Using Omeka S in Cataloguing Stanisław Lem's Letters

Szymon Piotr Kukulak[2], Luiz do Valle Miranda[1], Jakub Gomułka[2,*], Krzysztof Kutt[1] and Grzegorz J. Nalepa[1]

[1]*Department of Human-Centered Artificial Intelligence, Institute of Applied Computer Science, Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, prof. Stanisława Łojasiewicza 11, 30-348 Kraków, Poland*
[2]*Faculty of Humanities, AGH University of Kraków, Czarnowiejska 36, 30-054, Kraków, Poland.*

Abstract

The article explores the details of the process of collecting metadata of Stanisław Lem's correspondence as a part of a broader Lem Knowledge Graph (LKG) project. Data collection will be carried out via the Omeka S system, in which the intermediary lightweight ontology structures were implemented. The annotation process, in addition to assigning values to successive instances of the ontology's main class—*L0_Letter*—will also involve the creation of instances of several other classes. Upon completion of the cataloguing process, the collected data will be integrated into the main LKG ontology.

Keywords

Cultural Heritage, Stanisław Lem, Linked Data, Knowledge Graph, Omeka S, Cataloguing Correspondence

1. Introduction

The Lem Knowledge Graph (LKG) is an ongoing project of applying Semantic Web (SW) technologies for the creation of a comprehensive catalogue of Stanisław Lem's (1921–2006) literary heritage. Lem is mostly known for his numerous science fiction novels and stories, but also for his non-fiction works. During his lifetime, Lem has written several thousands of letters in which further information about his work and his personal life can be found. Some selected sets of Lem's letters have been published, such as his exchanges with Ursula Le Guin and Michael Kandel (the translator of his prose into English). However, most of the preserved letters have not yet been published. The LKG.letters module, one of the branches of the LKG project, is planned to collect metadata about Lem's letters.

The present paper aims to outline the implemented workflow for metadata acquisition in the case of Lem's letters. As the goal of LKG is to conform to the standards of Linked Data (LD), including the use of interoperable ontologies and thesauri, an essential requirement is the adhesion to rigorous metadata standards and consistent semantic annotation of the collected data to ensure interoperability, reusability, and long-term preservation. As manual metadata collection is a repetitive process and widely prone to errors, the use of software is strongly recommended to automate repetitive tasks such as metadata validation, ontology alignment, and data integration. Since current resources and timeframe exclude the possibility of developing dedicated software or the purchase of licences, we have been constrained to utilise open-source software.

We have decided to use Omeka S as an open source solution for the linked data management platform. Omeka S is a content management system specifically designed to manage and publish digital collections of cultural heritage. It supports the use of ontologies for controlled vocabularies, linking resource to external URIs, and easily prototyping an interface for resource search and browsing. Data can be inserted into Omeka S both manually and via API.

Proceedings of the Joint Ontology Workshops (JOWO) - Episode XI: The Sicilian Summer under the Etna, co-located with the 15th International Conference on Formal Ontology in Information Systems (FOIS 2025), September 8-9, 2025, Catania, Italy
*Corresponding author.
skukulak@agh.edu.pl (S. P. Kukulak); luiz.dovallemiranda@doctoral.uj.edu.pl (L. d. V. Miranda); jgomulka@agh.edu.pl (J. Gomułka); krzysztof.kutt@uj.edu.pl (K. Kutt); gjn@gjn.re (G. J. Nalepa)
0009-0009-2701-6567 (S. P. Kukulak); 0000-0003-1838-5693 (L. d. V. Miranda); 0000-0002-9100-0334 (J. Gomułka); 0000-0001-5453-9763 (K. Kutt); 0000-0002-8182-4225 (G. J. Nalepa)

2. Related works

Information of the use of software for data entry in the field of cultural heritage can be gathered both from comparison articles and from explicit use cases. [1] identifies three widely adopted platforms: DSpace[1], ePrints[2] and Greenstone[3]. These three platforms were developed and released in the late 1990s and early 2000s, therefore they did not natively support linked data. DSpace stands out as an open-source repository platform designed to store, manage, and disseminate digital content. It is commonly employed in over 1000 repositories distributed over academic institutions, libraries, and other organizations to preserve institutional research outputs and digital assets. DSpace supports various metadata standards, enables interoperability through protocols like OAI-PMH, and provides a user-friendly interface for collection management. [2] discusses how DSpace was utilized in building the Taiwan Digital Library History Library, serving both as a metadata repository and a user interface. Similarly, [3] investigates its impact at institutions such as the University of New Mexico, the University of Washington, and Ohio State University, where it was used for metadata harvesting and repository functions, contributing to improvements in metadata quality. Updates to DSpace have made possible the adoption of LD principles, as the platform can expose stored content and metadata in a way that makes them machine-readable and linkable with other datasets on the web.

ePrints is an open-source digital repository software developed to support the creation, curation, and dissemination of scholarly content. It offers a high degree of flexibility in metadata schema configuration and supports standard interoperability protocols. One of its updates was the feature allowing the exposition of metadata in RDF format. Designed with usability in mind, the platform provides intuitive tools for both repository managers and end users. As illustrated by [4], ePrints was employed by the Universitas Brawijaya Library to foster metadata interoperability between institutions using different digital library platforms. Greenstone is another open-source platform supporting data insertion and management and is widely used in cultural heritage projects, though it lacks built-in support for linked data. Information about recent enhancements in Greenstone is provided in [5], and various implementation examples can be found at https://www.greenstone.org/examples.

A comparison of more recent CH platforms that can be used for data insertion can be found in [6]. Among the presented platforms are Omeka S[4] and Arches[5]. Arches is a freely available software platform specifically developed for the management and dissemination of cultural heritage information, with a strong emphasis on LD to ensure semantic interoperability. In the study by [7], the Arches[6] system is leveraged to capture and harmonize Taiwan's cultural heritage records using the CIDOC CRM ontology[7]. Omeka S[8] is an open-source web publishing platform designed for organizing, describing, and presenting digital collections. It is particularly suited for working with LD, offering support for RDF-based metadata, customizable vocabularies, and integration with external semantic web services. Omeka S features a flexible, user-friendly interface that enables users to curate digital exhibits and manage complex collections. The implementation described in [8] details its use as a comprehensive solution for metadata input, storage, and public access at the Faculty of Mining and Geology's Digital Repository at the University of Belgrade. Another application is presented in [9], where Omeka S was employed at the Archives Henri Poincaré to support semantic enrichment, structured data storage, and advanced browsing and querying functionalities for Poincaré's correspondences.

It is worth mentioning CorrespSearch[9], a specialized platform developed to facilitate the scholarly exchange and analysis of correspondences. CorrespSearch provides the "CMIF Creator" tool for creating a digital index of letters, as well as browse and search functionalities for aggregated metadata.

[1] https://www.dspace.org/
[2] https://www.eprints.org/
[3] http://www.greenstone.org/
[4] https://omeka.org/s/
[5] https://www.archesproject.org/
[6] In this case, the system was based on the Heritage Inventory Package, an extension to the Arches system.
[7] https://cidoc-crm.org/
[8] https://omeka.org/s/
[9] https://correspsearch.net/

Furthermore, CorrespSearch allows the linking of personal and place records to supported authority files. The generated file from the insertion process, however, is XML-based, with no support for RDF. Additionally, the platform does not allow for personalization or extension of the metadata model beyond the predefined fields.

To address the needs of the LKG project, Omeka S was chosen as the main software for managing the data collection process. Omeka S is a user-friendly platform featuring comprehensive documentation, an expanding user community, and supports modular extensions to enhance its capabilities. Omeka S natively supports linked data and offers the possibility of quickly developing an interface for prototypes with the goal of presenting intermediate results for stakeholders.

3. Structure of collected records

3.1. Information to be collected

In the initial stage of the work, consultations were carried out among researchers of Lem's writings and biography to better define the user requirements for the future LKG module, which will make available metadata about the collection of letters. These consultations led to the creation of a series of competency questions that were subsequently taken into account in the design of the metadata collection process. A selection of the most characteristic questions is presented as follows:

1. In which letters (to whom and when) did Lem mention *Journey 22*?
2. With whom did Lem correspond about his earnings (apart from official correspondence)?
3. With whom and when did Lem correspond on the topic of artificial intelligence?
4. When does the name John Searle first appear in the correspondence?
5. What percentage of Lem's correspondents were women? What is the figure if we consider only private correspondence excluding family?
6. In how many letters does Lem mention Germans?
7. Are there more private letters or official ones?
8. What percentage of Lem's correspondents were scientists?
9. In which letters did Lem mention his youth in Lwów?
10. In which letters does Lem use neologisms or words in a manner unique to himself?
11. With whom does Lem engage in ideological and/or political disputes in his letters?
12. In which letters does Lem describe the process of creating *Solaris*?
13. Are there letters written by Lem within 30 days of the election of Karol Wojtya as Pope that refer to this event?
14. In how many letters does Lem mention the word "cybernetics"?
15. In letters to whom does Lem mention literary theory?

During the discussion of the proposed questions, it was agreed that not all of them could be taken into account when creating the final structure of the database. In particular, questions about specific words appearing in the letters (represented above by 10. and 14.) will have to remain unanswered, as the project at this stage does not involve including the full content of the letters in the database, only metadata. Although these metadata will include information about the content, they will do so in the form of a detailed index divided into several categories, rather than through direct search within the linguistic content of particular letters.

Taking into account the competency questions proposed by the researchers, as well as the limitation of the future knowledge base, a structure of the temporary database record was developed to represent a single item in the collection of letters. This structure includes the following fields:

- sender (selected from a list of individuals or institutions)
- recipient (selected from a list of individuals or institutions)

- title of the letter (a string serving as its main identifier)
- date of writing/sending (date field – various degrees of completeness allowed)
- length of the letter in typewritten/manuscript pages (numeric field)
- storage location of the letter (Geonames identifier)
- category of the letter based on Yaznevich's classification ([10], pp. 553-338):
    1. Official correspondence: a) government-related; b) adaptors (i. filmmakers, ii. illustrators); c) agents; d) editors, journalists; e) translators; f) publishers; g) other official
    2. Private correspondence: a) Lem scholars among readers; b) academics; c) writers; d) family; e) acquaintances, friends; f) other private
- people mentioned in the letter (selected from a list of individuals)
- cultural works mentioned in the letter (selected from a list of works)
    - work in general
    - work as a specific textual version (edition or translation)
    - work as a publication
- concepts or ideas mentioned in the letter (selected from a list of concepts)
- place of writing/sending (Geonames identifier)
- geographical places mentioned in the letter (Geonames identifier)
- language of the letter
    - selected from a list of languages
    - more than one language possible, ordered by their share in the letter
    - a commonly used phrase such as *dura lex sed lex* does not count as the use of another language
    - an expression counts as another language if it has at least the status of a subordinate clause
- typewritten or handwritten (options: typewritten, handwritten, both)
- to which other letter this one is a reply
- source of the annotation (options: original, carbon copy, scan of original, scan of copy)
- whether the letter is published (options: yes, no, partially, to be checked)
- comment (e.g. letter not sent, not received, an article was based on the letter, etc.)

Additionally, lists of values for certain fields—such as people, institutions, works, languages, and concepts—should be developed during the cataloguing process. The aim is to create controlled vocabularies in order to maintain consistency in the collected data. When adding new entries to these lists, it will be necessary to gather additional information to enable their later organisation.

3.2. Lightweight ontology for data collection process

Since the LKG.letters is part of the umbrella LKG project, the collected correspondence data is supposed to be interoperable with other modules of LKG and adhere to the standards of the semantic web. LRMoo[10], a CIDOC-CRM[11] extension, has been chosen as the main ontology for modelling bibliographic information within LKG.core (see [11]). To ensure compatibility, the collected data will be modelled using an ontology aligned with LRMoo and CIDOC CRM; however, no elaborate target ontology has yet been established.[12]

Given the absence of a target ontology for the final representation of the data and the need to define a structure for linked data resulting from the ongoing collection process, we decided to opt for the development of a lightweight ontology [14]. This ontology serves as an intermediary semantic layer,

[10] https://cidoc-crm.org/lrmoo

[11] https://cidoc-crm.org/

[12] See [12, 13] for existing projects that address the semantic representation of correspondence data using semantic web technologies.

designed to be flexible, extensible, and easily implementable while remaining relatively agnostic for future integration with more complex standards such as LRMoo and CIDOC CRM. By adopting this approach, we can support incremental enrichment of the knowledge graph without committing to a final data model.

The lightweight intermediary ontology was initially created using Protégé with the *L0_Letter* class corresponding to the letters themselves and several classes that correspond to real-life objects mentioned in the letters or to some of their aspects and qualities. There are classes *L1_Person*, *L2_Group*, *L3_Place* the instances of which correspond to real-life material entities: individuals, institutions, and locations, mentioned in the correspondence (or otherwise related). By contrast, *L4_Concept* is meant to contain abstract terms, ideas, or phenomena mentioned within the letters, with *L6_Discipline* class serving as an umbrella for them.

In the case of the latter, a closed list of possible states or qualities of the letters themselves is modelled as instances of specific group of classes, such as *L6_Category*, *L7_Language*, *L8_Physicality* (determining i.a. whether the annotated letter was a carbon copy, a scanned version thereof, the original etc.) and *L9_Publication_Status* (stating whether the letter was published, or cited in a published work).

Finally, instances of *L11_Work* class refer to various works of art (notably literary works), whenever they are mentioned in the content of a specific letter. That also includes Lem's own pieces, but only if mentioned in the most general terms. By contrast, the instances of two other classes are meant to be used instead, whenever Lem's work is mentioned in more specific context: as a specific version of the work, in case when there were several, including translations (*L12_Expression*), or as a specific published edition (*L13_Manifestation* class).

In addition to classes described above, the ontology consists of properties grouped into two distinct categories: data properties, containing plain information (strings, numbers), and object properties, that can link any number of instances of specific classes together. Among the former, the majority of properties has *L0_Letter* class defined as domain, such as *Q15_comment*, *Q16_date_of_creation*, *Q17_date_of_dispatch* or *Q18_number_of_pages* (the data to be provided in this case are either strings, dates, or non-negative integers). Likewise, among the latter group, most of the properties of objects are those with *L0_Letter* defined as their domain and other classes defined as their range.

Within the ontology, a specific letter (an instance of *L0* class) can be linked to the same type of real-world entities (like individuals or places, represented as instances of *L1* and *L3* classes) using different object properties, depending on whether the object in question is performing a specific role in regard to a given letter, or rather is mentioned within. Properties *Q0_author* and *Q1_recipient* link specific letters to people or institutions performing such a role (instances of *L1* or *L2* classes), while *Q11_actor_mentioned* is used to point actors mentioned in the letter. Likewise, properties *Q3_place_of_creation* and *Q4_place_of_storage* allow a specific letter to be linked with the said locations (instances of *L3* class), while *Q14_place_mentioned* is used for locations that can be found in the content. To cover other types of mentioned entities, like works of art (instances of *L11*, *L12* or *L13* classes) or specific abstract terms and concepts (instances of *L4* or *L5*), object properties *Q12_artwork_mentioned* and *Q13_concept_mentioned* are used, respectively. Finally, a property *Q9_is_answer_to* is meant to pinpoint another instance of *L0_Letter* class, allowing the annotator to track an exact sequence of correspondence between Lem and other people.

Some object properties are meant to describe the state of a specific annotated letter, rather than its relation to real-life objects, linking *L0_Letter* instance to entities denoting some more complex qualities than plain data (as was the case with data properties), such as *Q2_category* which links to a related *L6* class instance. Likewise, object properties *Q5_physicality*, *Q6_language_first* (with optional *Q7_language_second*) and *Q8_publication_status* link to specific instances of *L8*, *L7* and *L9* classes, respectively.

For some object properties within the ontology, the domain is not a *L0_Letter* class, but rather classes corresponding to other real-life entities (actors, places, or ideas), allowing us to model some basic relations between them. Notably, object properties *Q21_homeland* and *Q22_discipline* are meant to link actors (of *L1* or *L2* class) to specific countries (of *L3* class) and fields of activity (of *L4* class), respectively. Likewise, *Q24_member_of* property allows one to link people with their parent institutions,

Q23_city_hq allows to point a city in which a specific institution is located, and *Q27_subgroup_of* allows to create a hierarchy of institutions. In addition, data properties were defined specifically to provide basic information about instances of people (*Q28_sex* for *L1* class) or places (*Q29_geoname_id* for *L3* class) mentioned throughout Lem's correspondence.

Since the ontology is designed to serve as a tool to annotate Lem's letters as such (as opposed to creating an encyclopaedia-like model of their content), the reason for including basic relations between other types of real-life entity should be clarified. Inclusion of even such a scarce amount of extra information within the ontology is, in fact, a necessity, hence competency questions mentioned above set a high standard for accuracy. Certain amount of extra information about non-letters is essential, as it allows searching through the annotated letters in a way that can guarantee the results far more complex and refined than the results otherwise obtained without such additional data being stored. For example, looking for "physics" by a future user of the database should encompass not only the mentions of the discipline within Lem's letters, but also mentions of specific concepts from that field (such as black holes) or known physicists (such as Albert Einstein). Likewise, when looking for information related to "Germany" within Lem's correspondence, a future user might want to track not only the specific mentions of current federal state of Germany (or its former historical instances), but also mentions of German citizens or institutions. Adding object properties to classes other than *L0* establishes the basis within the ontology itself for such complex searches to be conducted properly in the future.

4. Annotation workflow with Omeka S

During the annotation process, two groups of objects (items) within Omeka S must usually be created concurrently: an instance of *L0_Letter* class, corresponding to the letter being annotated, and multiple instances of certain other classes, that correspond to real-life individuals, institutions, places and concepts (classes *L1*, *L2*, *L3* and *L5*, respectively) that are mentioned within the letter being annotated or that perform a specific function in relation to it. A creation of a new object of a specific class is performed in Omeka S by utilising templates, created by us on the basis of the lightweight ontology, which consist of fields to be filled in by an annotator, with each field corresponding to a specific data property and object property, that had a given class defined as their domain at the moment of template creation. In case of an object property, a specific item must be selected from the list of items already present within the Omeka S database.

Since a desired object might not be present yet at the given time, operating two instances of Omeka S is the most efficient way of annotating letters: with two instances of Omeka S running in different tabs within the browser, one can use the "main" tab for filling in a letter template, while the "auxillary" tab can be used to add multiple non-existent objects to the database (one after another) in order to be able to select them in the "main" tab, allowing for creation of an instance of desired object property.

4.1. Adding Lem's letters to the database

Omeka S provides the user with the option of adding multiple objects to a given field within the template at once, which makes "group creation" an efficient way of dealing with multiple objects to be annotated, whenever those are not yet present in the database. This is very often the case with the content of Lem's letters, which can be very "dense" in terms of the quantity of items to be added, sometimes forcing an annotator to include several dozens of new objects per letter. An exception from that rule is presently regarded as a future solution, meant for potential annotation process suitable for unpublished sets of letters marked as "sensitive" by the Lem's heirs, in which case certain objects (or classes of objects) might be deliberately omitted during the annotation process, or annotation of content might be omitted entirely, thus shortening the time needed to collect the basic information. However, in regular cases, the annotation of content is the most time-consuming part of the process.

In contrast, collecting the basic information within the letter template is usually faster as it does not require the creation of new objects. Such is the case with data properties mentioned earlier (such as number of pages, dates of creation and dispatch), and with those of object properties that

require selecting instances of classes that represent real-life properties of the letter itself, as opposed to representing other real-life entities. Properties such as these (*Q2*, *Q6*, *Q7*, *Q5*, *Q8*) require selection from the closed list, since they refer to standarised content, like the letter's category, language, physicality or publication status (instances of corresponding *L6*, *L7*, *L8* and *L9* classes, that are predefined). Defining object properties for a specific letter that does not relate to its content is usually also less time consuming, as it rarely involves adding new objects, hence those to be pointed by the user are, in this case, repeatable: for a "set" of letters between Lem and another individual, the *Q0* and *Q1* ("author" and "recipient" fields within the template) point to the fixed pair of individuals, while *Q3* and *Q4* properties ("place of creation" and "place of storage") are usually fixed, since the majority of Lem's letters were written at his place of residence in Kraków (and are presently stored within the archive still located within). Due to this repeatability, the part of the annotation process that covers collecting basic information alone can be done efficiently using Omeka S, as opposed to a much more time-consuming process of annotating letters' content. At the present stage, the latter provides a challenge for an annotator, and its partial automatisation seems to be a desirable solution to be looked for in the future.

4.2. Adding objects other than letters to the database

The addition of objects other than letters is usually less time consuming, although it requires an external source of information, which can be problematic in regard to often obscure individuals or concepts. Although the list of names mentioned in Lem's works was already prepared by Yaznevich ([10], pp. 606–722), it lacks information on gender, nationality, and profession (properties *Q28*, *Q21* and *Q22*). Moreover, information about the locations of institutions and their hierarchies (properties *Q23* and *Q27*) needs to be provided; the exact place of a newly mentioned concept within the multilayered structure of abstract terms (properties *Q25* or *Q26*), based on the already present multilayered OECD classification for scientific disciplines, must also be manually entered into the database.

The annotator must not only be able to decide when to annotate a reference to expression or manifestation (classes *L12* and *L13*), rather then a "regular" reference (to *L11* object), but also—in the case of the former—must be able to establish precisely which version of a given work, or which edition, is being referenced. The letter itself would very rarely provide necessary information, so the external database of various versions and editions of Lem's works must be used for reference.

5. Conclusion

Once the process of cataloguing Lem's correspondence is completed, the data collected within the lightweight ontology structure in Omeka S will be exported and automatically integrated into the main LKG ontology, specifically into its LKG.letters module. This will require the prior development of the target ontology, which must take into account the constraints imposed by the structure of the gathered metadata. The ultimate goal of the project is to create a knowledge base encompassing all of Lem's texts, within which it will be possible to formulate queries that cover both publications and correspondence.

Acknowledgments

This publication was funded by a flagship project "CHExRISH: Cultural Heritage Exploration and Retrieval with Intelligent Systems at Jagiellonian University" under the Strategic Programme Excellence Initiative at Jagiellonian University.

The research for this publication has been supported by a grant from the Priority Research Area DigiWorld under the Strategic Programme Excellence Initiative at Jagiellonian University.

Declaration on Generative AI

During the preparation of this work, the authors used GPT-4o, GPT-4o-mini and DeepSeek-V3 in order to: Grammar and spelling check, Paraphrase and reword. After using these services, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] L. Verma, N. Kumar, Comparative analysis of open source digital library softwares: A case study, DESIDOC Journal of Library & Information Technology 38 (2018) 361.

[2] C.-M. Chen, Y.-H. Chen, C.-L. Hong, C.-M. Liao, C.-M. Huang, Developing a taiwan library history digital library with reader knowledge archiving and sharing mechanisms based on the dspace platform, The Electronic Library 30 (2012) 426–442. URL: https://doi.org/10.1108/02640471211241681. doi:10.1108/02640471211241681.

[3] M. Kurtz, Dublin core, dspace, and a brief analysis of three university repositories, Information Technology and Libraries 29 (2010) 40–46. URL: https://ital.corejournals.org/index.php/ital/article/view/3157. doi:10.6017/ital.v29i1.3157.

[4] G. N. Pramudyo, M. R. Hendrawan, Metadata interoperability for institutional repositories: A case study in malang city academic libraries, in: E. Ishita, N. L. S. Pang, L. Zhou (Eds.), Digital Libraries at Times of Massive Societal Transition, Springer International Publishing, Cham, 2020, pp. 355–363.

[5] D. Bainbridge, I. H. Witten, A renewed look at greenstone: Entering the third decade, IPSI Transactions on Internet Research 16 (2020) 3–11.

[6] R. A. Martinez, Omeka s como alternativa para el desarrollo de colecciones digitales y proyectos de humanidades digitales, BiD 48 (2022). doi:10.1344/BID2022.48.05.

[7] W.-B. Yang, J.-F. Jan, T.-J. Wang, Y.-C. Lu, C.-L. Kuo, Y.-N. Yen, Digital interpretation and presentation for monuments built by arches - take kinmen area heritage as an example, in: M. Ioannides, E. Fink, R. Brumana, P. Patias, A. Doulamis, J. Martins, M. Wallace (Eds.), Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection, Springer International Publishing, Cham, 2018, pp. 366–375.

[8] P. Popović, M. Škorić, B. Rujević, The use of the omeka semantic platform for the development of the university of belgrade, faculty of mining and geology digital repository, Infotheca 20 (2020) 136–148. doi:10.18485/infotheca.2020.20.1_2.9.

[9] A. Bikakis, B. Markhoff, A. Mosca, S. Jean, E. Hyvönen, O. Bruneau, N. Lasolle, J. Lieber, E. Nauer, S. Pavlova, L. Rollet, A. Bikakis, E. Hyvonen, S. Jean, B. Markhoff, A. Mosca, Applying and developing semantic web technologies for exploiting a corpus in history of science: The case study of the henri poincaré correspondence, Semant. Web 12 (2021) 359–378. URL: https://doi.org/10.3233/SW-200400. doi:10.3233/SW-200400.

[10] V. Yaznevich, Stanislaw Lem Non Fiction. Bibliography and Personalities, Medisont, Minsk, 2021.

[11] L. do Valle Miranda, J. Gomułka, S. P. Kukulak, K. Kutt, G. J. Nalepa, LRMoo as the Conceptual Model for the Lem Knowledge Graph, in: Proceedings of the Second International Workshop on Semantic Digital Humanities, co-located with the Extended Semantic Web Conference (ESWC 2025), Heraklion, Crete, Greece, 2025. Forthcoming.

[12] E. Hyvönen, P. Leskinen, J. Tuominen, Lettersampo–historical letters on the semantic web: A framework and its application to publishing and using epistolary data, J. Comput. Cult. Herit. 16 (2023). URL: https://doi.org/10.1145/3569372. doi:10.1145/3569372.

[13] L. Fath, J. Strobel, Semantische erschließung frühromantischer korrespondenzen. wissensrepräsentation als element einer ›semantic scholarly digital edition‹, Zeitschrift für digitale Geisteswissenschaften 9 (2024). URL: https://www.zfdg.de/article/view/2024_006. doi:10.17175/2024_006, hTML / XML / PDF.

[14] F. Giunchiglia, I. Zaihrayeu, Lightweight Ontologies, Springer US, Boston, MA, 2009, pp. 1613–1619.

URL: https://doi.org/10.1007/978-0-387-39940-9_1314. doi:10.1007/978-0-387-39940-9_1314.