A comprehensive solution for semantic knowledge exploration*

Eleonora Bernasconi[1,†], Davide Di Pierro[1,†], Domenico Redavid[1,†] and Stefano Ferilli[1,*,†]

[1] *University of Bari, Bari, Italy*

Abstract

The need for advanced knowledge exploration and discovery tools has become paramount in an age defined by an overwhelming influx of information and ever-increasing data complexity. This paper presents the SKATEBOARD system that is designed to bridge the gap in semantic knowledge exploration. As a holistic solution, SKATEBOARD transcends conventional tools by offering an unparalleled approach to semantic exploration, encompassing data extraction, domain-specific ontology creation, ontology management, and interactive exploration. Through intelligent knowledge extraction, ontology construction, and interactive exploration, it equips researchers and practitioners with the means to confidently traverse the complexities of their domains, make informed decisions, and unearth knowledge that exceeds initial expectations.

Keywords

Semantic Web, Knowledge Graph, Knowledge Extraction, Knowledge Exploration, Semantic Tool,

1. Introduction

This paper introduces an innovative tool, SKATEBOARD (Semantic Knowledge Advanced Tool for Extraction Browsing Organization Annotation Retrieval and Discovery), designed to address the pressing need for advanced knowledge exploration and discovery in an era characterized by exponential information growth and escalating data complexity.

Graphs have emerged as a robust means of information representation, particularly when the goal is to derive knowledge and unearth hidden patterns. Unlike conventional tabular data structures, graphs excel at capturing intricate relationships between entities, rendering them an

2nd Italian Workshop on Artificial Intelligence for Cultural Heritage (IAI4CH 2023, https://ai4ch.di.unito.it/), co-located with the 22nd International Conference of the Italian Association for Artificial Intelligence (AIxIA 2023). 6-9 November 2023, Rome, Italy

[*] Corresponding author.

[†] These authors contributed equally.

✉ eleonora.bernasconi@uniba.it (E. Bernasconi); davide.dipierro@uniba.it (D. Di Pierro); domenico.redavid1@uniba.it (D. Redavid); stefano.ferilli@uniba.it (S. Ferilli)

🌐 https://www.uniba.it/it/docenti/eleonora-bernasconi (E. Bernasconi); http://conceptbase.sourceforge.net/mjf/ (D. Di Pierro); http://www.di.uniba.it/~redavid/ (D. Redavid); http://lacam.di.uniba.it/people/ferilli.html (S. Ferilli)

iD 0000-0003-3142-3084 (E. Bernasconi); 0000-0002-8081-3292 (D. Di Pierro); 0000-0003-2196-7598 (D. Redavid); 0000-0003-1118-0601 (S. Ferilli)

ideal choice for knowledge representation and discovery. Through graph-based approaches, users gain the ability to seamlessly navigate and explore interconnected data, revealing valuable insights that may remain concealed in other representations.

Artificial Intelligence (AI) has ushered in a revolution in research methodologies by empowering machines to comprehend, interpret, and reason with vast datasets. Notable advancements in AI, such as OpenAI's GPT (Generative Pre-trained Transformer) [1], exemplify AI models' capacity to "reason" semantically, extracting meaningful conclusions from raw data. These capabilities hold profound implications for knowledge discovery, presenting researchers with novel avenues to explore and enrich their comprehension of complex phenomena.

Yet, amidst these strides in AI and semantic technologies, a crucial gap persists—a deficiency in tools that facilitate knowledge extraction and exploratory visualization transparently. Frequently, existing AI-based solutions yield outputs that elude full comprehension or user control, triggering concerns about trust and accountability. To remedy this, we introduce SKATEBOARD, a novel framework and tool that empowers users with complete command over all stages of information extraction and manipulation.

SKATEBOARD embodies a multi-faceted approach, encompassing the extraction of pertinent information, the creation of domain-specific ontologies rooted in the extracted data, streamlined ontology management, and a robust platform for interactive exploration. By embracing Linked Data principles like graph-based exploration [1], the tool furnishes users with unparalleled transparency, permitting interactive navigation through information with comprehensive visibility into relationships and dependencies. Furthermore, SKATEBOARD unlocks the potential for recommendation systems and reasoning capabilities, facilitating serendipitous discoveries and novel insights.

Subsequent sections will delve into SKATEBOARD's functionalities, illustrating how this tool empowers users to harness semantic technologies fully, unlocking their data's latent potential. Through an amalgamation of intelligent knowledge extraction, ontology construction, and interactive exploration, SKATEBOARD emerges as a promising stride in the domain of knowledge discovery and management. Its user-centric design guarantees that researchers and practitioners can confidently navigate the intricacies of their domain, make informed decisions, and uncover knowledge surpassing initial expectations.

2. Related Work

In this section, we explore tools that share common characteristics with SKATEBOARD, with a particular focus on knowledge extraction, information retrieval, semantic data visualization, and broader utilization of semantic technologies. We analyze these tools, comparing their functionalities and approaches, while also highlighting distinctions and similarities with the SKATEBOARD system.

Many Linked Data interfaces focus on visualizing SPARQL endpoints [2] and Linked Data [2, 3], but SKATEBOARD distinguishes itself by integrating an API that connects to SPARQL endpoints, providing versatility to visualize and create Linked Data. This sets SKATEBOARD

[1] https://platform.openai.com/

[2] w3.org/wiki/SparqlEndpoints

apart, as it addresses challenges in extracting semantic knowledge from unstructured texts and semantic annotation, offering a comprehensive solution for the entire Linked Data lifecycle.

Our analysis of knowledge extraction tools [4, 5, 6] reveals a dynamic landscape. SKATEBOARD's comprehensive approach stands out, transforming unstructured text into structured data, performing Named Entity Recognition (NER), linking entities to knowledge bases, and enhancing data through semantic annotation and ontology-based integration via the GraphBRAIN system [7]. While the cited tools excel in specific aspects, SKATEBOARD offers a powerful and holistic solution that encompasses the entire knowledge extraction pipeline.

Traditional visual information seeking tools (e.g lerma https://www.lerma.it/ or Torrossa https://www.torrossa.com/) have long served data retrieval, but SKATEBOARD redefines the landscape by introducing semantic entities that enable precise and refined searches, serendipitous exploration, and intelligent recommendations, enriching the user experience. While implementing semantic technologies may involve initial investments, SKATEBOARD's enhanced search capabilities justify the cost over the long term.

Tools for the visualization of semantic data [8, 9, 10, 11, 12, 13, 14, 15, 16] are categorized by interaction paradigms and types of information. SKATEBOARD goes beyond these paradigms by customizing visualizations based on entity types, offering a dynamic and user-centric experience. As the field evolves, innovative solutions like SKATEBOARD are expected to cater to diverse knowledge exploration needs across various domains.

Semantic annotation tools [17, 18, 19, 20, 21] enrich documents with semantic information, and SKATEBOARD excels with its integration of GraphBRAIN [7] and collaborative validation of knowledge extraction. It offers a comprehensive solution for semantic annotation tasks, setting it apart from traditional tools.

Digital library exploration tools [22, 23, 24] transform digital libraries into dynamic spaces. SKATEBOARD enhances the exploration experience by dynamically adapting to individual preferences, enabling users to delve deeply into knowledge relationships, and facilitating collaborative validation of knowledge extraction (a feature inherited from the Arca system [25]). In comparison, other tools may have limitations in customization and collaborative data improvement.

In summary, SKATEBOARD distinguishes itself by providing a comprehensive solution for knowledge extraction, semantic annotation, and dynamic exploration within digital libraries, surpassing the capabilities of many existing tools. Its user-centric design and integration with GraphBRAIN [7] position it at the forefront of Linked Data interfaces, offering users a richer and more intuitive experience for navigating the vast expanse of knowledge.

3. The Pipeline

The SKATEBOARD pipeline serves as the framework's backbone, facilitating knowledge extraction, management, and interactive visualization through semantic technologies. It offers a transparent, structured path guiding users through essential phases for an efficient workflow.

Knowledge extraction. At the pipeline's onset, relevant information is extracted from diverse data sources, allowing users to specify structured data, unstructured text, databases, or various file formats. This step is pivotal in building the initial knowledge base.

Preprocessing and semantic enrichment. Following data extraction, SKATEBOARD initiates preprocessing to ensure data consistency with the source domain and enhance quality. Named Entity Recognition (NER) identifies entities in text sentences, and Named Entity Linking (NEL) disambiguates and links entities to databases or ontologies, structuring the data for better comprehension.

Ontology creation and management. Extracted and prepared data is sent to GraphBRAIN, enabling users to craft customized domain-specific ontologies, defining classes, properties, and relationships. GraphBRAIN offers tools for ongoing ontology management, adapting to evolving knowledge within the domain.

Connection to multiple endpoints. SKATEBOARD's distinctive feature is its ability to connect to multiple endpoints simultaneously. Beyond GraphBRAIN, it interacts with other systems and data sources, broadening its scope and utility (e.g. DBpedia [3]).

Visualization and interactive exploration. SKATEBOARD provides an advanced research platform for visualizing and exploring semantically enriched data interactively. The user interface enables users to filter, navigate, and analyze object relationships in real-time. Visualizations adapt based on selected entity types, enhancing data analysis.

4. Knowledge extraction in SKATEBOARD

This section delves into the intricacies of the knowledge extraction process using SKATEBOARD, with a specific focus on its integration with GraphBRAIN [7].

4.1. Data Identification

The initial phase of knowledge extraction entails identifying pertinent data. It is imperative to precisely define the domain of interest, pinpoint relevant data sources, and establish appropriate data acquisition methods. The choice of the research domain plays a pivotal role in shaping the selection of data sources and extraction techniques, ranging from the vast realm of literature to specialized fields like archaeology.

Data sources come in various forms: structured, semi-structured, or unstructured, and can be found in repositories such as digital libraries, online encyclopedias, structured databases, semi-structured documents, and fully unstructured texts. The selection of data acquisition methods hinges on the data's nature and its source; for instance, web crawlers may be deployed for web resource acquisition, while data mining techniques may be indispensable for database data extraction. The primary goal of this phase is to acquire and prepare the data essential for constructing the knowledge graph.

4.2. Construction of the Knowledge Graph Ontology

The subsequent phase revolves around constructing the ontology that underpins the knowledge graph, furnishing it with a high-level structure. This phase assumes paramount significance when an existing domain ontology can serve as the foundation for the knowledge graph ontology or when working with structured data that offers a framework for ontology creation.

[3] https://dbpedia.org/sparql

Building the ontology for the knowledge graph involves defining predefined entity types and their relationships. Common ontologies such as FOAF [4] or Geonames [5], along with established ontology languages like RDF(S) [26], OWL [27], and XML [28], can be employed in this endeavor. Notably, SKATEBOARD is seamlessly integrated with GraphBRAIN, permitting domain experts to manually develop and maintain the ontology, ensuring its alignment with specific domain requirements. Furthermore, SKATEBOARD can connect with various ontology sources, including DBpedia [10], to enhance available knowledge.

4.3. Knowledge Extraction

After data acquisition and ontology definition, the subsequent step entails the extraction of knowledge from the amassed data. The primary objective in this phase is to extract entities, establish relationships among them, and capture meaningful attributes.

Entity extraction involves identifying and categorizing entities from diverse data sources. SKATEBOARD harnesses Named Entity Recognition (NER) [29] to classify entities into predefined categories or types, subsequently linking them to relevant ontologies such as DBpedia and GraphBRAIN through named entity linking (NEL) [30].

Relation extraction, vital for connecting entities, varies depending on the data's nature but employs natural language processing (NLP) techniques [31] for unstructured data. The integration of ontologies within SKATEBOARD facilitates the assignment of relationships between extracted entities, based on predefined definitions.

4.4. Knowledge Processing

In the initial stage of knowledge extraction, knowledge processing is the essential step for enhancing the reliability of extracted data. It concentrates on addressing vagueness, optimizing information coherence, and mitigating gaps, thereby ensuring a more refined knowledge foundation.

4.5. Knowledge Integration

Subsequent to knowledge processing, knowledge integration plays a distinct role by amalgamating insights from diverse origins to construct a unified knowledge framework.

Knowledge integration, also referred to as knowledge fusion, entails amalgamating information from diverse sources while eliminating redundancy, contradictions, and ambiguities. This process encompasses entity resolution, and the assignment of unique identifiers to entities. Entity resolution is a pivotal step that aims to determine if different entities refer to the same real-world objects, effectively connecting them in the knowledge graph.

4.6. Knowledge Completion

The ultimate goal of this phase is to comprehensively enrich the knowledge within the knowledge graph, involving reasoning, triple validation, and optimization.

[4] http://www.foaf-project.org/

[5] https://www.geonames.org/

Reasoning on knowledge relies on predefined rules between relationships and may employ machine learning methods to unearth new knowledge from existing information. Triple validation ensures that only valid and pertinent information finds its place in the knowledge graph, subject to integrity constraints and other stipulated conditions [25].

Optimizing the knowledge graph might entail the removal of nodes or relationships unrelated to the domain of interest, thereby contributing to maintaining a coherent and logical structure.

We propose SKATEBOARD as a versatile tool for knowledge extraction. Its synergy with GraphBRAIN [7] and seamless integration with external ontologies like DBpedia streamline the entire knowledge graph creation and management process. This systematic approach ensures the generation of knowledge graphs suited for diverse applications across various domains.

5. Interactive Graph Exploration in SKATEBOARD

The SKATEBOARD interface introduce innovations in knowledge exploration, offering a powerful platform to dive into vast knowledge bases. It starts with a simple keyword entry in the search bar, triggering a tailored query based on the connected endpoint. Results are presented in a tabular list, displaying nodes within the reference knowledge graph containing the search term in their labels, along with related nodes ranked by similarity.

Exploration Modes. Users can easily select a resource of interest and drag it to the central dashboard. Each graph node offers two exploration modes: - **Primary Connections**: Visualizes relationships closely linked to the selected node. - **Dedicated Table**: Presents essential information related to the selected entity's type.

Entity-Specific Views. A unique feature is the availability of entity-specific views. These views allow the visualization of closely connected and complex relationships. For instance, selecting an author can lead to a map visualization of all publication locations related to their works.

User Profiling and Suggestions. SKATEBOARD anonymously profiles users, offering topic suggestions based on their interests. A history of searched topics helps track areas of interest and revisit prior searches.

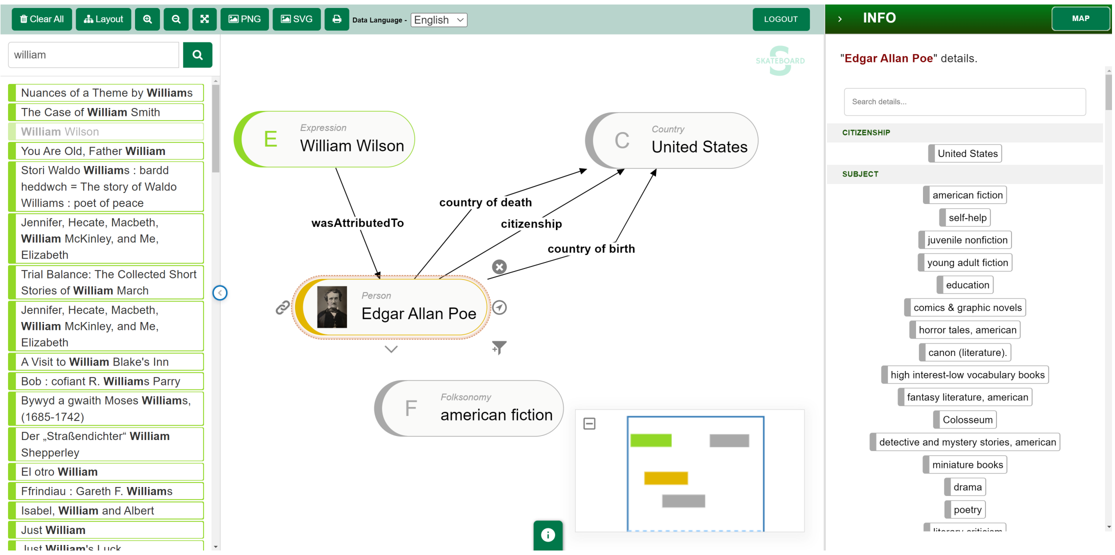
Property Graphs. SKATEBOARD goes beyond knowledge graphs, incorporating labeled property graphs (LPG) [32] like Neo4j [6]. This flexibility allows visualization with custom domain ontologies and seamless integration of data from public endpoints like DBpedia with proprietary ontologies.

Collaborative Validation. Users play a pivotal role in enhancing data quality within connected knowledge bases by contributing to high-quality content creation for domain experts.

Figure 1 illustrates the exploration of an entity of type 'Person' providing specific information tailored to that entity type and the associated endpoint, Figure 2 showcases an entity of type 'Expression' (e.g., a literary work) with pertinent information displayed on the right. Subsequently.

The SKATEBOARD interface emphasizes the connected endpoint, creating a collaborative space for domain experts to share knowledge and enrich their information repository. It com-

[6] https://neo4j.com/



william

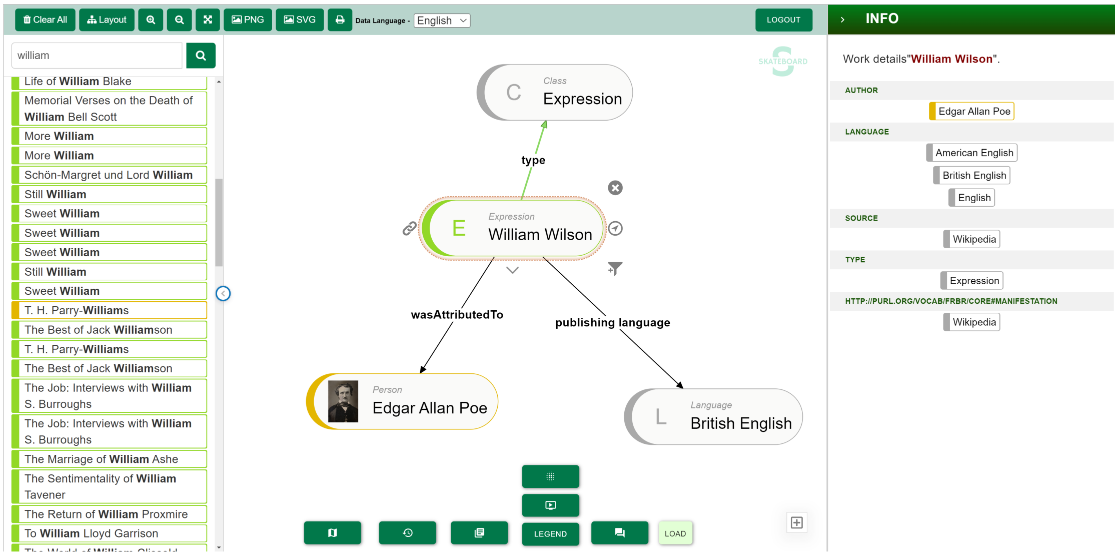
Nuances of a Theme by Williams
The Case of William Smith
William Wilson
You Are Old, Father William
Stori Waldo Williams : bardd heddwch = The story of Waldo Williams : poet of peace
Jennifer, Hecate, Macbeth, William McKinley, and Me, Elizabeth
Trial Balance: The Collected Short Stories of William March
Jennifer, Hecate, Macbeth, William McKinley, and Me, Elizabeth
A Visit to William Blake's Inn
Bob : cofiant R. Williams Parry
Bywyd a gwaith Moses Williams, (1685-1742)
Der „Straßendichter" William Shepperley
El otro William
Ffrindiau : Gareth F. Williams
Isabel, William and Albert
Just William
Just William's Luck

"Edgar Allan Poe" details.

CITIZENSHIP
United States

SUBJECT
american fiction
self-help
juvenile nonfiction
young adult fiction
education
comics & graphic novels
horror tales, american
canon (literature).
high interest-low vocabulary books
fantasy literature, american
Colosseum
detective and mystery stories, american
miniature books
drama
poetry
literary criticism

Figure 1: SKATEBOARD interface - part 1



william

Life of William Blake
Memorial Verses on the Death of William Bell Scott
More William
More William
Schön-Margret und Lord William
Still William
Sweet William
Sweet William
Sweet William
Still William
Sweet William
T. H. Parry-Williams
The Best of Jack Williamson
T. H. Parry-Williams
The Best of Jack Williamson
The Job: Interviews with William S. Burroughs
The Job: Interviews with William S. Burroughs
The Marriage of William Ashe
The Sentimentality of William Tavener
The Return of William Proxmire
To William Lloyd Garrison
The World of William Clissold

Work details"William Wilson".

AUTHOR
Edgar Allan Poe

LANGUAGE
American English
British English
English

SOURCE
Wikipedia

TYPE
Expression

HTTP://PURL.ORG/VOCAB/FRBR/CORE#MANIFESTATION
Wikipedia

Figure 2: SKATEBOARD interface - part 2

bines various visualization paradigms for linked data and graphs, including node-link, tabular, and multilevel visualizations. This approach enables incremental exploration of connected resources, unveiling paths that link graph entities and harnessing the potential of semantic integration and graph reasoning within LPG.

6. Conclusion and future work

As we enter the evaluation phase of the system, which is currently accessible online at the following link: http://digitalmind.di.uniba.it:3000/, initial feedback has shown significant enthusiasm among the researchers utilizing the platform. However, it is important to note that a comprehensive evaluation, enriched by the collection of both qualitative and quantitative data, is currently in progress.

In conclusion, we believe that the system presented in this paper represents a promising solution for the creation and exploration of domain-specific content. The power of artificial intelligence extends beyond knowledge extraction; it also influences the results displayed within the graph navigation interface. The system's modularity and integration of knowledge from various knowledge bases enable SKATEBOARD to harness the capabilities of advanced reasoning algorithms provided by artificial intelligence.

Looking ahead, the future work on SKATEBOARD involves refining the system based on user feedback and enhancing its capacity to extract, manage, and visualize knowledge. We aim to conduct comprehensive user studies to gather qualitative insights and quantify the system's performance metrics. Additionally, we plan to expand the system's capabilities by integrating advanced AI algorithms that can offer even more intelligent knowledge recommendations and insights.

In summary, the journey of SKATEBOARD is ongoing, and we anticipate that it will continue to evolve into a robust tool for knowledge creation and exploration, driven by the synergy of artificial intelligence and a modular, knowledge-based approach.

Acknowledgments

This research was partially supported by projects FAIR – Future AI Research (PE00000013), spoke 6 – Symbiotic AI, and CHANGES – Cultural Heritage Active innovation for Next-GEn Sustainable society (PE00000020), Spoke 3 – Digital Libraries, Archives and Philology, under the NRRP MUR program funded by the NextGenerationEU.

References

[1] A. Veremyev, A. Semenov, E. L. Pasiliao, V. Boginski, Graph-based exploration and clustering analysis of semantic spaces, Applied Network Science 4 (2019) 1–26.

[2] C. Bizer, T. Heath, T. Berners-Lee, Linked data-the story so far, in: Linking the World's Information: Essays on Tim Berners-Lee's Invention of the World Wide Web, 2023, pp. 115–143.

[3] E. Bernasconi, M. Ceriani, D. D. Pierro, S. Ferilli, D. Redavid, Linked data interfaces: A survey, Information 14 (2023). URL: https://www.mdpi.com/2078-2489/14/9/483. doi:10.3390/info14090483.

[4] J. Hoffart, M. Yosef, I. Bordino, H. Fürstenau, M. Pinkal, M. Spaniol, B. Taneva, S. Thater, G. Weikum, Robust disambiguation of named entities in text, in: Proceedings of the 2011 conference on empirical methods in natural language processing, 2011, pp. 782–792.

[5] A. A. Sinaci, S. Gonul, Semantic content management with apache stanbol, in: The Semantic Web: ESWC 2012 Satellite Events: ESWC 2012 Satellite Events, Heraklion, Crete, Greece, May 27-31, 2012. Revised Selected Papers 9, Springer, 2015, pp. 371–375.

[6] J. Hoffart, F. M. Suchanek, K. Berberich, G. Weikum, Yago2: A spatially and temporally enhanced knowledge base from wikipedia, Artificial intelligence 194 (2013) 28–61.

[7] S. Ferilli, D. Redavid, The graphbrain system for knowledge graph management and advanced fruition, in: Foundations of Intelligent Systems: 25th International Symposium, ISMIS 2020, Graz, Austria, September 23–25, 2020, Proceedings, Springer, 2020, pp. 308–317.

[8] T. Berners-Lee, Y. Chen, L. Chilton, D. Connolly, R. Dhanaraj, J. Hollenbach, A. Lerer, D. Sheets, Tabulator: Exploring and analyzing linked data on the semantic web, in: Proceedings of the 3rd international semantic web user interaction workshop, volume 2006, Athens, Georgia, 2006, p. 159.

[9] T. Berners-Lee, J. Hollenbach, K. Lu, J. Presbrey, E. Prud'ommeaux, M. Schraefel, Tabulator redux: Browsing and writing linked data (2008).

[10] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, Z. Ives, Dbpedia: A nucleus for a web of open data, in: international semantic web conference, Springer, 2007, pp. 722–735.

[11] A. Nuzzolese, V. Presutti, A. Gangemi, A. Musetti, P. Ciancarini, Aemoo: exploring knowledge on the web, in: Proceedings of the 5th Annual ACM Web Science Conference, 2013, pp. 272–275.

[12] A. Micsik, Z. Tóth, S. Turbucz, Lodmilla: Shared visualization of linked open data, in: Theory and Practice of Digital Libraries – TPDL 2013 Selected Workshops, Springer International Publishing, Heidelberg, 2014, p. 89–100.

[13] F. Viola, L. Roffia, F. Antoniazzi, A. D'Elia, C. Aguzzi, T. Salmon Cinotti, Interactive 3d exploration of rdf graphs through semantic planes, Future Internet 10 (2018).

[14] M. Ceriani, P. Bottoni, Sparqlblocks: using blocks to design structured linked data queries, Language (XSD) 1 (2017) 11.

[15] F. Haag, S. Lohmann, T. Ertl, Sparqlfilterflow: Sparql query composition for everyone, in: Extended Semantic Web Conference (ESWC), 2014, p. 362–367. doi:10.1007/978-3-319-11955-7_49.

[16] N. Marie, F. Gandon, M. Ribiere, F. Rodio, Discovery hub: on-the-fly linked data exploratory search, in: Proceedings of the 9th Int. Conf. on Semantic Systems, ACM, 2013, p. 17–24.

[17] J. M. Cejuela, P. McQuilton, L. Ponting, S. J. Marygold, R. Stefancsik, G. H. Millburn, FlyBase Consortium, tagtog: interactive and text-mining-assisted annotation of gene mentions in plos full-text articles, Database 2014 (2014).

[18] A. Loreggia, S. Mosco, A. Zerbinati, Sentag: A web-based tool for semantic annotation of textual documents, in: ThirtySixth AAAI Conference on Artificial Intelligence, AAAI Press, 2022.

[19] A. Kumar, M. Spaniol, Annotag: Concise content annotation via lod tags derived from entity-level analytics, in: Linking Theory and Practice of Digital Libraries, Springer International Publishing, 2021, p. 175–180.

[20] R. Simon, et al., Linked data annotation without the pointy brackets: Introducing recogito 2, Journal of Map Geography Libraries 13 (2017) 111–132.

[21] G. Giannopoulos, N. Bikakis, T. Dalamagas, T. Sellis, Gontogle: a tool for semantic annotation and search, in: Extended Semantic Web Conference, Springer, Berlin, Heidelberg,

2010, pp. 376–380.

[22] E. Bernasconi, M. Ceriani, M. Mecella, Academic research creativity archive (arca), in: International Conference on Research Challenges in Information Science, Springer, Barcelona, 2021, pp. 713–714. doi:10.1007/978-3-030-77554-3_44.

[23] M. Bolina, Yewno Discover, Nordic Journal of Information Literacy in Higher Education 11 (2019).

[24] E. Ikkala, E. Hyvönen, H. Rantala, M. Koho, Sampo-ui: A full stack javascript framework for developing semantic portal user interfaces, Semantic Web (2021) 1–16. doi:10.3233/SW-210428.

[25] E. Bernasconi, M. Ceriani, M. Mecella, A. Morvillo, Automatic knowledge extraction from a digital library and collaborative validation, in: International Conference on Theory and Practice of Digital Libraries, Springer, Padua, 2022, p. 480–484. doi:10.1007/978-3-031-16802-4_49.

[26] B. McBride, The resource description framework (rdf) and its vocabulary description language rdfs, in: Handbook on ontologies, Springer, 2004, pp. 51–65.

[27] B. Parsia, P. Patel-Schneider, B. Motik, Owl 2 web ontology language structural specification and functional-style syntax, in: W3C, W3C Recommendation, 2012.

[28] R. Ghawi, N. Cullot, Building ontologies from xml data sources, in: 2009 20th International Workshop on Database and Expert Systems Application, IEEE, 2009, pp. 480–484.

[29] K. Byrne, Nested named entity recognition in historical archive text, in: International Conference on Semantic Computing (ICSC 2007), IEEE, 2007. doi:10.1109/icsc.2007.107.

[30] M. Röder, R. Usbeck, A. Ngomo, Gerbil – benchmarking named entity recognition and linking consistently, Semantic Web 9 (2018) 605–625. doi:10.3233/sw-170286.

[31] A. Menon, J. Choi, H. Tabakovic, What you say your strategy is and why it matters: natural language processing of unstructured text, in: Academy of management proceedings, volume 1, Academy of Management Briarcliff Manor, NY 10510, 2018, p. 18319.

[32] D. Di Pierro, S. Ferilli, D. Redavid, Lpg-based knowledge graphs: A survey, a proposal and current trends, Information 14 (2023) 154. doi:10.3390/info14030154.

A. Online Resources

You can access the proof of concept for our system at the following URL: http://digitalmind.di.uniba.it:3000.