Pertanika J. Sci. & Technol. 33 (5): 2049 - 2068 (2025)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

*Review Article*

# Assessing the Role of Ontologies in Enhancing Various Modern Systems

**Sarah Dahir[1]* and Abderrahim El Qadi[2]**

*[1]Laboratory of Intelligent Systems and Applications, Faculty of Sciences and Technologies, Sidi Mohamed Ben Abdellah University, Fez 30000, Morocco*

*[2]Department of Applied Mathematics and Computer Engineering, National Graduate School of Arts and Crafts, Mohammed V University in Rabat 10000, Morocco*

## ABSTRACT

Increasing the discoverability, accessibility, and understandability of data for both humans and machines is the ultimate objective of the Semantic Web (SW). Therefore, the purpose of this work is to survey and gain a clear understanding of the current state of the use of Linked Open Data (LOD) across a range of domains. We discovered that, of the four domains we evaluated, the two that use ontologies the most are machine learning (ML) and artificial intelligence (AI) in general. On the other hand, because it is a relatively new domain, the Metaverse uses ontologies the least. Despite ontologies' capacity to guarantee consistency in the virtual world, increase revenue, ensure inclusivity for people with disabilities, and save time. Additionally, the majority of domains are not utilizing SW to its full potential, and additional customization is required in light of each domain's unique challenges and traits. For instance, AI, cybersecurity, and the Metaverse have an unstructured nature and lack stability. Also, cybersecurity and the Metaverse lack consensus. In addition to this, the Metaverse is highly scalable. Another common difficulty of incorporating ontologies in general is choosing the right mapping technique as there are many. Given these domains' characteristics, Business Intelligence (BI) finds it easier to integrate them, whereas cybersecurity and the Metaverse find it more difficult. Lastly, dynamic ontologies are believed to make ontologies more appropriate and adaptable for domains lacking stability.

*Keywords*: Artificial intelligence, augmented reality, Business Intelligence, cybersecurity, ontology, Semantic Web

ARTICLE INFO

*Article history:*
Received: 12 December 2024
Accepted: 08 April 2025
Published: 11 August 2025

DOI: https://doi.org/10.47836/pjst.33.5.02

*E-mail addresses*:
sarah.dahir@usmba.ac.ma (Sarah Dahir)
a.elqadi@um5r.ac.ma (Abderrahim El Qadi)
* Corresponding author

e-ISSN: 2231-8526

## INTRODUCTION

First and foremost, it is safe to confess that data is the ultimate fortune that anyone can capitalize on to succeed. Hence, Database Management Systems (DBMS) such as Oracle hang on their log files to be able to get back data if it gets corrupted or lost. As for businesses, they always have a history of data (in their data warehouse) to keep track of their evolution and build knowledge from it. It is also worth mentioning that the Facebook owner purchased WhatsApp and spent plenty of money on it even though it was not profitable, because he will gain something that is even more beneficial: data.

As for the SW, it is a main focus of Web 3.0 for many good reasons. Not only does it offer multiple Ontology Web Language (OWL) profiles that are centred on triples (subject, predicate, object), adding semantics to data. But it also allows us to claim data using SPARQL Protocol and Resource Description Framework (RDF) Query Language. Furthermore, it provides Named Entity Recognition (NER), e.g., France and Barack Obama. Moreover, it provides a wide range of axioms that can be tailored to our needs. For instance, there are: (i) fuzzy ontologies that consider the degree of membership of an instance to a class, and (ii) contextual ontologies that consider the degree of trust in a source when there are many. Additionally, OWL2 Description Logics (OWL2 DL) adds the ability to chain multiple properties thanks to (iii) property chain axioms. Also, one can use (vi) Semantic Web Rule Language (SWRL) to make decisions based on multiple conditions. In addition to that, SW requires the use of (v) reasoners to make inferences and check consistency. Last but not least, it allows the (vi) organization of ontologies into modules, which makes them easy to manage and reuse. Thus, offering Linked data to benefit from in diverse domains and for different purposes.

In recent years, ontologies have become a fundamental component of modern systems, aiding data organization, knowledge representation, and semantic interoperability across various domains. Worldwide, sectors have integrated ontologies to enhance system capabilities and improve decision-making. For example, the MITRE ATT&CK framework uses ontologies to categorize and analyse cyber threats, allowing organizations to better understand and mitigate potential attacks. Similarly, AI systems, such as Google's Knowledge Graph, leverage ontologies to create structured knowledge representations, improving search relevance and machine understanding.

This survey investigates how ontologies are incorporated into four different domains of information systems: AI, BI, cybersecurity, and the Metaverse. The choice to focus on these sectors stems from their pivotal role in the digital economy, where the complexity of data and the need for enhanced decision-making are rapidly growing. In cybersecurity, the increasing volume and sophistication of cyber threats demand structured knowledge frameworks for threat analysis. AI applications require ontologies to ensure semantic clarity and interpretability in decision-making models. Businesses rely heavily on data integration and insights provided by BI systems, which are enhanced by ontologies. Lastly, the Metaverse, an emerging sector, necessitates interoperable virtual environments where ontologies can support digital asset management and rule-setting. This survey also determines which domains are more often introducing ontologies in their applications. Additionally, it examines how ontologies can enhance systems' performance and functionality despite challenges. Our survey shows the potential of ontologies to enhance modern systems by thoroughly examining various domains, addressing challenges, and proposing solutions. This study remains analytical, with no practical implementation. Further research is needed to validate and apply our findings.

## BACKGROUND

### Artificial Intelligence

Concerning AI, it is undeniably a huge success. Nevertheless, it has some liabilities. For example, it fails at predicting that « safe » is irrelevant to the query « dangerous cars », since « dangerous » and « safe » co-occur. Not only this, but AI systems require at times both an elevated number of iterations to start making correct predictions and a large scale of data for training.

One example of AI-powered systems is Chatbots. They are dialog systems (1,2). In order to answer a user's query: first, Natural Language Processing (NLP) is used to process it and extract keywords from it. Second, the keywords are matched to a knowledge base like Wikipedia, Frequently Asked Questions (FAQ), and manuals or a knowledge graph with nodes and edges. The edges may or may not be labelled. These Chatbots often assist disabled people by reading text for example, thereby enhancing daily activities and promoting inclusivity for people with disabilities. The drawbacks to building a chatbot are the difficulty of classifying the query especially when the utterance is long. Consequently, it is easy to be misled by the term frequency or the lack of certain terms that would best describe the user's intent. Also, it is hard to retrieve the right answer, and provide related information.

Another example of AI technologies is IBM Watson that uses Medical like Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT). The terminology consists of over 325,000 clinical concepts and is used in analysing medical records and suggesting treatment options. But such terminologies although organized in hierarchical way; they only define basic relationships, e.g., "part of", "is a". Also, they cannot be used for automated reasoning as they do not support reasoning engines like Hermit and Pallet.

AI systems that adhere to ethical standards and protect individual privacy are of need. Hence, Radanliev et al. (2024) explore innovative algorithmic techniques such as homomorphic encryption, which allows computations on encrypted data. They also explore federated learning that protects individuals' data privacy through training the model locally and sharing only the results with a central server. Last but not least, they dive into differential privacy that consists of adding noise to data so as to make sure that individual data points cannot be identified. These methods limit breaches and misuse of AI systems, therefore allowing responsible AI deployment.

**Cybersecurity**

The Common Vulnerabilities and Exposures (CVE) identified 16,555 vulnerabilities in 2018. This was the highest number in the past 10 years. Later on, the increase in cyber-attacks during the pandemic and its remote working aftermath (Taylor, 2022) increased security attacks. This number kept rising every year and attained 24,000 in 2023, according to the National Vulnerability Database.

Multiple sources provide vulnerabilities to sensitize users; they come through when it comes to assisting them in vulnerability management. But, sources that gather and yield from all available and trustworthy sources are scarce (R. Syed, 2020). According to them, available ontologies in the domain are not enough for vetting vulnerabilities properly; like the Common Vulnerability Scoring System (CVSS) suggests.

One attempt to lessen security issues was made by Gao et al. (2023), who emphasize on swarms ability to solve cybersecurity.

**Business Intelligence**

Decision support Systems (DSS) aim at offering « the right information at the right time, with the right format (Turban et al., 2011). These systems can be classified into many categories, including data-driven DSS: data warehouse/BI (DW/BI) systems (Power, 2009) that allow decision making in companies through: First, the integration of data from various sources into a data lake (DL). These data include: (i) operational systems' data i.e., data from the information system (IS) or an Entreprise Resource Planification (ERP). They also include (ii) Client Relationship Management (CRM) data which are the company's strategy for client retention. The latter involves storing clients' phone numbers as well as their complaints and targeted or personalized marketing offers by sending e-mails to clients based on their age for example. Furthermore, they include (iii) external data like e-reputation data. Second, Extract, Transform, Load (ETL) these data into the DW. The latter could be split into many subject-oriented datasets called data marts. Third, the use of Online Analytical Processing (OLAP) tools to get reports and dash boards for decision making. These tools store data in cubes for dimensional modelling. They distinguish between quantitative facts (e.g., salesperson's quarterly target, product category's quarterly target, and regional quarterly target), and contextual dimensions (e.g., salesperson, product's category, region, time, and their hierarchies). Consequently, one can roll down to analyse a salesperson's quarterly target in a specific month or even a day. Alternatively, roll up to find that of a certain year. Also, users may drill up/down, i.e., move upwards or downwards in the overall hierarchy of the cube and not only for a specific dimension but for all of them. Furthermore, one gets to aggregate many dimensions to check the quarterly target of a salesperson for a product in a certain month and a specific city. Users also get to slice for a two-dimensional view, e.g., sales per product and month. Alternatively, dice to extract a sub-cube that includes multiple dimensions. Hence, multidimensional databases (Kimball & Ross, 2013) are powerful compared to relational databases.

### Metaverse

Augmented reality (AR) is valuable given that it yields a seamless integration of virtual environment in the real world objects (Marques et al., 2021). One main use of the Metaverse is allowing remote collaboration, which can be challenging since it involves vetting multiple aspects like team dynamics, task management, and communication, to name a few (Marques et al., 2021). Furthermore, the huge volume of interlinked data, especially when dealing with Mobile AR (MAR), makes appropriate filtering even more crucial to not overwhelm users using small screens. Yet, give them the ability to explore further links without having to switch to a different application. However, the heterogeneous nature of data in the Metaverse is yet another challenge.

One example of a virtual world is Decentraland. It is designed with its own set of data structures and protocols. This can hinder assets like skins and avatars from being transferred between different virtual environments or applications. In fact, the restrictive nature of such applications prevents the fluidity and scalability of the Metaverse, hindering the potential for a truly unified and open ecosystem.

As of now, many existing platforms lack the necessary accommodations for individuals with disabilities. To ensure that users with disabilities have equal access to virtual spaces, Radanliev et al. (2024) suggest a framework that enhances inclusivity for disabled people. For instance, disabled people can be allowed to customize their avatars to reflect their condition (e.g., adding a wheelchair). Another possibility for inclusivity would be creating interfaces that can accommodate people with different disabilities using assistive technologies like screen readers, voice commands, or haptic feedback devices for those with visual, auditory, or physical impairments.

## METHODOLOGY

To identify relevant literature for this survey, a search method (Figure 1), which is commonly used in surveys within the scientific domain, was conducted. The method includes finding the right search strings (Table 1) to answer the following research questions (RQ):

RQ1: How and which domains are incorporating ontologies the most?

RQ2: How can the use of ontologies be optimized within each domain for enhanced results?

RQ3: How do the complexity and characteristics of specific fields impact the suitability and success of their integration?

As for the search engine, we used Google Scholar and filtered results to only get papers that were first and for most published between 2014 and 2024 and that contained one of the search strings from Table 1 along with « ontology », « Linked data », « Knowledge graph », « Semantic Web » in the title using the inclusion criterion « allintitle : », e.g., « allintitle : ontology natural language processing ».

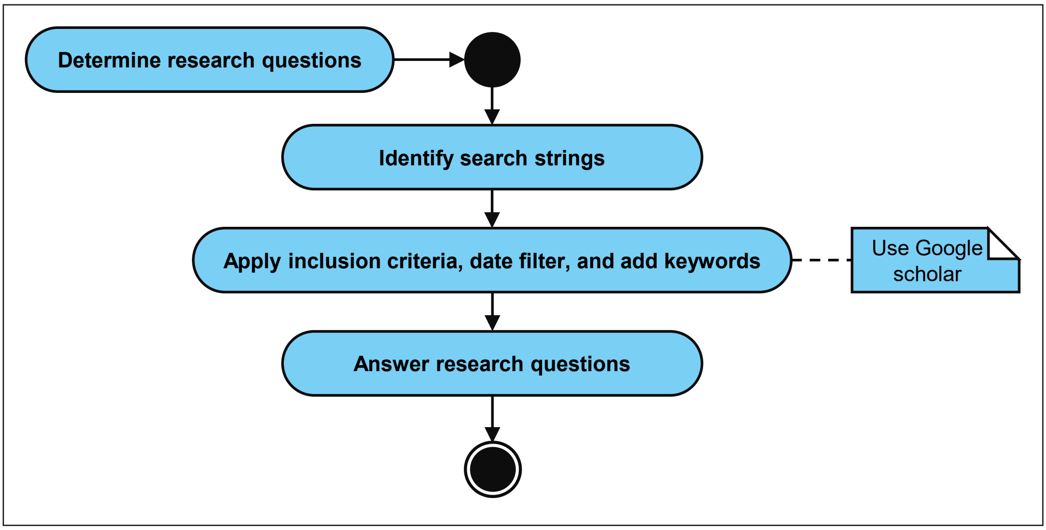


*Figure 1*. Methodology steps

Table 1
*A comparison between the use of ontologies in different domains*

| Domain | Reference | Method | Purpose |
|---|---|---|---|
| AI | Ali et al. (2017) | Fuzzy ontologies and SWIRL rules | Sentiment analysis and decision making |
| | Nguyen et al. (2021) | NER | Classification of user entry in Chatbots |
| Cybersecurity | De Rosa et al. (2022) | Developing Ontologies | Circumventing attacks with an ontological representation instead of a syntactic one |
| | Mugwagwa et al. (2023) | Swarm ontologies | Identifying and mitigating threats |
| | K. Liu et al. (2022) Mitra et al. (2021); R. Syed (2020); Rastogi et al. (2021) | Developing ontologies | Gathering and sharing cybersecurity issues and solutions |

Table 1 *(continue)*

| Domain | Reference | Method | Purpose |
|---|---|---|---|
| | K. Liu et al. (2022) | Temporal-event ontologies | Representing dynamic knowledge through |
| BI | Antunes et al. (2022) | NER | Avoiding semantic mismatches in ETL |
| | Amaral and Guizzardi (2019); Moreira et al. (2015) | Representing multidimensional models in the form of ontologies by adding OWL DL's constraints in the DW | Better expressiveness |
| | Prat, Akoka, et al. (2012) | Uncovering relations through reasoner's inference | Better semantic expressiveness |
| | Prat, Akoka, et al. (2012); Prat, Megdiche, et al. (2012) | A domain ontology | Validating data's consistency within the DW |
| | Kurze et al. (2010) | Defining core concepts through an ontology | Insurance of interoperability between DW systems |
| Metaverse | Marques et al. (2021) | Developing an ontology | Facilitating remote collaboration |
| | Vlachos et al. (2024) | Combining existing cultural heritage ontologies | Ensuring interoperability |

*Note*. AI = Artificial intelligence; BI: Business Intelligence; SWIRL = Semantic Web Inference Rule Language; NER = Named Entity Recognition; ETL = Extract, Transform, Load; OWL = Ontology Web Language; DL = Data lake; DW = Data warehouse

## RESULTS

### Artificial Intelligence

LOD could serve as training data. It was believed that their limits could be tackled by processing the training data to add and/or drop data based on LOD.

Also, by formalizing knowledge and relationships between concepts, ontologies enhanced AI systems' contextual understanding and ethical decision-making processes. In other words, ontologies could assist in embedding ethical standards within automated decision-making frameworks.

#### *Ontologies in Sentiment Analysis*

Ali et al. (2017) emphasized on the paramount importance of fuzzy ontologies as opposed to crisp ones in enhancing Intelligent Transportation Systems (ITS). They used tweets bigrams and trigrams' features to determine the degree of opinions' polarity for travellers. Also, they used SWRL rules to determine the reasons behind road congestion, for example. The authors' suggested system provided a 23% improvement in precision compared to using a classic ontology system. This leads to better decision-making for transport offices.

Nguyen et al. (2021) looked for the intent in the utterance by performing NER and classified queries into greeting, concepts, out of scope, comparison, and related knowledge. Consequently, classification issues are solved. Furthermore, they relied on relations between the provided answer and the meaning of the query to yield further knowledge. The authors' chatbot achieved an accuracy of 82% when tested on six types of queries.

**Cybersecurity**

De Rosa et al. (2022) suggested an ontology-based tool for tackling the rise in the number of attacks and presented knowledge gathered from external security sources. Their end goal was to opt for a semantic representation instead of a syntactic one.

Mugwagwa et al. (2023) used swarm ontology to sort out cybersecurity issues, given that the collective capabilities of swarms were higher than those of individual models in threat detection and bypass. They also developed a simulator to assess the role of ontologies in threat identification and mitigation.

K. Liu et al. (2022) outline the importance of knowledge graphs in cybersecurity. They gave vent to their possible applications given: (i) the asymmetric relationship between offence and defence in the security domain. Also, (ii) there has been a span in cyberspace to accommodate more fields, ranging from health to aviation and many more. Additionally, (iii) there is a shortage of cybersecurity experts. And (vi) the aggregation difficulty of heterogeneous data from open-source libraries and datasets into one model could be tackled by using, for instance, the Unified Cybersecurity Ontology (UCO). This ontology aggregates data from multiple cybersecurity standards and systems (Iannacone et al., 2015; Z. Syed et al., 2016). It therefore facilitates information sharing and exchange. However, ontologies were most effective when they were aimed towards a specific scenario. In other words, when they were application oriented, e.g., intrusion detection, malware categorization, vulnerability analysis, and threat actor analysis (Hooi et al., 2019; Pinkston et al., 2003; Sanagavarapu et al., 2021) (i.e., analysing his/her tools and level of expertise) rather than domain oriented (i.e., general). This specificity in ontologies was yet another issue since they might vary based on the field's specificities (e.g., health cybersecurity). Moreover, according to Sanagavarapu et al. (2021), ontologies should have been automatically enriched as this was a rapidly evolving domain. For instance, vulnerability management and prediction were highlighted in R. Syed (2020) as Syed developed a vulnerability ontology for cyber intelligence alert systems. Rastogi et al. (2021) established a malware knowledge graph for predicting malware attributes and sorting potential vulnerabilities. Their model achieved 80.4 for the hits@10 metric, which predicted the top 10 options for an information class. K. Liu et al. (2022) shed light on the need for representing dynamic knowledge through temporal and event subordination relations in cybersecurity knowledge graphs. Additionally, Mitra et al. (2021) suggested an ontology to vet out fake cybersecurity intelligence, i.e., false or misleading information presented as threats or attacks by either malicious actors or unintentional errors and misinterpretations.

Social media was another source that could have been of paramount importance in vulnerability management. Truth be told, every social media per se yielded a mere tad number of ways to counter and hold out against attacks. But their combination had the potential to pay off. Hence, R. Syed (2020) took up benefitting from social media in an ontology along with other information from multiple other ontologies.

### Business Intelligence

Given that data in a DL was heterogeneous, the same entity could have been in different formats and different presentations. So, in order to have a single version of truth (Antunes et al., 2022) and avoid semantic mismatches, NER could be used in the ETL phase to automate the process. Next, the determined entities along with their extracted properties could have been mapped to other ontologies to pinpoint their corresponding formalized entities' names and properties' names. This alignment of ontologies to have a common terminology (T-box) for knowledge building also benefited on the one hand, the interoperability between stakeholders or between DW/BI systems and other DSS (Kimball & Ross, 2013). On the other hand, it benefited the enrichment of data with semantic similarity (e.g., cat and kitten) and semantic relatedness (e.g., cat and dog). This enrichment offers new knowledge for decision makers. Additionally, at the DW level, the representation of multidimensional models in the form of ontologies by adding OWL DL's constraints allowed a higher level of expressiveness (Amaral & Guizzardi, 2019; Moreira et al., 2015). This semantic expressiveness would, in turn, help uncover relations through the reasoner's inference (Prat, Akoka, et al., 2012). For this purpose, an extra mapping layer was added between the ontological constraints within DW and the formal domain ontology, with the reasoner being used in the latter. As a result, users would get insight into other facts and hierarchical dimensions to use in semantic OLAP cubes. Alternatively, a domain ontology could be used to validate the data within the DW by checking its consistency and enforcing constraints (Prat, Akoka, et al., 2012; Prat, Megdiche, et al., 2012). On a similar note, users could take advantage of ontologies to interpret OLAP results and collaborate/share knowledge with other stakeholders. Last but not least, Kurze et al. (2010) guided the insurance of interoperability between DW systems by defining core concepts for data warehousing in an ontology.

### Metaverse

According to Azuma et al. (2001), AR was not constrained to a particular type or a particular sense. Hence, it could be applied with all human senses (McGee, 1999; van Krevelen & Poelman, 2010) and could even substitute people's lacking senses (aka sensory substitution) (Carmigniani et al., 2011).

One more subpart of AR was diminished or mediated reality, which consisted on the removal of physical objects from the perceived environment (Azuma et al., 2001).

Marques et al. (2021) yielded an ontology to facilitate remote collaboration through structuring and understanding the scenarios of such collaborations.

Assuming that data without ontological structure and constraints might lack depth and consistency, Vlachos et al. (2024) combined existing cultural heritage ontologies for better AR in that domain, ensuring both interoperability and reusability across different projects and scenarios.

As for ethics and accessibility, particularly for disabled people, developing ontologies that were specific to the Metaverse, like MetaOntology, aimed at standardizing associated technologies and infrastructure. This formalization enhanced interoperability and accessibility, enabling users with disabilities to navigate and interact more effectively and within virtual environments.

From Figure 2, it could be seen that AI was the domain where ontologies were the most used in recent years and that BI was the domain that used them the least.

Based on Table 2, ML was up there in terms of taking advantage of ontologies to improve related applications. Moreover, deep learning and sentiment analysis were other sub-domains that benefited from ontologies the most.

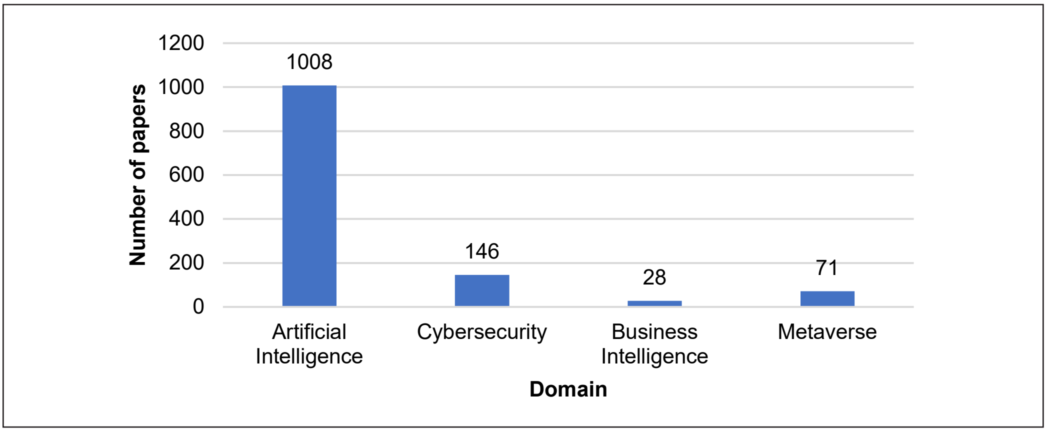


*Figure 2*. Number of English papers published between 2014 and 2024 and using in the title the keywords: Ontology/Linked Data/Semantic Web/Knowledge Graph, along with domain keywords

Table 2
*Number of papers by sub domain using in the title: Ontology/Linked Data/Semantic Web/Knowledge Graph*

| Domain | Search string | Number of papers |
|---|---|---|
| Artificial intelligence | Artificial intelligence | 94 |
| | Deep learning | 282 |
| | Machine learning | 484 |

Table 2 *(continue)*

| Domain | Search string | Number of papers |
|---|---|---|
| | Opinion mining | 27 |
| | Sentiment analysis | 152 |
| | Natural language processing | 116 |
| | Chatbot | 46 |
| Cybersecurity | Cybersecurity | 83 |
| | Cyber security | 63 |
| Business Intelligence | Business Intelligence | 26 |
| Metaverse | Augmented reality | 39 |
| | Virtual reality | 32 |

## DISCUSSION

### Comparison and Analysis

When it comes to AI, reasoners' inference would add more semantics to knowledge graphs. Thus, it benefits chatbot's capacity to classify user's input. Concerting healthcare AI, it is believed that by transforming SNOMED CT into an ontology with more advanced constraints like cardinality constraints (e.g., a patient has at most two diagnoses), and inverse properties (e.g., "has parent" and "has child"). Additionally, reasoners can infer new facts using defined axioms (e.g., disjointness and equivalence). For instance, they can find new connections. They can also detect logical inconsistencies such as contradictory definitions, redundant terms, and missing relationships. Furthermore, ontologies would allow advanced querying and improve interoperability by linking different datasets and allowing for more seamless integration across systems using Uniform Resource Identifiers (URIs).

As for Cybersecurity, ontologies are assumably great categorizers. This advantage leads to bettered reasoning (Abburu, 2012) and consideration of diverse implicit as well as explicit relations (DeStefano et al., 2016) for semantically boosted knowledge bases.

Assumably, SW remains by far the best choice if one has structured ontological data (like BI data) and is looking for semantic consistency across different areas of life (which is the case for BI for example). In other words, it is the best choice when the aim is to have a common vocabulary for a common understanding. This would also improve data integration through harmonizing data coming from different sources and ensuring its quality. One example of existing ontologies is Financial Industry Business Ontology (FIBO).

And concerning the Metaverse, matching ontologies can solve data heterogeneity. And URIs can allow access to real-time data despite their huge amount. Furthermore, a shared ontology can allow cross-world usage. For example, a wearable skin bought in Decentraland could be used by our created avatar in another Metaverse platform if both platforms agreed upon the same ontology definitions. Consequently, instead of having to buy a new jacket or skin for each world, one can use the same wearables on different Metaverse platforms, making virtual items more valuable and interoperable. This, in turn, leads to easier content creation as creators can make wearables once and have them work across multiple virtual worlds, increasing their audience and revenue opportunities. Moreover, the ontology-based system could also suggest related or complementary assets based on semantic relationships defined in the ontology. For instance, if the creator selects a "cyberpunk jacket", the system might recommend other items such as "glowing gloves" or "high-tech boots". This would be time-saving and consistency-improving. But, this domain is new compared to the other ones, in this paper, which is why ontologies are underused in it despite their huge potential.

Our findings on the usefulness of ontologies in enhancing semantic interoperability, data integration, and decision-making are not limited solely to cybersecurity, AI, BI, and the Metaverse. In fact, similar principles and methodologies can be applied across a wide array of sectors (Gruber, 1993):

- Healthcare: Ontologies are extensively used in healthcare to standardize medical terminologies, integrate patient records, and support clinical decision-making. For example, the Unified Medical Language System (UMLS) organizes diverse biomedical vocabularies to facilitate data sharing and interoperability among Electronic Health Records (EHR) (Bodenreider, 2004; Noy & McGuinness, 2001). Also, the work by Kouremenou et al. (2024) presents a data modelling process aimed at achieving interoperability. The authors emphasize the importance of semantic and syntactic interoperability and address challenges such as compatibility issues and the need for global standards. Their approach contributes to resolving data management and exchange problems among healthcare entities, enhancing data accessibility and accuracy.
- Manufacturing and supply chain management: In manufacturing, ontologies help in structuring data from Internet of Things (IoT) devices and sensors, which can lead to more effective predictive maintenance, process optimization, and smart supply chain management. This approach enables better integration of heterogeneous data sources across production systems (Gómez-Pérez et al., 2004).
- Legal informatics: The legal field benefits from ontologies by systematizing complex legal information. They enable automated reasoning over case laws, regulations, and contractual documents, thereby supporting legal research and compliance monitoring.
- Education: Ontologies can support the development of adaptive learning systems by organizing educational content and tailoring it to individual learner profiles. This facilitates personalized learning and improved outcomes by mapping curricular standards to instructional materials.

- E-commerce: In e-commerce, ontologies are applied to enhance product categorization, semantic search, and recommendation engines. They enable more precise matching between customer queries and product offerings, ultimately leading to improved user experience and sales.
- Digital twins (DT): The study by Karabulut et al. (2024) offers a comprehensive review of how ontologies are utilized within DT, highlighting their role in knowledge representation, interoperability, and automated reasoning.

Incorporating ontological data representations aligns with the European Union's (EU) objectives of promoting digital transformation and establishing a Digital Single Market. It supports the EU's strategy for data-driven innovation and the development of interoperable digital public services. Moreover, the insights from these studies contribute to global discussions on data standardization and interoperability, influencing policies and practices beyond the EU.

## Challenges and Open Issues

Although users opt for dynamic ontologies that include spatial, temporal, and event data to capture various dimensions of cyber threats and vulnerabilities; there is an absence of evaluation standards for dynamic ontologies (K. Liu et al., 2022).

Regarding BI, most related work studies use ontologies to limit heterogeneity and achieve interoperability while designing multidimensional models. But there is still a lack of papers that consider ontologies for the enrichment of DW with new interlinked data (Antunes et al., 2022). This enrichment would not only give a new ground on which users can build their decisions, but would also sort out semantic ambiguity, especially in complex domains (Bargui et al., 2011) like the healthcare domain. Moreover, to the best of our knowledge, even fewer studies were conducted on defining the architecture (T-box) of the DW to add instances (A-box) depending on the T-box. This is because full transformation of the DW would require further changes to the workflow and the tools used. For the time being, ontologies are rather used either to describe the DW's architecture (Szwed et al., 2015) or to support its design (X. Liu & Iftikhar, 2013). Concerning fuzzy and contextual ontologies, they are almost non-existent in the literature. Their use would drastically benefit the determination of valuable external sources from non-valuable ones right from the jump and would bring more accuracy to opinions' polarity on social media. Another possibility would be using ontologies, early on, for strategy modelling and metadata for guidance during information retrieval (IR) to automate the retrieval process and avoid retrieving irrelevant data for decision making.

As for the Metaverse, although it obtained a layered ontology from combining multiple ones (Vlachos et al., 2024). It can be less time-consuming, but it is hard to manage. Consequently, we will need to figure out ways to facilitate complex data management.

In general, ontologies can be challenging to incorporate for several reasons. First, they can be hard to create (Kiourtis et al., 2019). In fact, in order to build and maintain ontologies; a significant amount of manual intervention and expertise is needed to learn ontologies (Khadir et al., 2021) either from text or from relational databases. This process is time-consuming and impractical for resource-constrained domains. For more emphasize, ontologies are easier to incorporate when a domain is stable, as they can be hard to update if classification keeps changing. Second, ontologies have scalability and flexibility issues. This is specifically challenging when real-time data changes frequently or when we want to incorporate new concepts, due to their structured and rigid nature. To sort this out, tagging systems are often preferred (Höning, n.d.; Noy & McGuinness, 2001). Third, ontologies are hard to integrate because of the lack of standardization in some domains, as ontologies require a broad consensus across stakeholders to be effective. Fourth, incorporating ontologies in modern systems comes with the challenge of choosing adequate mapping techniques (Mavrogiorgou et al., 2020) to allow interoperability.

Table 3 indicates the domains vetting in which ontologies are easy/moderately easy/hard to incorporate based on this survey.

Table 3
*Classification of ontologies in terms of their ease of incorporation per domain*

| **Ontology's ease of use** | **Domain** | **Domain's characteristics** |
|---|---|---|
| Easy | BI | • Stability<br>• Structured data<br>• Consensus through common frameworks, e.g., OLAP, DW |
| Moderately easy | AI | • Lack of stability<br>• Unstructured nature of some subfields like deep learning<br>• Potential for flexibility; through combining ontologies with more flexible models like probabilistic reasoning or using probabilistic ontologies. These dynamic ontologies integrate ontological structures with probabilistic reasoning. This leads to better decision-making in scenarios with incomplete or ambiguous information |
| Hard | Cybersecurity | • Constantly evolving threats make ontologies hard to keep up-to-date<br>• A lot of unstructured data, e.g., threat intelligence reports and network logs. These types of data can be hard to classify and manage within an ontology<br>• Lack of consensus as some systems may priorities certain threats over others |
| | Metaverse | • Unstructured data<br>• Dynamic virtual environment<br>• Lack of consensus, as there is not much agreement on classification, given that many developers contribute to it<br>• Scalability, as the domain's rapid growth and nature makes scaling ontologies hard |

*Note*. BI = Business Intelligence; OLAP = Online Analytical Processing; DW = Data warehouse; AI = Artificial intelligence

Constructing ontologies for sectors such as cybersecurity, AI, BI, and the Metaverse presents several challenges:

- Data heterogeneity: Integrating diverse data formats and sources necessitates the development of robust mapping techniques to ensure consistency and interoperability.
- Dynamic environments: The rapidly evolving nature of these fields requires ontologies to be adaptable, accommodating continuous technological advancements, and emerging threats.
- Privacy and security concerns: Particularly in the Metaverse, safeguarding user data and ensuring secure interactions pose significant challenges. Biometric methods, while unique, are susceptible to misuse, highlighting the need for secure data handling practices.
- Decentralization issues: The lack of centralized authority in decentralized systems like the Metaverse complicates the establishment of uniform security standards and regulatory frameworks.

**Future Directions**

SW can be used interchangeably with AI to gain insight from the reasoner, for better anomaly or fraud detection, and for tweets' sentiments analysis in a fuzzy or a contextual way, to name a few use cases. Alternatively, it can be used along with AI to enrich data and make it more discoverable.

Ontologies can facilitate the distinction between threats, assess risks, scale vulnerabilities, and provide a step-by-step assistance tailored to every possible attack scenario or so. Nevertheless, it is believed that even if OWL DL is highly used as opposed to OWL Full. Critical domains like cybersecurity need the higher expressiveness of the full version of ontologies. Hence, more studies need to be done in that regard.

As for businesses, they may benefit from LOD in decision-making. Companies rely on DW and BI systems such as OLAP to make decisions. Using LOD prior to data marts would potentially add a lot of semantics that can lead to more accurate choices in the future. Especially, data warehouses consist, among other things, of operational systems' data and external data. The latter may help drastically when it comes to smartly assessing e-reputation, analysing social media, and studying the target population if incorporated with LOD. Dimensions bring context to the facts. So, using ontologies would bring more context that stakeholders were not even aware of. Users tend to create a bridge (i.e., an interoperability layer) between DW and the formal domain ontology, instead of fully transforming it into an ontology. Thus, we will not waste our time. It can also be used in DBMS and DL to avoid wasting space by creating only links to other ontologies. But before all that, it is worth digging deeper to investigate why BI is the least domain to benefit from ontologies, out of the four domains in this survey, despite their ease of incorporation in it.

Incorporating ontologies in Metaverse not only facilitates a common representation and classification of diverse data types but also bridges the gap between virtual entities and the physical world. Indeed, the Semantic Web serves as a valuable tool for advanced data exploitation, reasoning, and inference (Lampropoulos et al., 2020). This is specifically the case with emerging Metaverse technologies for context awareness through enriching our understanding of the difference between physical and virtual objects. However, do we need mere ontologies ensuring standardisation? Or are complex ontologies mandatory for adaptability, which is a key element in domains like the Metaverse?

Furthermore, ontologies can be used across various domains to boost expressiveness, reasoning, and create common knowledge. In fact, SW provides a wide range of features to choose from, and SPARQL to interrogate ontologies. Therefore, it should be used more often in various systems to boost their performance.

## CONCLUSION

The Semantic Web leverages the capabilities of knowledge graphs by allowing them to be reusable and shareable using URIs. This enables internet-connected devices to understand, just like humans do, for ideal human-computer interaction and decision-making. It would also increase revenue and audience, amongst other things, across platforms in the Metaverse, for example. However, ontologies are constrained in systems and not fully taken advantage of as this would take either a bigger change in the architectures, as in BI. Alternatively, it would take their use early on in the process and not just as a complement. Additionally, critical domains like cybersecurity need thorough vetting and are very low in false tolerance. Hence, OWL Full would be better suited, thanks to its high expressiveness, despite it being computationally demanding. Furthermore, using dynamic ontologies more often can enhance systems given their diversity and the ability of each to encounter a certain issue. However, evaluation benchmarks for dynamic ontologies are lacking as of now. Finally, ontologies range from easy to hard to integrate depending on the domain in question. But again, dynamic ontologies have the potential to make them less difficult to incorporate as in AI.

It is believed that stakeholders such as cybersecurity professionals, AI developers, BI analysts, and Metaverse platform designers can benefit from this research by gaining insights into the integration of Semantic Web technologies to enhance systems' interoperability and security.

In the future, the focus will be on improving the MetaOntology's accessibility, inclusiveness, interoperability, and respect for user privacy:

- Months 1-2: Assessment and planning
    - Conduct a thorough review of existing the Metaverse platforms to identify accessibility shortcomings.

    - Evaluate current data collection practices within the Metaverse to identify potential privacy vulnerabilities.
- Months 3-5: Development of standards and guidelines
    - Create comprehensive standards that define accessibility features, such as customizable user interfaces, screen reader support, and alternative input methods.
    - Formulate robust privacy policies that outline data collection limitations, user consent mechanisms, and data protection measures.
    - Design a framework that facilitates seamless interaction between various Metaverse platforms, ensuring consistent user experiences.
- Months 6-8: Implementation and testing

## ACKNOWLEDGEMENTS

We would like to thank the Laboratory of Intelligent Systems and Applications at the Faculty of Sciences and Technologies in Sidi Mohamed Ben Abdellah University for providing the technical support for this research.

## REFERENCES

Abburu, S. (2012). A survey on ontology reasoners and comparison. *International Journal of Computer Applications*, *57*(17), 33-39.

Ali, F., Kwak, D., Khan, P., Islam, S. M. R., Kim, K. H., & Kwak, K. S. (2017). Fuzzy ontology-based sentiment analysis of transportation and city feature reviews for safe traveling. *Transportation Research Part C: Emerging Technologies*, *77*, 33-48. https://doi.org/10.1016/j.trc.2017.01.014

Amaral, G., & Guizzardi, G. (2019). On the application of ontological patterns for conceptual modeling in multidimensional models. In T. Welzer, J. Eder, V. Podgorelec, & A. Kamišalić Latifić (Éds.), *Advances in databases and information systems* (Vol. 11695, pp. 215-231). Springer. https://doi.org/10.1007/978-3-030-28730-6_14

Antunes, A. L., Cardoso, E., & Barateiro, J. (2022). Incorporation of ontologies in data warehouse/business intelligence systems — A systematic literature review. *International Journal of Information Management Data Insights*, *2*(2), 100131. https://doi.org/10.1016/j.jjimei.2022.100131

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, *21*(6), 34-47. https://doi.org/10.1109/38.963459

Bargui, F., Ben-Abdallah, H., & Feki, J. (2011). A decision-making ontology for analytical requirements elicitation. In E. Gelenbe, R. Lent, & G. Sakellari (Éds.), *Computer and information sciences II* (pp. 495-501). Springer. https://doi.org/10.1007/978-1-4471-2155-8_63

Bodenreider, O. (2004). The Unified Medical Language System (UMLS): Integrating biomedical terminology. *Nucleic Acids Research*, *32*(suppl_1), D267-D270. https://doi.org/10.1093/nar/gkh061

Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E., & Ivkovic, M. (2011). Augmented reality technologies, systems and applications. *Multimedia Tools and Applications*, *51*, 341-377. https://doi.org/10.1007/s11042-010-0660-6

De Rosa, F., Maunero, N., Nicoletti, L., Prinetto, P., & Trussoni, M. (2022). *Ontology for cybersecurity governance of ICT systems*. https://ceur-ws.org/Vol-3260/paper4.pdf

DeStefano, R. J., Tao, L., & Gai, K. (2016). Improving data governance in large organizations through ontology and linked data. In *IEEE 3rd International Conference on Cyber Security and Cloud Computing* (pp. 279-284). IEEE. https://doi.org/10.1109/CSCloud.2016.47

Gao, X., Xiao, G., Xie, K., Wang, W., Fu, Y., Chang, C., & Wang, Z. (2023). A framework of modeling and simulation based on swarm ontology for autonomous unmanned systems. *Applied Sciences*, *13*(16), 9297. https://doi.org/10.3390/app13169297

Gómez-Pérez, A., Fernández-López, M., & Corcho, O. (2004). *Ontological engineering: With examples from the areas of knowledge management, e-commerce and the Semantic Web* (1st ed.). Springer. https://doi.org/10.1007/b97353

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199-220.https://doi.org/10.1006/knac.1993.1008

Höning, N. (n.d.). *Clay Shirky - "Ontology is overrated": A review*. https://www.nicolashoening.de/?twocents&nr=11

Hooi, E. K. J., Zainal, A., Maarof, M. A., & Kassim, M. N. (2019). *TAGraph: Knowledge graph of threat actor*. In *International Conference on Cybersecurity* (pp. 76-80). IEEE. https://doi.org/10.1109/ICoCSec47621.2019.8970979

Iannacone, M., Bohn, S., Nakamura, G., Gerth, J., Huffer, K., Bridges, R., Ferragut, E., & Goodall, J. (2015). Developing an ontology for cyber security knowledge graphs. In *Proceedings of the 10th Annual Cyber and Information Security Research Conference* (pp. 1-4). Associations for Computing Machinery. https://doi.org/10.1145/2746266.2746278

Karabulut, E., Pileggi, S. F., Groth, P., & Degeler, V. (2024). Ontologies in digital twins: A systematic literature review. *Future Generation Computer Systems*, *153*, 442-456. https://doi.org/10.1016/j.future.2023.12.013

Khadir, A. C., Aliane, H., & Guessoum, A. (2021). Ontology learning: Grand tour and challenges. *Computer Science Review*, *39*, 100339. https://doi.org/10.1016/j.cosrev.2020.100339

Kimball, R., & Ross, M. (2013). *The data warehouse toolkit: The definitive guide to dimensional modeling* (3rd ed.). Wiley.

Kiourtis, A., Mavrogiorgou, A., & Kyriazis, D. (2019). Constructing healthcare ontologies of any data format. In *eTELEMED 2019: The Eleventh International Conference on eHealth, Telemedicine, and Social Medicine* (pp. 43-48). IARIA Press.

Kouremenou, E., Kiourtis, A., & Kyriazis, D. (2024). A data modeling process for achieving interoperability. In H. N. Costin, R. Magjarević, & G. G. Petroiu (Éds.), *Advances in digital health and medical bioengineering* (Vol. 109, pp. 711-719). Springer. https://doi.org/10.1007/978-3-031-62502-2_80

Kurze, C., Gluchowski, P., & Bohringer, M. (2010). Towards an ontology of multidimensional data structures for analytical purposes. In *43rd Hawaii International Conference on System Sciences* (pp. 1-10). IEEE. https://doi.org/10.1109/HICSS.2010.485

Lampropoulos, G., Keramopoulos, E., & Diamantaras, K. (2020). Enhancing the functionality of augmented reality using deep learning, semantic web and knowledge graphs: A review. *Visual Informatics*, *4*(1), 32-42. https://doi.org/10.1016/j.visinf.2020.01.001

Liu, K., Wang, F., Ding, Z., Liang, S., Yu, Z., & Zhou, Y. (2022). *A review of knowledge graph application scenarios in cyber security*. arXiv. http://arxiv.org/abs/2204.04769

Liu, X., & Iftikhar, N. (2013). Ontology-based big dimension modeling in data warehouse schema design. In W. Abramowicz (Ed.), *Business information system* (Vol. 157, pp. 75-87). Springer. https://doi.org/10.1007/978-3-642-38366-3_7

Marques, B., Silva, S., Dias, P., & Santos, B. S. (2021). An ontology for evaluation of remote collaboration using augmented reality. In *Proceedings of the 19th European Conference on Computer-Supported Cooperative Work* (pp. 1-8). European Society for Socially Embedded Technologies. https://doi.org/10.18420/ecscw2021_p04

Mavrogiorgou, A., Kiourtis, A., & Kyriazis, D. (2020). Identification of IoT medical devices APIs through ontology mapping techniques. In P. Inácio, A. Duarte, P. Fazendeiro, & N. Pombo (Eds.), *5th EAI International Conference on IoT Technologies for HealthCare* (pp. 39-54). Springer. https://doi.org/10.1007/978-3-030-30335-8_4

McGee, M. K. (1999). *Integral perception in augmented reality* [Unpublished Doctoral dissertation]. Virginia Polytechnic Institute and State University.

Mitra, S., Piplai, A., Mittal, S., & Joshi, A. (2021). Combating fake cyber threat intelligence using provenance in cybersecurity knowledge graphs. In *IEEE International Conference on Big Data* (pp. 3316-3323). IEEE. https://doi.org/10.1109/BigData52589.2021.9671867

Moreira, J., Cordeiro, K., Campos, M. L. M., & Borges, M. (2015). Hybrid multidimensional design for heterogeneous data supported by ontological analysis: An application case in the Brazilian electric system operation. In *Workshop Proceedings of the EDBT/ICDT 2015 Joint Conference* (pp. 72-77). CEUR Workshop Proceedings.

Mugwagwa, A., Chibaya, C., & Bhero, E. (2023). A survey of inspiring swarm intelligence models for the design of a swarm-based ontology for addressing the cyber security problem. *International Journal of Research in Business and Social Science (2147-4478)*, *12*(4), 483-494. https://doi.org/10.20525/ijrbs.v12i4.2473

Nguyen, H. D., Tran, T.-V., Pham, X.-T., Huynh, A. T., & Do, N. V. (2021). Ontology-based integration of knowledge base for building an intelligent searching chatbot. *Sensors and Materials*, *33*(9), 3101-3123. https://doi.org/10.18494/SAM.2021.3264

Noy, N. F., & McGuinness, D. L. (2001). *Ontology development 101: A guide to creating your first ontology*. Knowledge Systems Laboratory Stanford University.

Pinkston, J., Undercoffer, J., Joshi, A., & Finin, T. (2003). *A target-centric ontology for intrusion detection*. https://ebiquity.umbc.edu/_file_directory_/papers/626.pdf

Power, D. J. (2009). *Decision support basics*. Business Expert Press.

Prat, N., Akoka, J., & Comyn-Wattiau, I. (2012). Transforming multidimensional models into OWL-DL ontologies. In *Sixth International Conference on Research Challenges in Information Science* (pp. 1-12). IEEE. https://doi.org/10.1109/RCIS.2012.6240451

Prat, N., Megdiche, I., & Akoka, J. (2012). Multidimensional models meet the semantic web: Defining and reasoning on OWL-DL ontologies for OLAP. In *Proceedings of the Fifteenth International Workshop on Data Warehousing and OLAP* (pp. 17-24). Association for Computing Machinery. https://doi.org/10.1145/2390045.2390049

Radanliev, P., De Roure, D., Novitzky, P., & Sluganovic, I. (2024). Accessibility and inclusiveness of new information and communication technologies for disabled users and content creators in the Metaverse. *Disability and Rehabilitation: Assistive Technology*, *19*(5), 1849-1863. https://doi.org/10.1080/17483107.2023.2241882

Rastogi, N., Dutta, S., Christian, R., Gridley, J., Zaki, M., Gittens, A., & Aggarwal, C. (2021). *Predicting malware threat intelligence using KGs*. https://www.semanticscholar.org/paper/Predicting-malware-threat-intelligence-using-KGs-Rastogi-Dutta/0b4d8612f4d2cd69b6eade7f01f023fcca2e16cc

Sanagavarapu, L. M., Iyer, V., & Reddy, R. (2021). *A deep learning approach for ontology enrichment from unstructured text*. arXiv. https://doi.org/10.48550/arXiv.2112.08554

Syed, R. (2020). Cybersecurity vulnerability management: A conceptual ontology and cyber intelligence alert system. *Information and Management*, *57*(6), 103334. https://doi.org/10.1016/j.im.2020.103334

Syed, Z., Padia, A., Finin, T., Mathews, L., & Joshi, A. (2016). *UCO: A unified cybersecurity ontology*. In *The Workshops of the Thirtieth AAAI Conference on Artificial Intelligence* (pp. 195-202). Association for the Advancement of Artificial Intelligence.

Szwed, P., Komnata, W., & Dymek, D. (2015). DWARM: An ontology of data warehouse architecture reference model. In S. Kozielski, D. Mrozek, P. Kasprowski, B. Małysiak-Mrozek, & D. Kostrzewa (Éds.), *Beyond databases, architectures and structures* (Vol. 521, pp. 222-232). Springer. https://doi.org/10.1007/978-3-319-18422-7_20

Taylor, A. (2022, February 17). There's a huge surge in hackers holding data for ransom, and experts want everyone to take these steps. *Fortune*. https://fortune.com/2022/02/17/ransomware-attacks-surge-2021-report/

Turban, E., Sharda, R., & Delen, D. (2011). *Decision support and business intelligence systems* (9th ed.). Pearson Education India.

van Krevelen, D. W. F., & Poelman, R. (2010). A survey of augmented reality technologies, applications and limitations. *International Journal of Virtual Reality*, *9*(2), 1-20. https://doi.org/10.20870/IJVR.2010.9.2.2767

Vlachos, A., Perifanou, M., & Economides, A. A. (2024). A review of ontologies for augmented reality cultural heritage applications. *Journal of Cultural Heritage Management and Sustainable Development*, *14*(2), 160-174. https://doi.org/10.1108/JCHMSD-06-2021-0110