Enhancing Social Networks Services: An Ontology-Based Approach

Eman Alduweib
Department of Computer Science
University of Petra
Jordan
eman.alduweib@uop.edu.jo

Abdurahman Alahmadi
Department of Computer Science and Information, Taibah University
Saudi Arabia
aahmadi@taibahu.edu.sa

Waseem Alromema
Department of Computer Science and Information, Taibah University
Saudi Arabia
wromema@taibahu.edu.sa

Abstract: *As for the Internet is grown this much and the massive data of social networks accounts have been accumulated about users' behavior on digital platforms, it is found that this data require to be derived on efficient way to increase the personalization for improving the internet user experience. This paper introduces an ontology-based architecture that extracts user's data from multiple sources such as websites, applications and social networks accounts. The data is then preprocessed and used for developing the domain ontology, which facilitates the structure of user concepts and allows for efficient data integration and intelligent reasoning. The proposed architecture makes the domain ontology applicable in personalized digital marketing, digital cloning, AI (Artificial Intelligence) based virtual assistants and other user-driven applications. In addition, it takes privacy and ethical concerns into account, where user consent mechanisms, data anonymization techniques and compliance with regulations like the General Data Protection Regulation (GDPR) can be utilized. By incorporating such considerations, the architecture ensures responsible use of the collected data for maximum contribution to the digital services, and prevents any form of exploitation against internet users.*

Keywords: *Social networks, internet users, user behavior, ontology, digital clones, digital immortality.*

Received May 30, 2025; accepted October 1, 2025
https://doi.org/10.34028/iajit/23/2/8

1. Introduction

In the digital world, data is the pivot of innovation including personal user experience, prediction, and analytics, and various optimization aspects of service. As the data in the form of multiple websites, web applications, social networks and others are growing exponentially, more sophisticated frameworks are needed to integrate different kinds of data. There are billions of online interactions created every single day, and bring about the need for intelligent systems to escape with various data streams, aggregation and processing, and eventually derive meaning knowledge from them [12, 25]. However, it is commonly recognized that traditional data integration methods are heavily limited, including the problem of integrating heterogeneous data formats, preserving user privacy, and extracting useful information from massive unstructured datasets. Today, the rising dependency on digital services equips business and researchers to develop data driven solutions that eventually; raise concerns surrounding the ethical ways of data usage and privacy regulations such as General Data Protection Regulation (GDPR) whereby it restricts to collect, store and process data not within the boundaries [11] and furthermore, businesses and researchers still going through to balance between innovation and morals [20].

Consequently, the proposed ontology-based architecture is a new paradigm towards querying about internet users by systematically collecting and integrating the data generated from their online behavior. The proposed architecture in this case is based on semantic web and domain ontology, which allow machine readable parsing and user behavior analysis. Ontology developed using protege tool [18], which allows to organizing the user concepts, relationships and attributes in such a way as to provide richer information retrieval and automated reasoning. It helps in the semantic interoperability of user data among multiple platforms by the ontology engineering-based architecture [2, 23]. In big data, ontology is a way of doing things, instead of using a more conventional method, which is to rely on statistical models, a structured and scalable means to represent and infer the user behavior in a more explainable way is given, which empowers AI (Artificial Intelligence) systems to make personalized recommendations, predict user preferences, and enforce ethical data usage [13, 27].

Our contributions in the proposed architecture can be summarized as follows:

- Proposing different mechanisms for collecting/extracting data of the internet user from various sources.
- Building the domain ontology based on the collected/extracted data after being preprocessed.
- Analyzing data in order to identify patterns and extract insights to be used in decision-making or other

real time applications.

The paper is organized as follow, next section is devoted to reviewing related works, presenting the present methods and summarizing their restrictions. Section 3 presents the proposed architecture and its main parts, such as data collection, data preprocessing and applications. Section 4 presents the implementation and results discussion, and Section 5 concludes the paper with the summary of the findings and further research directions.

2. Related Works

This section presents a synthesis of prior research in social networks data integration, user behavior analysis, and ethical frameworks, classifying key gaps that this proposed ontology-based architecture aims to fill. To enable cross-domain mappings in social networks analysis, structured classes and rules are developed into multi-domain ontology, Andrades *et al*. [2]. However, even though it is effective in semantic reasoning, it inevitably limits its adoption in sensitive data scenarios as the absence of fine grained consent mechanisms prevent the model from being used in wide scale.

Discussion of digital immortality by Caon [5] focuses on posthumous profiling, not the technical design for ethical data aggregation. This philosophical approach lacks integrity like data constraints and dynamic consent mechanisms, which makes data abuse a real possibility. Furthermore, it extends these ideas by incorporating privacy preserving controls (e.g., time bound data access) and ontology based consistency checks, as well. Additionally, Yang *et al.* [27] proposed a machine learning framework that infers depression from social networks behaviors.

However, the mean shift is creative in that it is statistical rather than deterministic, and it does not generate a formal semantic representation, limiting the interpretability of results and their representation of the degrees of heterogeneity of the input features of data. Rather, our architecture integrates semantic reasoning and learning-based methods to augment the generalizability of the architecture and ability to be explained at the same time [1]. Bendjamaa and Taleb [4] have been developing an ontology covering the maximum amount of data possible, which named OntoDin, they selected for a method using protege plugins. Arafeh *et al*. [3] used the Ontology based recommender systems but restricted to the specific applications without their practical use cases across the multi domains. Whereas, our architecture applies integrated ontologies so the applications could be in any domain (e.g., personalized marketing or AI avatar) but able to intercommunicate via inter domain interoperability. It extends an ontology-based access control method in social networks and a secure method of sharing datasets presented by Masoumzadeh and Joshi [15], it does not provide solutions for upstream challenges like preprocessing data or semantic annotation that would be needed to deploy in a fully ethical end to end manner. Through the whole, we follow privacy-by-design principles, which entails that data is automatically anonymized and consent tracking is context aware.

Moreover, Smith *et al.* [24] developed classification fusion model for a social network (Tweeter/X) data, but in their case, they used a static schema which poses a scalability problem under dynamic settings. Our architecture offers adaptive ontology versioning and associated learning to serve smooth updates on data supported over different datasets. Miller and Mork [16] perspective, they proposed an ethical AI framework for social networks that focused on the mitigation of bias. However, their rule based system does not appear to be semantically interoperable, hence, integration is possible only in the case of homogeneous platforms Huang *et al*. [13]. Our architecture empowers data security using the decentralized identity management, which can help keep GDPR compliance through auditable consent log.

Although previous works have given some insights for ontology-driven and analysis, there are three main limitations:

1. They do not consider privacy-utility trade-offs sufficiently.
2. They do not apply to single domain.
3. They do not sufficiently present ethical restrictions.

To fill these gaps, our architecture makes use of privacy-oriented ontologies such as dynamic consent layers [2, 15], modular, non-domain specific, semantic frameworks with potential cross domain scalability [3, 27] as well as embedded ethical reasoning via accountability metrics as part of the ontology [10, 16]. Our work brings these dimensions together in order to contribute to socially accountable user-generated data integration at scale.

3. The Proposal Architecture

This paper provides a three-phase architecture on mining the data of Internet users and its analysis without privacy and ethical conflicts based on the big data value chain. The architecture includes 1) Data collection, 2) Data preprocessing and analysis, 3) Applications as in Algorithm (1). At the beginning, data is collected from many sources, then advanced analytics techniques are applied for processing data, finally, results can be used in many real time applications. The architecture directs attention to employing ontology in order to insure the semantic interoperability and to structure a service knowledge as demonstrated in Figure 1. The following sections provide a detailed explanation of each phase and its key components.

Algorithm 1: Proposed architecture workflow algorithm.

1. *Start Process*
2. *Phase 1: Data Collection*

- *Collect data from multiple sources (social media, applications, browsing history).*
- *Normalize and convert data into a unified format.*
- *Classify data into predefined ontological categories.*

3. *Phase 2: Data Preprocessing and Analysis*
   - *Categorization and Filtering:*
     - *Cluster data based on interactions, interests, and activities.*
     - *Detect outliers, reduce dimensionality, and normalize data.*
   - *Data Preprocessing:*
     - *Process textual data (tokenization, stop-word removal, feature extraction).*
     - *Develop ontology using the Protégé tool.*
   - *Data Analysis:*
     - *Apply NLP, machine learning, and data visualization techniques.*
     - *Identify patterns and extract insights for decision-making.*
4. *Phase 3: Applications*
   - *User Digital Replication (Create a digital user replica).*
   - *Personalized Advertising and Product Promotion (Targeted ads based on user behavior).*
   - *AI-Enhanced Websites (Personalized website experiences using AI).*
   - *Electronic Personal Agents (Digital assistants that reflect user preferences).*
   - *Customized Social Media Feeds (Content filtering based on user preferences).*
5. *Ensure Compliance with Privacy and Data Protection Policies*
6. *End Process*

The main goal of this paper is to develop a robust architecture for mining Internet user data collected from various platforms while strictly adhering to ethical and privacy principles. The proposed architecture is based on the big data value chain [8, 16], which consists of three phases: data collection, preprocessing and analysis, and applications.

Given the sensitive nature of the data, particularly as it is gathered from multiple platforms, privacy and ethical considerations are necessary. The architecture incorporates mechanisms to ensure data protection, user consent, and compliance with relevant privacy regulations. GDPR and other security and privacy mechanisms can be used. To mitigate privacy risks, data anonymization techniques are employed, enhancing the reliability and trustworthiness of the framework.

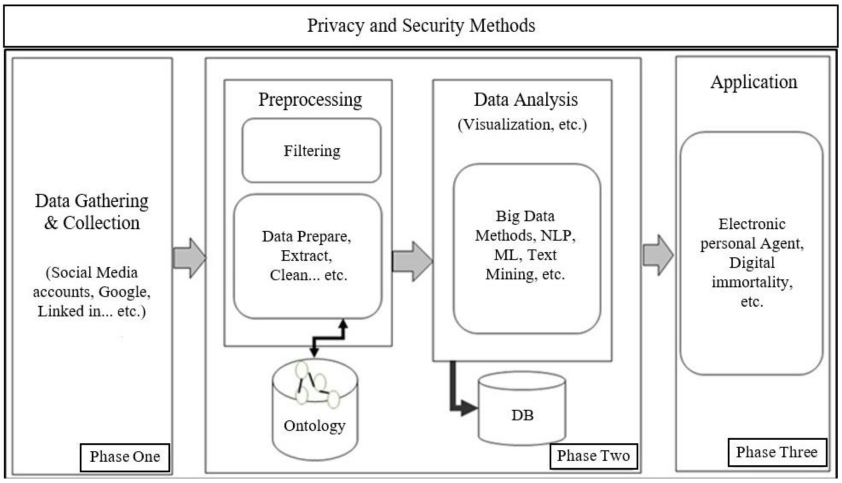

Figure 1. Proposed architecture.

- **Phase 1: Data Collection**

The first phase aims to collect the Internet user data from many heterogeneous sources (e.g., web browsing history, social networks interactions, apps, etc.). At this stage, the user-related behaviors such as likes, comments and check ins, based on the location are collected. The main challenge in this phase is to deal data heterogeneity. Since data comes from heterogeneous platform, it is transformed into a uniform structure, in order to map data to ontology and maintain semantic. To ensure conformity and semantic interoperability across the dataset, we need to normalize the data and to classify it to already existing predefined ontological categories.

- **Phase 2: Data Preprocessing and Analysis**

Data preprocessing and analysis phase receives the data collected/extracted from phase1. It includes many steps, namely, categorization and filtration, data preprocessing, and data analysis:

a) Categorization and Filtration: dividing the gathered data into separate piles corresponding to the social behavior, interest expressed, and its pattern. The filtering process includes outlier detection, dimensionality reduction and data normalization [10, 27].
b) Data Preprocessing: preprocessing tasks differ based on the type of data. In the numeric datasets, preprocessing is mostly omitted, but in the textual datasets, tokenization, stop word removal, feature extraction and feature reduction are included in the preprocessing. Next, extraction of all relevant concepts comes, such concepts are needed for the ontology development, using the Protege development tool. It is a conceptual architecture that formally represents human/digital users' information in the domain by using various properties and

associations. The developed ontology supports sophisticated reasoning and relationship extraction, allowing developers to build smart and data-driven applications.

c) Data Analysis: In this phase, the processed data are examined to obtain valuable findings. Natural Language Processing (NLP), machine learning algorithms, and data visualization methods are used to detect patterns and produce actionable insights in these processes.

• **Phase 3: Applications**

The last step of the architecture includes using the knowledge gained from data analysis to develop a practical application. Examples of potential applications include:

a) User digital replication: it can create a digital representation of an Internet user by aggregating collected data to architecture a virtual representation of an individual in a form of a digital replica of a dead and coma person. This digital immortality concept brings up major ethical and privacy issues that are handled by the architecture with secure data protection protocols and user consent mechanisms [2, 27].

b) It enables development of personalized avatars/virtual agents as digital assistants/agents are able to replace some of the real-life interactions. These agents would be able to represent personalities, preferences, and lifestyles of their users in virtual environments [27].

c) AI-enhanced websites: personalized websites can be constructed based on the user data that can be used to train the AI architectures of the website to adapt dynamically to the user preference. Such a representation of user activities and interests is utilized to carry on personalized interactions and thereby to improve the user experience [14, 15, 16].

d) Customized social networks feed: based on user preference, algorithms analyze user's behavior to fulfil what matches user news feed. This means, to prioritize the content and decide what to view on top, based on user's interests. By discarding the irrelevant content, the user is left with the content that is more likely to lead to meaningful engagement with the user [14, 23].

e) Personalized advertising and product promotion: with a good understanding of the user's behavior and with the help of digital marketing, one can advertise and promote a particular product or service as per the user preferences, needs, and purchase behavior. It makes advertisements more relevant and decreases frustration of user as only highly targeted content is served to the user [21].

A domain ontology serves as a body of knowledge that allows semantic interoperability and as the basis of the proposed architecture which is composed of structuring the knowledge and reasoning capabilities. This semantics approach enables automation in decision making and is based on a new common knowledge graph with new integration pathways so as to enhance the user experience in many areas.

4. Implementation

The core of the implementation is to apply an aggregating and analysis architecture of Internet user data based on the structured ontology. The ontology is developed using the Protege-Web Ontology Language (Protege-OWL) tool [18, 26], consisting of the resource importing, as shown in figure 3, to encapsulate key behavioral patterns, connections, and interactions between whole multiple online platforms. Such a methodology involves conceptualization, specification and implementation towards a full representation of user activities. All datasets for such integration are standardized and fully compliant with privacy regulations. Advanced reasoning techniques allow inference of such complex relationships and provide meaningful insight on what the architecture can do, such as personalized recommendations and behavioral analytics.

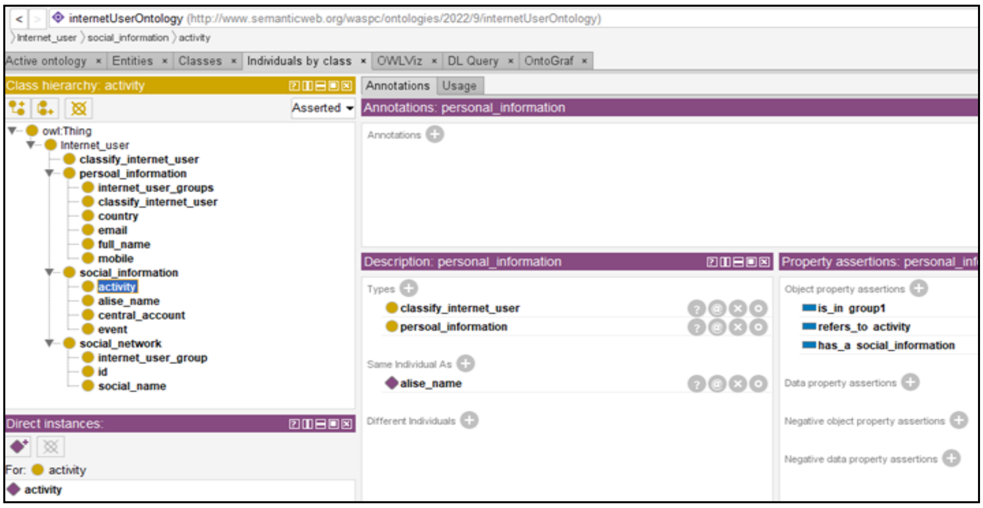

Figure 2. Classes and sub-classes of the domain ontology.

The architecture is evaluated with regard to key performance metrics, including ontology complexity, data integration efficiency, reasoning accuracy, and scalability. It is demonstrated that the ontology mediates large heterogeneous amounts of user data with a degree of reliability and accuracy that is very high. The architecture is validated on a case study in the usability of the architecture in creating digital user profiles in order to examine its applicability to personalized digital services and in user profiling for AI-driven user modeling. Figure 2 shows the content of Internet user domain ontology in the Protege tool. In the sections below, we discuss the steps of implementation, the evaluation metrics as well as important findings.

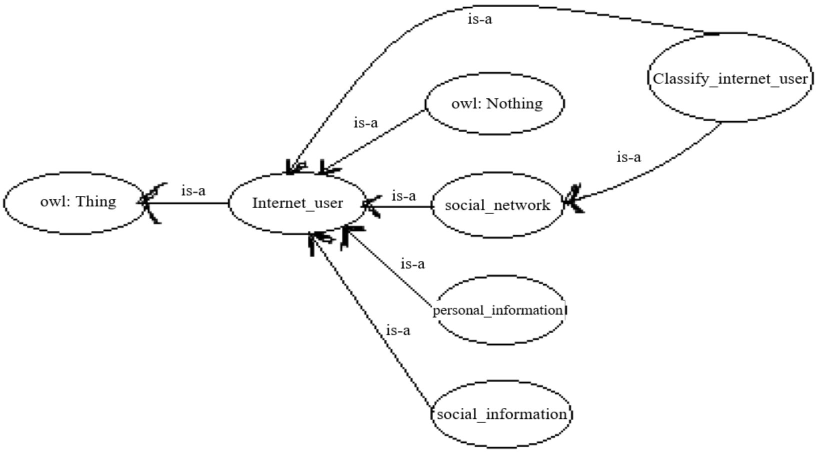

Figure 3. Sample of (is-a) relationship.

Based on the structured methodology, the domain ontology is developed in three phases (i.e., conceptualization, specification and implementation). Fundamental concepts are identified in the conceptualization phase of the Internet user behavior, including entities such as "UserProfile," "activity," "location," and "interaction." Object properties are used to establish relationships between these concepts. For instance, the "is in" property is applied to connect a user's location with a given specific location, whereas "refers to" is used to connect different activities to the suitable categorizations. Figure 3 illustrates a sample of is a relationship. One instance of is a relationship is modeling the activity of a user participating in an online discussion wherein the "UserProfile" entity participates on an "activity" such as "posting" by virtue of the "hasActivity" property.

For example, User1 has_AccountOn X means that a user has an account on Tweeter/X social network platform. Another example is User2 connected_With User3 can be interpreted as one user is a friend of another one.

Table1. Sample relationships.

Individual	Object property	Example
Activity	is_in, refers_to, has_a is_live_in, isPartOf, has_AccountOn, Generates_Data	User1 is_in group1, User1 has_AccountOn X
Event	is_a, has_a, has_name, has_generator,follows, connected_With	User1 follows user2, User2 connected_With User3

A sample scenario applying Internet user digital copy is presented in Figure 4, which can be explained as follows:

- **Scenario**: when an Internet user1 creates any Internet-based account with his own information (e.g. Full name, email) and makes an activity (e.g. writes a post/tweet/etc.), makes an action (e.g. reacts/likes/share/etc.), he joins any social networks group.
- **The action:** the proposed architecture gathers and stores every bit of information (entered by user1, as well as extracted by the architecture) and uses it, after preprocessing and analysis phases, to the creation of a digital copy of the Internet user.

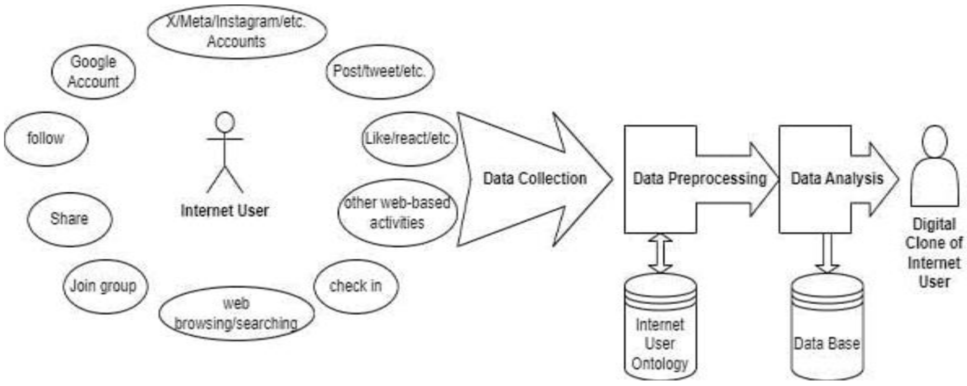

Figure 4. Internet ser digital copy/clone sample scenario.

To test the architecture, simulated datasets from major social networks platforms (Facebook, Twitter, Instagram), and web browsing histories are integrated into the architecture. These datasets are standardized and mapped to the ontology structure for representation uniformity. For instance, a dataset containing user posts, timestamps, and geolocations is transformed into ontology instances in which every post belongs to an instance of the class named "activity", is related to an instance of the class named "UserProfile", and is assigned to an instance of the class named "location", that is recognized using geotagging. Privacy considerations dominates this process; for ensuring adherence to GDPR rules, strong disguising methods are performed, such as data hiding and encryption. This means, instead of storing a user's exact IP, a generalized location identifier (e.g., city level granularity) can be stored to protect user privacy, while still allowing meaningful analysis.

However, we use ontology reasoning to extract useful information from the given data. Based on SPARQL Protocol and RDF Query Language (SPARQL) queries, the architecture that generates high quality advanced relational inferences and recommendations is configured to operate effectively. For example, a query can retrieve trending topics by users in a certain area that filter activities labelled in the course of 'Social Interaction' and rank those topics based on their occurrence. This demonstrates that the architecture can dynamically aggregate and analyze the pattern of user behavior. Additionally, the reason engine is tested in terms of its ability to generate personalized recommendations. For example, if a user always reads posts about "technology" and "AI", interpretation can deduce that the user has high interest in this area and will be promising similar content. This improves the architecture's predictive power and its ability to adapt to changing environments through such reasoning mechanisms.

The developed ontology comprises a total of 50 classes, 21 subclasses, 30 object properties, and 37 assertions, reflecting its depth and representational power. The structural complexity is further reinforced by 154 logical axioms, which enables efficient reasoning and relationship inference. Key relationships such as "is_a" (class hierarchies), "has activity" (users to their actions), and "is_in" (activities with locations) facilitate intuitive mapping of user behaviors. For example, if a "UserProfile" is associated to an "OnlinePurchase" activity and if this event is associated to a "product

category" of "electronics", the architecture can infer a preference for tech-related items, enabling user behavior profiling. The ontology statistics are summarized in Table 2.

Table 2. The ontology statistics.

Metric	No.	Metric	No.
Axiom	154	Object axiom proberty	-
Logical axiom count	103	Functional object property	2
Declaration axioms count	50	Transitive object property	2
Class count	21	Symmetric object property	2
Object property count	9	Asymmetric object property	1
Individual count	21	Individual axiom	-
Annotation property count	1	Class assertion	30
Class axiom	-	Object property assertion	37
SubClassOf	20	Same individual	7
-		**Annotation assertion**	**1**

Figure 5 displays the building of Internet user ontology. The constructed ontology forms a clear architecture, in which all classes and sub classes, that are involved in so, can extract the relationships and collect the data.

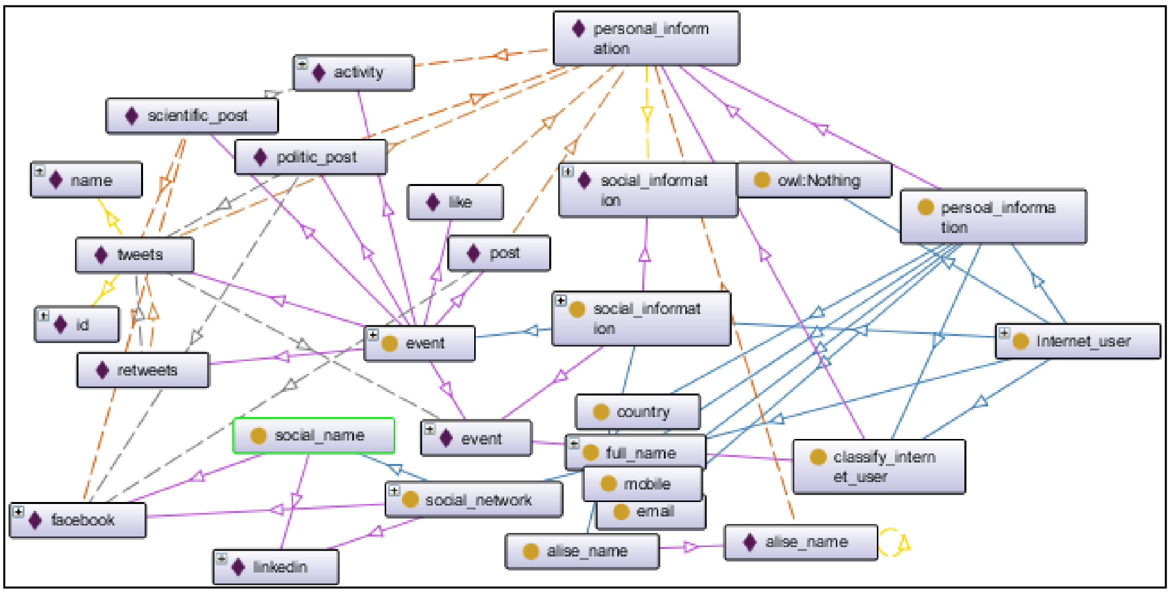

Figure 5. Internet user ontology.

Figure 6 demonstrates example mapping of a post that is extracted from Tweeter/X platform as activity, the resulting JSON parameters (or any other type extracted from the APIs) are then entered to the internet user domain in order to search for the corresponding classes/sub_classes, properties, relations, and individuals. Finally, after extracting all this data the Tweeter post can be mapped to the developed ontology using knowledge graph representation.

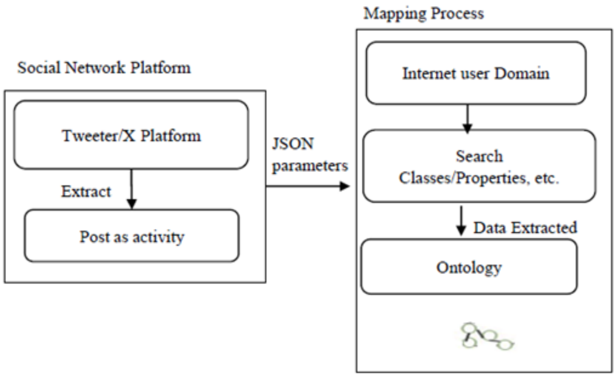

Figure 6. Example mapping from social network platform to ontology.

5. Results Discussions

For assessing the proposed architecture's effectiveness, multiple performance metrics are utilized, including data integration efficiency, ontology reasoning accuracy, and system scalability, where:

- Data integration efficiency: time required for integrating user data from various sources into ontology is recorded. 2.5 seconds are needed to integrate 100 users, which shows that this architecture is built to scale in terms of data ingestion of big data in large volumes.
- Reasoning accuracy: it measures the accuracy of the reasoning process based on path logic to infer user attributes and behavior. More than 93% of the extracted relationships are verified to be correct upon manual validation, demonstrating the extent to which the ontology is capable of representing complex user behavior.
- Scalability: the scalability of the architecture is analyzed by gradually increasing the dataset size. Inference time grows linearly with respect to the number of data points fed to the architecture, but this indicates that the architecture can be run for larger datasets without significant loss of performance.

Case Study

Creation of a digital copy of a user: using the information collected, a case study is performed to create a digital copy of a user that captures user preferences, behaviors, and activities. The digital version created is able to perform basic tasks, like answering questions about the user's interests, suggesting content based on their browsing history, and mimicking the responses the user would give based on his/her own posts. In over 85% of the time, the digital copy responses correctly to relevant questions about the user's behavior, proving the architecture's effectiveness in mimicking user behavior. The digital copy can be applied for digital immortality, where a conversation mimicking a deceased user interacting with the digital copy can be done. However, the ethical concerns emphasizing the user consent and regulation would be crucial for this kind of technology.

The results as a whole demonstrate that the architecture generalizes well on heterogeneous user data and aggregates it well. Ontology-based architecture presents great use in personalized digital services, targeted marketing and digital cloning with high accuracy, high scalability and structured representation.

6. Conclusions

This paper proposes an ontology-based architecture that can mine and analyze heterogeneous Internet user data from different platforms. This architecture offers a solution for merging the heterogeneous user data and reasoning, by applying the idea of semantic web and constructing a domain ontology using Protege tool. The proposed architecture takes into account data protection mechanisms like GDPR, which highly preserves privacy. The architecture presents a scalable cross-platform analysis, while automatically handling a key challenge of personalized service, namely semantic interoperability across heterogeneous data sources. The proposed architecture shows ability for the optimal consolidation of heterogeneous user data with high reasoning accuracy and scalability. Personalized digital marketing, AI-based personal assistants, digital content recommendation, and digital immortality are some application examples where the proposed architecture can be deployed. However, there are few open issues with the privacy of data lifecycle, the data quality (in terms of noise and completeness), and ethical issues of digital avatar of an individual. Future works include enhancing encryption techniques and providing user consent management and guidelines for tracking user data ethically. This provides a basis for AI-based systems that achieve a good balance between personalization and privacy and can evolve according to user needs while maintaining ethical principles in an ever-changing digital environment. Moreover, the performance of the proposed architecture and the reliability and ethical compliance of the real world deployment will be further refined using more sophisticated anonymization techniques and better handling of noisy data.

References

[1] Alromema W. and Alahmadi A., "Ontology Building for Patient Bioinformatics of the Smart Card Domain: Implementation Using OWL," *International Journal of Cloud Computing*, vol. 11, no. 4, pp. 316-329, 2022. https://doi.org/10.1504/IJCC.2022.124796

[2] Andrades J., Rodriguez I., Benavides C., Moreton H., and Gayo J., "An Ontology-Based Multi-

Domain Architecture in Social Networks Analysis: Experimental Validation and Case Study," *Information Sciences*, vol. 540, pp. 390-413, 2020. https://doi.org/10.1016/j.ins.2020.06.008

[3] Arafeh M., Ceravolo P., Mourad A, Damiani E., and Bellini E., "Ontology Based Recommender System Using Social Networks Data," *Future Generation Computer Systems*, vol. 115, no. 1, pp. 769-779, 2021. https://doi.org/10.1016/j.future.2020.09.030

[4] Bendjamaa F. and Taleb N., "OntoDin: An Islamic Ontology of Quran and Hadith," *The International Arab Journal of Information Technology*, vol. 21, no. 5, pp. 773-785, 2024. DOI: https://doi.org/10.34028/iajit/21/5/1

[5] Caon M., "Designing Systems in the Digital Immortality Era," *in Proceedings of the ACM Conference Companion Publication on Designing Interactive Systems*, Hong Kong, pp. 237-241, 2018. https://doi.org/10.1145/3197391.3205442

[6] Deng L., Liu B., and Li Z., "Multimodal Sentiment Analysis Based on a Cross-Modal Multihead Attention Mechanism," *Computers, Materials and Continua*, vol. 78, no. 1, pp. 1157-1170, 2024. https://doi.org/10.32604/cmc.2023.042150

[7] Dwivedi Y., Ismagilova E., Hughes D., Carlson J., and et al., "Setting the Future of Digital and Social Networks Marketing Research: Perspectives and Research Propositions," *International Journal of Information Management*, vol. 59, pp. 102-168, 2021. https://doi.org/10.1016/j.ijinfomgt.2020.102168

[8] Faroukhi A., El Alaoui I., Gahi Y., and Amine A., "Big Data Monetization Throughout Big Data Value Chain: A Comprehensive Review," *Journal of Big Data*, vol. 7, no. 1, pp. 1-22, 2020. https://doi.org/10.1186/s40537-019-0281-5

[9] Galvao V. and Maciel C., "The Acceptability of Digital Immortality: Today's Human is Tomorrow's Avatar," *in Proceeding of the 16th Brazilian Symposium on Human Factors in Computing Systems*, Joinville, pp. 1-4, 2017. https://doi.org/10.1145/3160504.3160580

[10] Garcia S., Gallego S., Luengo J., Benitez J., and Herrera F., "Big Data Preprocessing: Methods and Prospects," *Big Data Analytics*, vol. 1, no. 1, pp. 1-22, 2016. https://doi.org/10.1186/s41044-016-0014-0

[11] Ghani N., Hamid S., Hashem I., and Ahmed E., "Social Networks Big Data Analytics: A Survey," *Computers in Human Behavior*, vol. 101, pp. 417-428, 2019. https://doi.org/10.1016/j.chb.2018.08.039

[12] Google, Data Centers, https://www.google.com/about/datacenters/, Last Visited, 2025.

[13] Huang C., Yang C., and Hsiao Y., "A Novel Framework for Mining Social Media Data Based on Text Mining, Topic Modeling, Random Forest, and DANP Methods," *Mathematics*, vol. 9, no. 17, pp. 20-41, 2021. https://www.mdpi.com/2227-7390/9/17/2041

[14] Labayen V., Magana E., Morato D., and Izal M., "Online Classification of User Activities Using Machine Learning on Network Traffic," *Computer Networks*, vol. 181, pp. 107557, 2020. https://doi.org/10.1016/j.comnet.2020.107557

[15] Masoumzadeh A. and Joshi J., "Ontology-Based Access Control for Social Networks Systems," *International Journal of Information Privacy, Security and Integrity*, vol. 1, no. 1, pp. 59-78, 2011. https://doi.org/10.1504/IJIPSI.2011.043731

[16] Miller H. and Mork P., "From Data to Decisions: A Value Chain for Big Data," *IT Professional*, vol. 15, no. 1, pp. 57-59, 2013. https://doi.org/10.1109/MITP.2013.11

[17] Othman S. and Al-Dhaqm A., "An Improved Machine Learning Method by Applying Cloud Forensic Meta-Architecture to Enhance the Data Collection Process in Cloud Environments," *Engineering, Technology and Applied Science Research*, vol. 14, no. 1, pp. 13017-13025, 2024. https://doi.org/10.48084/etasr.6609

[18] Protege, https://protege.stanford.edu/, Last Visited, 2025.

[19] Puri P., Hassler G., Katragadda S., and Shenk A., "Digital Cloning of Online Social Networks for Language-Sensitive Agent-based Modeling of Misinformation Spread," *PLoS ONE*, vol. 19, no. 6, pp. 1-19, 2024. https://doi.org/10.1371/journal.pone.0304889

[20] Rahman M. and Reza H., "A Systematic Review Towards Big Data Analytics in Social Networks," *Big Data Mining and Analytics*, vol. 5, no. 3, pp. 228-244, 2022. https://doi.org/10.26599/BDMA.2022.9020009

[21] Saenz C., Wong S., Chang Y., and Bravo E., "The Effect of Fair Information Practices and Data Collection Methods on Privacy-Related Behaviors: A Study of Mobile Apps," *Information and Management*, vol. 58, no. 1, pp. 103284, 2021. https://doi.org/10.1016/j.im.2020.103284

[22] Sekerci D. and Alp S., "Investigation of European Union Horizon 2020 Information and Communication Technology Projects with the Social Network Analysis Method," *Engineering Technology and Applied Science Research*, vol. 13, no. 4, pp. 11182-11190, 2023. https://doi.org/10.48084/etasr.5967

[23] Shadbolt N., Lee T., and Hall W., "The Semantic Web Revisited," *IEEE Intelligent Systems*, vol. 21, no. 3, pp. 96-101, 2006. https://doi.org/10.1109/MIS.2006.62

[24] Smith G., Kabban C., Hopkinson K., Oxley M., and et al., "Sensor Fusion for Context Analysis in

Social Media Covid-19 Data," *in Proceedings of the NAECON-IEEE National Aerospace and Electronics Conference*, Dayton, pp. 415-422, 2021. https://doi.org/10.1109/NAECON49338.2021.9696396

[25] Tene O., "What Google Knows: Privacy and Internet Search Engines," *International Association of Privacy Professionals*, vol. 2008, no. 4, pp. 1-72, 2008. https://doi.org/10.63140/eatu6hcsq2

[26] Uschold M. and Gruninger M., "Ontologies: Principles, Methods and Applications," *The Knowledge Engineering Review*, vol. 11, no. 2, pp. 93-136, 1996. https://doi.org/10.1017/S0269888900007797

[27] Yang X., McEwen R., Ong L., and Zihayat M., "A Big Data Analytics Architecture for Detecting User-Level Depression from Social Networks," *International Journal of Information Management*, vol. 54, pp. 102-141, 2020. DOI: 10.1016/j.ijinfomgt.2020.102141

Eman Alduweib received the B.S. and the M.S. degrees in Computer Information Systems from University of Jordan, Amman, Jordan, and the Ph.D. degree in Computer Science from University of Jordan in 2019. She worked as part-time Lecturer in different educational institutions, including University of Jordan and Luminus Technical University College. Currently, she works as Assistant Professor in the Department of Computer Science at the University of Petra, Amman, Jordan. Her research interests include Many Artificial Intelligence Fields, Including Machine Learning and Natural Language Processing.

Abdurahman Alahmadi obtained his Ph.D. in Computer Science and Engineering at Southern Illinois University at Carbondale in 2019. During his studies, he was working in Cloud Computing and Big Data research lab for 5 years. Currently, he works as Associate Professor in Computer Science and Information Department at Taibah University, Saudi Arabia. His Ph.D. research was in Cloud Computing data centers scheduling for energy consumption reduction and resource utilization improvement. His research interests are Machine Learning Resource Management in Cloud Computing, Task Scheduling in Fog Computing and IoT-Supported Edge Offloading Techniques. Dr. Alahmadi has published various peer-reviewed research papers in Edge and Fog Cloud Computing.

Waseem Alromema received his B.S. in CIS from Zarqa University-Jordan in 2005, MSc degree in Information Systems from University of Jordan, Amman, Jordan in 2010, and his Ph.D. in Information Systems from Faculty of Computer and Information Science, Ain Shams University in 2016. He is now work as Associate Professor in Data Science at Taibah University-Saudi Arabia. Dr. Alromema has international publications research interests and Contributions in Various Research Areas, including Information Retrieval, NLP, Machine Learning, and Semantic Web.