001
	Law Enforcement /CounterIntelligence Forensics Analyst	IN-FOR-001
	Cyber Defense Forensics Analyst	IN-FOR-002
Exploitation	Information Systems Security Developer	SP-SYS-001
	Systems Developer	SP-SYS-002
	Database Administrator	OM-DTA-001
	Technical Support Specialist	OM-STS-001
	Network Operations Specialist	OM-NET-001
	System Administrator	OM-ADM-001
	Exploitation Analyst	AN-EXP-001
	All-Source Analyst	AN-ASA-001
	Partner Integration Planner	CO-OPL-003
	Cyber Operator	CO-OPS-001
Cybersecurity Architect	Enterprise Architect	SP-ARC-001
	Security Architect	SP-ARC-002

26 Appendix O: Graphical database model

https://drive.google.com/file/d/1dbdCaqIQyYqOsjGlQ3KzJZc3MXW1_06b/view?usp=sharing



27 Appendix P: Competency evaluation questionnaire

Hi Marc-Andre, when you submit this form, the owner will be able to see your name and email address.

* Required

Role-specific Tasks

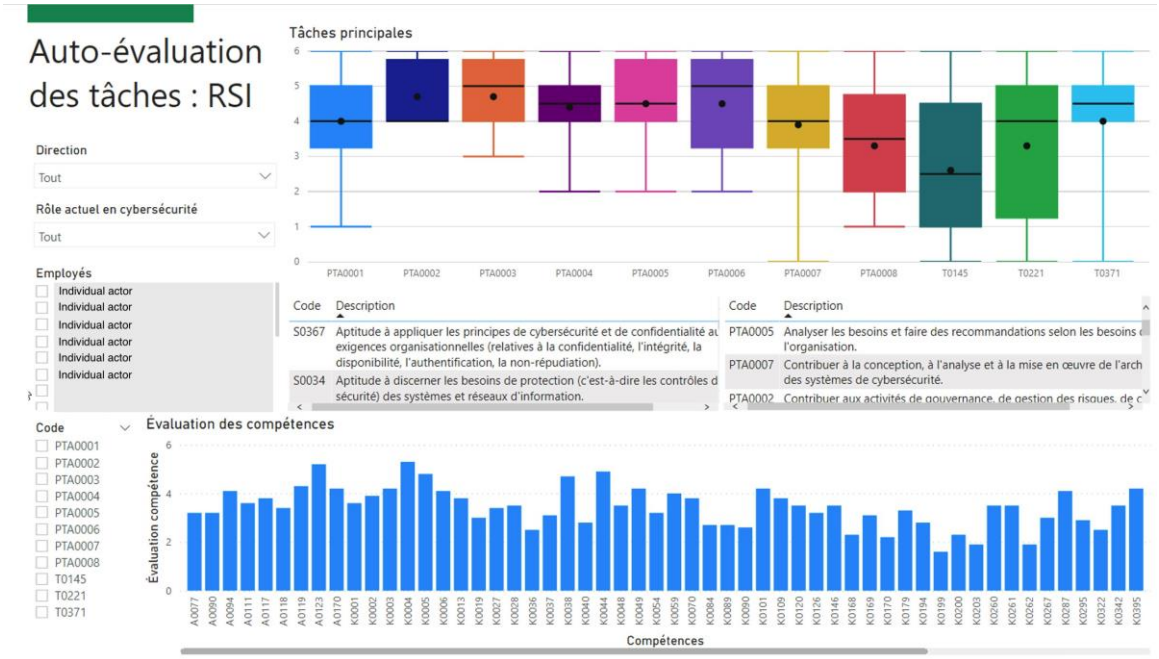
1. Please indicate your self-estimated level of competency on a scale from zero (0) to six (6). Zero (0) being the absence of any competency for a particular element and six (6) being the highest level of competency or expertize in the mentionned area. *

	0	1	2	3	4	5	6
Provide business context to cybersecurity. [PTA0001]	○	○	○	○	○	○	○
Contribute to governance, Risk Management, compliance and cybersecurity activities. [PTA0002]	○	○	○	○	○	○	○
Provide input on cybersecurity requirements. [PTA0003]	○	○	○	○	○	○	○

Translate business and user requirements into actionable cybersecurity information and technical recommendation. [PTA0004]	○	○	○	○	○	○	○
Analyse needs and make recommendations as required by the organization. [PTA0005]	○	○	○	○	○	○	○
Translate and explain the cybersecurity requirements to the stakeholders. [PTA0006]	○	○	○	○	○	○	○
Contribute to cybersecurity system architecture design, analysis ans implementation. [PTA0007]	○	○	○	○	○	○	○
Perform project management and change management activities. [PTA0008]	○	○	○	○	○	○	○
Manage and approve certification and conformity packages. [T0145]	○	○	○	○	○	○	○
Review authorization and assurance documents to confirm that the level of risk is within acceptable limits for each software application, system, and network. [T0221]	○	○	○	○	○	○	○
Establish acceptable limits for the software application, network, or system. [T0371]	○	○	○	○	○	○	○

Next

28 Appendix Q: Power BI dashboard for competency management



29 Appendix R: Summary of the ADR approach

ADR steps	Mapped principles	Operationalization	Outcome
Problem formulation	1: Research inspired by practice 2: Artifacts rooted in theory	• Scope, roles, research questions • Identify material and references	• research proposal • Initial concept map
Building, Intervention, Evaluation	3: Reciprocal shaping 4: Mutually influential roles 5: Authentic, simultaneous assessment	• Interviews • Workgroups	• Concept map • UML model • WebProtégé ontology • Protégé ontology • Stardog ontology
Reflexion and learning	6: Guided emergence	• Identify contributions • Adjust the research process • Ontology validation and testing	• Ontology engineering methodology • Ontology • Graph database • GitHub
Formalization of learning	7: Generalization of results	• Describe the achievements • Formalize the OWL ontology	• Dissertation • Presentations • Article(s)

30 Appendix S: Overview of the BIE cycles

BIE cycle	Principal activities	Outcome
1	• Individual interviews • Sharing of initial model with stakeholders • Continuous improvement of model	• Develop an understanding of the cybersecurity competency • Identify the frameworks are usefull • Create an initial model
2	• Workshop with managers • Sharing of model with stakeholders • Continuous improvement of model • Initial design of ontology	• Concept of the 2 categories and 9 specialties • Need to perform 2 more workgroups to understand the roles of the categories
3	• Workshop with business specialists • Workshop with technical specialists • Continuous improvement of model • Initial build of ontology in WebProtégé	• Understand the work roles • Determine the competency elements • Ontology engineering methodology
4	• Validation • Migration from WebProtégé to Protégé (desktop) • Design of SPARQL queries • Evaluation of ontology tools and graph databases • Implementation into Stardog • Design and test of queries • 3, then 4, then 5 series of queries • F1-Scores, then MCC	• SPARQL in Protégé • SPARQL in Stardog • Matching jobs to MITRE Att&ck scenarios • F1-Scores • Matthews Correlation Coefficient

31 Appendix T: Profile of participants (n=42) in the study

Semi-structured interviews with the managers of all areas (**n=8**)

No	Work role	Responsibilities	Years of experience	Years with organization
1	Director	Cybersecurity Projects	7	7
2	Director	Cybersecurity Strategy and Transformation	12	7
3	Manager	Cybersecurity incident response	26	7
4	Manager	Red Team management	22	9
5	Manager	Compliance	11	8
6	Team leader	Risk management	15	8
7	Team leader	Governance	10	5
8	Information Security Officer	Strategic business unit support (cybersecurity)	7	3

Workshop 1 with 12 participants (**n = 12**) in leadership positions in the cybersecurity

No	Work role	Responsibilities	Years of experience	Years with organization
1	CISO	Executive Cyber Leadership	15	10
2	Director	Cyberdefense	26	11
3	Director	GRC	12	7
4	Director	IAM	7	7
5	Manager	Information Systems Security	22	9
6	Manager	Information Systems Security	16	7
7	Manager	IT Investment/Portfolio	33	5
8	Manager	IT Investment/Portfolio	19	4
9	Team leader	Team Lead Cyberdefense	16	5
10	Team leader	Team Lead Red Team	9	3
11	Information Security Officer	Authorizing Official/Designating Representative	20	2
12	Information Security Officer	Authorizing Official/Designating Representative	23	4

Workshop 2 with staff from the business-oriented cybersecurity roles (**n = 10**)

No	Work role	Responsibilities	Years of experience	Years with organization
1	Analyst	Systems Security	4	2
2	Analyst	Systems Security	15	3
3	Analyst	Systems Security	8	3
4	Advisor	IT Program Auditor	12	2
5	Advisor	Systems Requirements Planner	8	3
6	Advisor	Systems Requirements Planner	11	3
7	Advisor	Project Manager	16	4

8	Advisor	Project Manager	15	4
9	Advisor	Cyber Policy and Strategy Planner	12	3
10	Awareness	Cyber Instructional Curriculum Developer	8	2

Workshop 3 with staff from the technical-oriented cybersecurity roles (**n = 12**)

No	Work role	Responsibilities	Years of experience	Years with organization
1	Analyst	Security Control Assessor	8	3
2	Advisor	Security Control Assessor	15	10
3	Analyst	Cyber Intel Planner	10	3
4	Analyst	Cyber Intel Planner	9	2
5	Advisor	Cyber Ops Planner	16	4
6	Analyst	Data Analyst	7	4
7	Advisor	Ethical hacker	5	3
8	Advisor	Ethical hacker	4	2
9	Advisor	Research & Development Specialist	20	4
10	Advisor	All Source-Collection Manager	8	2
11	Advisor	All Source-Collection Manager	10	2
12	Advisor	Security Architect	12	3