

# Leveraging Large Language Models for Processing and Evaluating FAIR Digital Objects

Nicolas Blumenröhr<sup>1,\*</sup>, Felix Kraus<sup>1</sup>

<sup>1</sup>Karlsruhe Institute of Technology, Scientific Computing Center, Hermann-von-Helmholtz Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

## Abstract

This paper explores the potential of generative AI, i.e., Large Language Models, for processing FAIR Digital Objects and enhancing their reusability assessment. By leveraging ChatGPT's o3-mini model, the authors propose a lightweight workflow of prompt-based operations that automates the resolution of persistent Handle identifiers, extracts the corresponding metadata into structured key-value pairs, and evaluates the content. For demonstration, we experimented with a detailed case study. We processed FAIR Digital Objects from different domains using a series of targeted prompts, requesting ChatGPT to interpret the syntactic and semantic features of the corresponding metadata elements. Through a set of competency questions, the preliminary results reveal that while ChatGPT successfully indexes and analyzes most of the metadata content, challenges remain in the semantic interpretation of record elements, fully automated resolving of non-URL based references, and targeted prompting for optimizing the information content of the outputs. The findings underscore the potential of generative AI-assisted methods to reduce manual effort, improve data reuse, and ultimately foster data infrastructures in the spirit of FAIR Principles.

## Keywords

FAIR Digital Objects, FAIR Principles, LLMs, ChatGPT

## 1. Introduction

The adoption and implementation of FAIR Digital Objects (FDOs) as a strategy to realize FAIR Principles has potential to enhance interoperability and data reuse, facilitating the work for scientists by automation. Despite substantial advancements in FAIR-compliant data infrastructures, inspecting, analyzing, and effectively evaluating the reuse potential of FDOs remains challenging, mainly due to a lack of associated machine-actionable decisions [1]. One problem in these aspects is the interpretation of the typed information in the FDO that typically needs to be realized using proper operations. These operations must be defined, implemented, and associated with the FDO types, which in turn are also often insufficiently specified, or require a complex service infrastructure. However, each content of an FDO is essentially text-based and often contains references to more text-based information, such as vocabularies or landing pages. Although the initial effort to provide machine-interpretable information for FDOs aimed to overcome the limitations of evaluating plain text content, large language models (LLMs), which are a type of generative AI, have recently offered novel perspectives on processing text-based information. [2].

ChatGPT [3] is currently one of the most recognized LLMs and may constitute a viable lightweight approach for processing the contents of FDOs without the additional overhead imposed by conventional processing methods.

In this work, we demonstrate the potential of using LLMs as a environment to process and evaluate metadata contained in FDOs. To illustrate this, we present a case study that involves two FDOs from the domains of digital humanities and energy research; we leverage ChatGPT's o3-mini model, design a

---

2nd International Workshop on Natural Scientific Language Processing and Research Knowledge Graphs (NSLP 2025), co-located with ESWC 2025, June 01–02, 2025, Portorož, Slovenia

\*Corresponding author.

✉ nicolas.blumenroehr@kit.edu (N. Blumenröhr); felix.kraus@kit.edu (F. Kraus)

ORCID 0009-0007-0235-4995 (N. Blumenröhr); 0000-0002-2102-4170 (F. Kraus)



© 2025 Copyright © 2025 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

sequence of LLM prompts as a workflow, and evaluate the results using a set of competency questions.

## 2. Background and Related Work

The concept of Digital Objects was first introduced by [4] and later identified as being consistent with the requirements of the FAIR principles [5], constituting a strategy for their implementation. This gave rise to the FDO concept [6], with the potential to facilitate broader abstraction for interoperability in data management [7]. Essentially, each FDO is a persistent, high-level representation of a digital resource. Per definition, each FDO is assigned a Persistent Identifier (PID) that is registered at the Handle Registry and resolves to an information record that constitutes the essential metadata of the represented digital resource. However, as pointed out in [1], the machine-actionable capabilities of FDOs depends on the provision of a framework that makes use of object-associated operations. Such conventional FDO operation environments require a lot of overhead due to the standardization of types, semantics, and the provision of a service infrastructure with interfaces [8]. Examples of such FDO operation environments are described in [9, 10].

LLMs offer advanced capabilities in text understanding, generation, summarization, and semantic interpretation, making them particularly suited for tasks involving metadata analysis and indexing [11]. For instance, GPT-based tools have been successfully applied in contexts such as automated document summarization, metadata enrichment, and semantic classification of data resources [12, 13, 14]. Such tools have the potential to significantly reduce manual workloads, increased metadata accuracy, and facilitate deeper analytical insights. However, to the best of our knowledge, there currently exists no systematic exploration and evaluation of GPT-driven methods for FDO analysis.

## 3. Case Study

We consider a case study for using ChatGPT to analyze and evaluate the content of FDOs that represent digital resources from different domains. First, we are going to detail the structure of the metadata in the FDO and its characteristics that makes it suitable for processing in a GPT model. We then propose a prototypical workflow for prompting ChatGPT, and to evaluate its outputs by a set of competency questions.

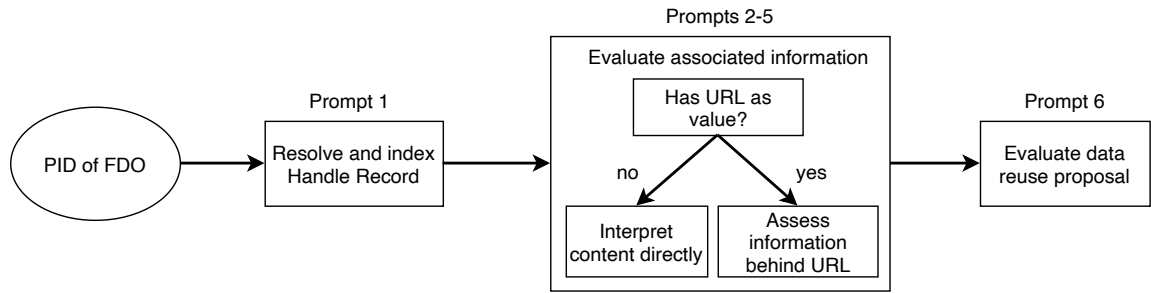
### 3.1. FAIR Digital Objects

We picked one FDO that was created as part of a project in energy research [15], representing a set of drone images. Another FDO represents a controlled vocabulary benchmarking dataset in the digital humanities [16]. Both FDOs were built on the basis of a Kernel Information Profile [17] that contains a set of attributes that adhere to PID-Information Types (PITs) [18], following the FDO data model described in [1]. The PITs form the core of the FDO type system and can be modeled hierarchically with a finite combination of PITs and Basic PITs down to the elementary level of JSON types for automated schema extraction. Consequently, each FDO contains an information record of typed key-value pairs that is persistently registered at the Handle Registry and can be resolved using the FDO's PID<sup>1,2</sup>. Whilst the type of each value is validated through the type system of the corresponding PITs, the plain text of each value is directly assessable from the Handle interface. For a human reader, interpreting these text-based values is generally straightforward, but for conventional machine-actionable procedures, it is crucial that operations for the underlying type system are provided. A main aspect of operations is thereby to leverage an FDO's entity-relationship characteristics, i.e., its relations to other entities on the web using URLs, e.g. a landing page, or Handle PIDs that typically point to FDOs that represent related (meta)data.

---

<sup>1</sup><https://hdl.handle.net/21.11152/6858a0b5-cc60-40e9-afef-8c2dd8b35e8e?noredirect>

<sup>2</sup><https://hdl.handle.net/21.11152/a3f19b32-4550-40bb-9f69-b8ffd4f6d0ea?noredirect>



**Figure 1:** The ChatGPT prompting workflow.

### 3.2. Proposed Workflow

To process the example FDOs, we used the GPT-o3-mini model that has advanced reasoning capabilities. We performed our experiments on the 13-03-2025. Note that we did not explicitly explain to the GPT model the theory behind FDOs or other related concepts, such as metadata standards or energy research methodologies. The reference FDO and any of its content has up to this point not been explicitly provided to the GPT model. We then formulated and executed a series of prompts in a sequential workflow (cf. fig. 1). The exact prompt texts are listed in listing 1.

Listing 1: The ChatGPT FDO operation prompts.

Prompt 1: "Here is the URL to a Handle record. Resolve and index this record.  
[https://hdl.handle.net/Handle\\_PID](https://hdl.handle.net/Handle_PID)"

Prompt 2: "Evaluate the information associated with these key-value pairs.  
For each value that references an external resource, e.g. via a URL, evaluate  
also the referenced information."

Prompt 3: "Resolve the external URL links you have found and report  
on the content."

Prompt 4: "Resolve and index the records of the handle references you have found  
by adding them to the following URL: <https://hdl.handle.net/>"

Prompt 5: "Resolve these handle URLs and index their records."

Prompt 6: "Based on your previous assessment, how would you say the described  
data can be reused in which use cases?"

The entire conversation that constitutes the proposed prompt workflow including all responses from ChatGPT can be found at [19].

### 3.3. Competency Questions

To qualitatively evaluate the result of our prompt workflow, we formulated a set of competency questions that relate to the typical expectations towards an FDO operation's capabilities, and evaluated the plausibility of the generated output in the context of these questions:

- **Q1 - was the Handle record correctly resolved, indexed and its content listed by the GPT-model?:** Yes, for both FDOs.
- **Q2 - were the given key-value pairs correctly interpreted with respect to the syntactic and semantic specification of the corresponding PITs?:** partly, most of the key-value pairs were correctly analyzed and described, whilst others were only vaguely described, not catching relevant aspects. No false interpretations were observed though.
- **Q3 - were external resources referenced by specific values via URLs and PIDs analyzed and accurately described?:** partly, external resources referenced via URLs were resolved and analyzed, providing accurate and useful information. Referenced PIDs were added to the base URL, but could not be directly resolved and analyzed.
- **Q4 - were the suggested reuse cases reasonable?:** Yes, most of the suggested reuse cases and proposed subsequent steps either complement the original use case or provide useful inspiration for alternative applications of the digital resources, although from a very high-level perspective.

## 4. Results and Discussion

Our case study underlines that the GPT-o3-mini model is capable of resolving and indexing a Handle record of an FDO when provided with the URL of the corresponding PID, which was a crucial baseline for any additional investigations. When asked for the provision of a structured listing of the record contents, we received the correct key-value pairs. It is important to point out that the advanced reasoning of this GPT model seems to be a crucial aspect at this stage, because when we tested other GPT models, such as the GPT 4.5 model, we obtained a text block that was missing key elements instead of the correctly structured record content.

The GPT model successfully interpreted the semantics of most values without needing explicit knowledge of the associated PIT keys. Since these PITs are at this stage not widely recognized, we assume that the GPT model inferred the meaning of the associated values purely on the provided information record. Whilst this interpretation was very accurate and complete for certain values that can be easily inferred based on the value text, e.g. the date-time, others were accurate but rather sparse, obviously due to the absence of knowing the corresponding PIT specification and a lack of context, e.g. the identifier of an related FDO.

For each value that constitutes a URL to a related web entity, the GPT model was able to resolve the URL and provide additional information of the underlying content which was accurate and useful, e.g. the assessment of associated *UNESCO Thesaurus* concepts<sup>3,4</sup> for the *topic* PIT. This shows the capability of the GPT model to harvest web content that it did not receive directly from the client but through the FDO's information record. A separate prompt for this specific task was required though (cf. prompt 3). Further elaboration on these contents and harvesting of additional websites that may be discoverable through Linked Data principles were not further explored. Whilst the harvesting of contents was possible for values that contain URLs, those values that contain PIDs were not resolved, also not when the GPT model was explicitly asked to do so. In order to yield information on these FDOs,

<sup>3</sup><http://vocabularies.unesco.org/thesaurus/concept10081>

<sup>4</sup><http://vocabularies.unesco.org/thesaurus/concept1557>

their PID URLs must be prompted manually. Therefore, it seems that the PID-triples that are inherently constellated by FDOs as described in [1] can not be automatically discovered and analyzed by the used GPT model at this stage. However, we want to point out that we did not perform exhaustive prompt engineering, and cannot confidently exclude that this could not be achieved by bypassing guidelines. A general observation was also that the more elements are contained in the record, the less was elaborated on each by ChatGPT. Separate prompts for analyzing each element could most likely increase their information content.

With respect to the suggested reuse of the provided FDO, the GPT model gave reasonable answers that align with the original use case and considered the license conditions. Typically, these must be inferred by the user, or a programmed digital client, taking into account the results of earlier performed operations. Again, we did not further elaborate on these outputs to receive more concrete suggestions for further steps or specific aspects. Whilst useful and important information was captured and provided by ChatGPT, the answers contained a lot of redundant information and were often described with verbose phrases. This could be substantially improved by proper prompt-engineering.

During this study, the GPT model only processed the text-based metadata in the FDO's information record and interpreted its content, but did not apply any specific operations to yield a modified version of the contents. Depending on the requirements, certain operations can be already performed by current GPT models, e.g. numerical operations on the date-time. Whilst this is possible using the metadata text, operations on the referenced bit sequence of the represented digital resource may be more challenging to be accomplished. However, the capabilities of GPT models to identify potential use cases, and writing executable code pieces could in the long-term result in dynamic workflows that analyze the given information and create proper operations on the fly that could be executed in a robust and secure pipeline.

## 5. Conclusions and Future Work

In this work, we explored the feasibility of using LLMs, i.e., ChatGPT's o3-mini model, for resolving, processing, and evaluating the reuse potential of digital resources that are represented as FAIR Digital Objects. The proposed case study demonstrates that generative AI can successfully resolve persistent Handle identifiers, extract and structure metadata records into key-value pairs, and perform content analysis by incorporating external data sources. These results indicate that AI-assisted methods can enhance the efficiency of FDO metadata processing, reduce manual effort, and improve the overall reusability of digital resources. Therefore, we see a great potential in the combination of these technologies, where FDOs are a fundament for persistent, reliable and standardized information entities in the spirit of FAIR Principles, and LLMs can be used as a lightweight approach to operate on these entities.

Despite the promising outcomes, several limitations remain. The reliance on AI-generated content raises concerns about potential biases and inaccuracies, especially when interpreting complex metadata relationships. Future studies should concentrate on broader case studies across various disciplines to investigate the possible extent of processing FDO information, benchmarking different LLM models, considering prompt engineering, and evaluating the reliability and robustness of the generated outputs. There should also be a more detailed analysis of the differences to- and compatibility with a conventional operation system based on FDO types. Our work provides the first effort in this direction.

## Acknowledgments

This project is funded by the Helmholtz Metadata Collaboration Platform (HMC), and supported by the research program "Engineering Digital Futures" of the Helmholtz Association of German Research Centers.

# Declaration on Generative AI

The authors have not employed any Generative AI tools outside of the case study described in this work.

## References

- [1] Blumenröhr, Nicolas, Ost, Philipp-Joachim, Kraus, Felix, Streit, Achim, FAIR Digital Objects for the Realization of Globally Aligned Data Spaces, in: IEEE International Conference on Big Data (BigData), IEEE Xplore, Washington, DC, USA, 14-18 December 2024, 2025, pp. 374–383.
- [2] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, D. Amodei, Language Models are Few-Shot Learners, 2020. doi:10.48550/arXiv.2005.14165, arXiv:2005.14165 [cs].
- [3] OpenAI, J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Al-tenschmidt, S. Altman, S. Anadkat, R. Avila, I. Babuschkin, S. Balaji, V. Balcom, P. Baltescu, H. Bao, M. Bavarian, J. Belgum, I. Bello, J. Berdine, G. Bernadett-Shapiro, C. Berner, L. Bogdonoff, O. Boiko, M. Boyd, A.-L. Brakman, G. Brockman, T. Brooks, M. Brundage, K. Button, T. Cai, R. Campbell, A. Cann, B. Carey, C. Carlson, R. Carmichael, B. Chan, C. Chang, F. Chantzis, D. Chen, S. Chen, R. Chen, J. Chen, M. Chen, B. Chess, C. Cho, C. Chu, H. W. Chung, D. Cummings, J. Currier, Y. Dai, C. Decareaux, T. Degry, N. Deutsch, D. Deville, A. Dhar, D. Dohan, S. Dowling, S. Dunning, A. Ecoffet, A. Eleti, T. Eloundou, D. Farhi, L. Fedus, N. Felix, S. P. Fishman, J. Forte, I. Fulford, L. Gao, E. Georges, C. Gibson, V. Goel, T. Gogineni, G. Goh, R. Gontijo-Lopes, J. Gordon, M. Grafstein, S. Gray, R. Greene, J. Gross, S. S. Gu, Y. Guo, C. Hallacy, J. Han, J. Harris, Y. He, M. Heaton, J. Heidecke, C. Hesse, A. Hickey, W. Hickey, P. Hoeschele, B. Houghton, K. Hsu, S. Hu, X. Hu, J. Huizinga, S. Jain, S. Jain, J. Jang, A. Jiang, R. Jiang, H. Jin, D. Jin, S. Jomoto, B. Jonn, H. Jun, T. Kaftan, Kaiser, A. Kamali, I. Kanitscheider, N. S. Keskar, T. Khan, L. Kilpatrick, J. W. Kim, C. Kim, Y. Kim, J. H. Kirchner, J. Kiros, M. Knight, D. Kokotajlo, Kondraciuk, A. Kondrich, A. Konstantinidis, K. Kosic, G. Krueger, V. Kuo, M. Lampe, I. Lan, T. Lee, J. Leike, J. Leung, D. Levy, C. M. Li, R. Lim, M. Lin, S. Lin, M. Litwin, T. Lopez, R. Lowe, P. Lue, A. Makanju, K. Malfacini, S. Manning, T. Markov, Y. Markovski, B. Martin, K. Mayer, A. Mayne, B. McGrew, S. M. McKinney, C. McLeavey, P. McMillan, J. McNeil, D. Medina, A. Mehta, J. Menick, L. Metz, A. Mishchenko, P. Mishkin, V. Monaco, E. Morikawa, D. Mossing, T. Mu, M. Murati, O. Murk, D. Mély, A. Nair, R. Nakano, R. Nayak, A. Neelakantan, R. Ngo, H. Noh, L. Ouyang, C. O’Keefe, J. Pachocki, A. Paino, J. Palermo, A. Pantuliano, G. Parascandolo, J. Parish, E. Parparita, A. Passos, M. Pavlov, A. Peng, A. Perelman, F. d. A. B. Peres, M. Petrov, H. P. d. O. Pinto, Michael, Pokorny, M. Pokrass, V. H. Pong, T. Powell, A. Power, B. Power, E. Proehl, R. Puri, A. Radford, J. Rae, A. Ramesh, C. Raymond, F. Real, K. Rimbach, C. Ross, B. Rotsted, H. Roussez, N. Ryder, M. Saltarelli, T. Sanders, S. Santurkar, G. Sastry, H. Schmidt, D. Schnurr, J. Schulman, D. Selsam, K. Sheppard, T. Sherbakov, J. Shieh, S. Shoker, P. Shyam, S. Sidor, E. Sigler, M. Simens, J. Sitkin, K. Slama, I. Sohl, B. Sokolowsky, Y. Song, N. Staudacher, F. P. Such, N. Summers, I. Sutskever, J. Tang, N. Tezak, M. B. Thompson, P. Tillet, A. Tootoonchian, E. Tseng, P. Tuggle, N. Turley, J. Tworek, J. F. C. Uribe, A. Vallone, A. Vijayvergiya, C. Voss, C. Wainwright, J. J. Wang, A. Wang, B. Wang, J. Ward, J. Wei, C. J. Weinmann, A. Welihinda, P. Welinder, J. Weng, L. Weng, M. Wiethoff, D. Willner, C. Winter, S. Wolrich, H. Wong, L. Workman, S. Wu, J. Wu, M. Wu, K. Xiao, T. Xu, S. Yoo, K. Yu, Q. Yuan, W. Zaremba, R. Zellers, C. Zhang, M. Zhang, S. Zhao, T. Zheng, J. Zhuang, W. Zhuk, B. Zoph, GPT-4 Technical Report, 2024. doi:10.48550/arXiv.2303.08774, arXiv:2303.08774 [cs].
- [4] R. Kahn, R. Wilensky, A framework for distributed digital object services, International Journal on Digital Libraries 6 (2006) 115–123. doi:10.1007/s00799-005-0128-x.
- [5] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg,

- J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. 't Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, B. Mons, The FAIR Guiding Principles for scientific data management and stewardship, *Scientific Data* 3 (2016) 160018. doi:10.1038/sdata.2016.18.
- [6] E. Schultes, P. Wittenburg, FAIR Principles and Digital Objects: Accelerating Convergence on a Data Infrastructure, in: Y. Manolopoulos, S. Stupnikov (Eds.), *Data Analytics and Management in Data Intensive Domains*, Springer International Publishing, Cham, 2019, pp. 3–16.
- [7] P. Wittenburg, G. O. Strawn, Digital Objects as Drivers towards Convergence in Data Infrastructures, 2019. doi:<https://doi.org/10.23728/B2SHARE.B605D85809CA45679B110719B6C6CB11>.
- [8] S. Soiland-Reyes, C. Goble, P. Groth, Evaluating FAIR Digital Object and Linked Data as distributed object systems 10 (2024) e1781. doi:10.7717/peerj-cs.1781.
- [9] N. Blumenröhr, R. Aversa, From implementation to application: FAIR digital objects for training data composition, *Research Ideas and Outcomes* 9 (2023) e108706. doi:10.3897/rio.9.e108706.
- [10] S. Islam, J. Beach, E. Ellwood, J. Fortes, L. Lannom, G. Nelson, B. Plale, Assessing the FAIR Digital Object Framework for Global Biodiversity Research, *Research Ideas and Outcomes* 9 (2023). doi:10.3897/rio.9.e108808.
- [11] H. Song, S. Bethard, A. Thomer, Metadata Enhancement Using Large Language Models, in: T. Ghosal, A. Singh, A. Waard, P. Mayr, A. Naik, O. Weller, Y. Lee, S. Shen, Y. Qin (Eds.), *Proceedings of the Fourth Workshop on Scholarly Document Processing (SDP 2024)*, Association for Computational Linguistics, Bangkok, Thailand, 2024, pp. 145–154. URL: <https://aclanthology.org/2024.sdp-1.14/>.
- [12] M. Martorana, T. Kuhn, L. Stork, J. v. Ossenbruggen, Zero-Shot Topic Classification of Column Headers: Leveraging LLMs for Metadata Enrichment, 2024. doi:10.48550/arXiv.2403.00884, arXiv:2403.00884 [cs].
- [13] S. S. Sundaram, B. Solomon, A. Khatri, A. Laumas, P. Khatri, M. A. Musen, Use of a Structured Knowledge Base Enhances Metadata Curation by Large Language Models, 2025. doi:10.48550/arXiv.2404.05893, arXiv:2404.05893 [cs].
- [14] H. Shakil, A. M. Mahi, P. Nguyen, Z. Ortiz, M. T. Mardini, Evaluating Text Summaries Generated by Large Language Models Using OpenAI's GPT, 2024. doi:10.48550/arXiv.2405.04053, arXiv:2405.04053 [cs].
- [15] Z. Mayer, J. Kahn, M. Götz, Y. Hou, T. Beiersdörfer, N. Blumenröhr, R. Volk, A. Streit, F. Schultmann, Thermal Bridges on Building Rooftops, *Scientific Data* 10 (2023) 268. doi:10.1038/s41597-023-02140-z.
- [16] F. Kraus, N. Blumenröhr, G. Götzelmann, D. Tonne, A. Streit, A Gold Standard Benchmark Dataset for Digital Humanities, in: E. Jiménez-Ruiz, O. Hassanzadeh, C. Trojahn, S. Hertling, H. Li, P. Shvaiko, J. Euzenat (Eds.), *Proceedings of the 19th International Workshop on Ontology Matching*, volume 3897 of *CEUR Workshop Proceedings*, CEUR, Baltimore, USA, 2024, pp. 1–17. doi:10.5445/IR/1000178023.
- [17] T. Weigel, B. Plale, M. Parsons, G. Zhou, Y. Luo, U. Schwardmann, R. Quick, M. Hellström, K. Kurakawa, RDA Recommendation on PID Kernel Information FINAL, 2019. URL: <https://zenodo.org/records/3581275>.
- [18] U. Schwardmann, Automated schema extraction for PID information types, in: *2016 IEEE International Conference on Big Data (Big Data)*, 2016, pp. 3036–3044. doi:10.1109/BigData.2016.7840957.
- [19] N. Blumenröhr, ChatGPT Prompts on FAIR Digital Objects, 2025. doi:10.5281/zenodo.15056647.