# Systematically modeling and extracting bibliographic metadata of power grid standard documents with LLMs

*Guowei Chen, Wei Xie, Yanan Liu, Xiaoqun Yuan, and Liang Zhao*

## Abstract

**Introduction.** This study addresses the critical need for systematic bibliographic metadata representation and extraction from power grid standard documents, essential for operational efficiency and knowledge management in the power industry.

**Method.** We developed a two-stage methodology utilizing large language models (LLMs) for extracting bibliographic metadata. The first stage involves constructing state grid-oriented instructions for the LLM, and the second stage includes a trustworthiness estimation to ensure the reliability of the extracted metadata.

**Analysis.** Experiments were conducted using 96 state grid PDF samples to test the accuracy of metadata extraction. The performance of different LLMs was evaluated using single and multiple instructions.

**Results.** The results showed over 70% accuracy across all models, with GPT-4 achieving the highest accuracy of 84%. Multiple instructions outperformed single instructions, highlighting the effectiveness of our approach.

**Conclusion(s).** This study demonstrates the promising potential by LLM for data management in the power grid field, with the trustworthiness estimation mechanism significantly enhancing the reliability of the data extracted.

# Introduction

Power grid standard documents in actual production provides a set of standardized operation and management processes for power grid enterprises (China Power, 2024). Through unified regulations, standardization helps optimize resource allocation, risk control and decision-making support, thus enhancing overall operational efficiency and service quality (Gal & Rubinfeld, 2019; Reyes et al.,2023).

In the era of burgeoning data and information systems, metadata plays a pivotal role in organizing, navigating, and understanding the plethora of massive standard documentation resources (Zeng & Qin, 2020; Riley, 2017; Baca, 2016). Particularly, bibliographic metadata of the standard documents, which includes information about documents such as titles, authors, and publication details, is crucial for effective knowledge management and retrieval (Liu et al.,2022; Reyes et al.,2023).

However, systematically and effectively extract structured bibliographic metadata from a variety of unstructured standard documents (e.g., in PDF format) is a non-trivial task. Existing methods of knowledge extraction from PDFs (Büchter et al., 2020) typically uses OCR technique to convert these documents into editable text (He ,2020; Karthick et al.,2019;Bartz et al.,2017), followed by manual rule-based systems (Proctor et al.,2019) (e.g., by regular expressions; Chapman et al., 2017) or the construction of machine learning models (Fanni et al.,2023; Kang et al., 2020; Chowdhary& Chowdhary, 2020) to recognize the required information. Such methods generally incur high costs and require model readjustment when the set of standard documents change, lacking a convenient and universal framework.

Ovbiously, there are two major challenges that need addressing. Conceptually, a unified and structured representation framework of bibliographic metadata specifically for power grid standard documents is still ambiguous. Technically, a unified and low-cost knowledge extraction pipeline needs to be developed.

The advent of large language models (LLMs) has marked a significant leap forward in the field of knowledge extraction, especially in areas where traditional data labelling is limited or costly. Efforts have been made to harness the abilities of LLMs for generative information extraction tasks, (e.g. Xu et al., 2023; Sivarajkumar et al., 2023). Researchers have been investigating the usage of LLMs in both zero-shot and in-context learning settings to tackle the problem of extracting procedures from unstructured text in an incremental question-answering fashion (Dagdelen et al., 2024; Kabongo & D'Souza, 2024). For example, Papaluca et al. (2023) developed a dynamic pipeline for knowledge graph triplet extraction in zero-shot and few-shot scenarios by using contextual data from a knowledge base. Xue et al. (2024) introduced AutoRE, a method for relationship extraction that goes beyond traditional sentence-level analysis to document-level analysis. Through simple prompt-based strategies, Shao et al. (2023) explored using pre-trained LLMs for extracting astronomical knowledge entities from astrophysical journal articles. Zhang et al. (2024) proposed a zero-shot learning framework to enable LLMs to assimilate knowledge even without direct training on specific datasets.

Inspired by the promising capabilities of LLMs, in this study, we propose a comprehensive pipeline for both modelling and extracting the bibliographic  metadata of power grid standard documents, marking a pivotal step from conceptual representation to application realization within the domain of power grid standard documentation. As shown in Figure 1, we adopted a concise and efficient pipeline to realize metadata extraction from power network standard documents, which consists of two modules, addressing the two major challenges, respectively. Firstly, we conceptually construct a unified and structured framework of bibliographic metadata specifically for power grid standard documents, a step that is the basis of the entire process and involves determining what key information should be extracted from the documents. Secondly, a two-stage methodology

utilizing LLMs for extracting bibliographic metadata is designed, including instruction design upon LLMs and trustworthiness estimation to refine the reliability.
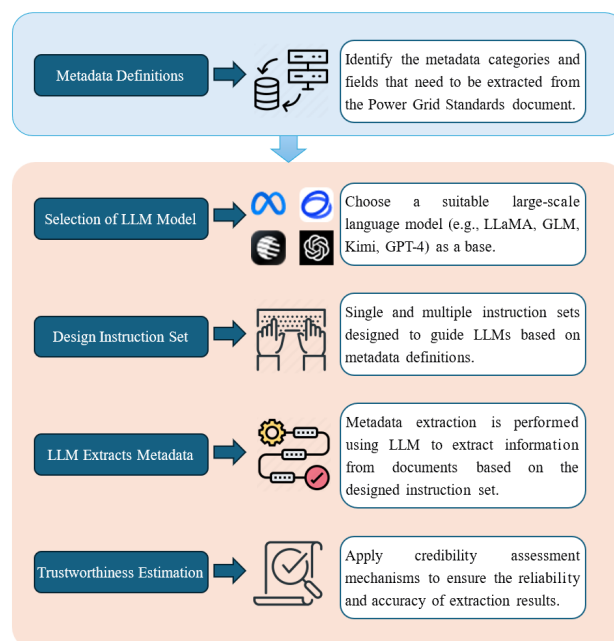


**Figure 1.** Research framework

# Modeling bibliographic metadata of power grid standard documents

## Conceptualization of bibliographic metadata

It is necessary to construct a unified and structured conceptualizaiton that categorizes metadata into distinct yet interconnected components. According to the national standard GB/T 22373-2021 'Standard Document Metadata' (http://c.gb688.cn/bzgk/gb/showGb?type=online&hcno=174FB61306FB900BFD86FF84C3BD12F 7) of China, which is designed to standardize the metadata of general standard documents (Liu et al.,2022), we summarize and design the conceptual framwork of bibliographic metadata by involving the specific attributes in the field of power grid, capturing the essence of the metadata systematically and hierarchically. Totally, 24 bibliographic metadata items are conceptualized, organizd in four categories including basic information, semantic information, correlated information, as well as property information. Figure 2 illustrates hierarchical diagram for clarity.
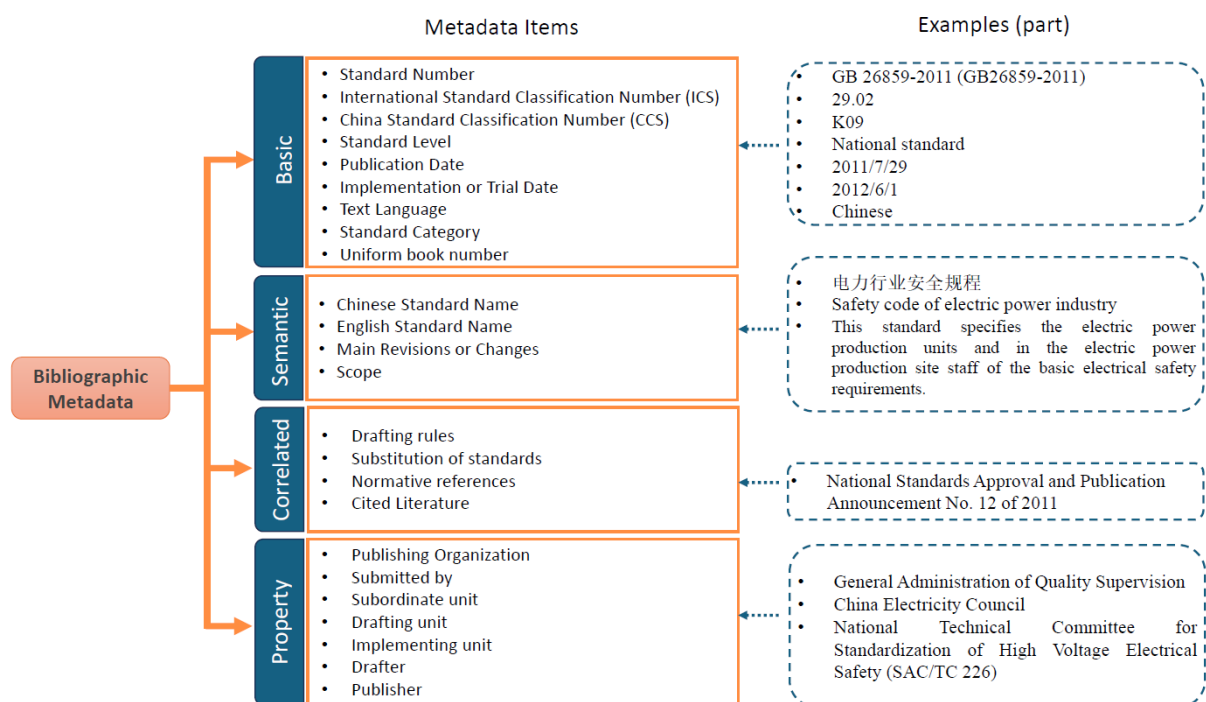
**Figure 2.** Hierarchical conceptualization of bibliographic metadata

## Detailed representations of title metadata

In this section, we will detail the representation of each metadata item in the standard documents according to the above hierarchical conceptualization.

### Basic information

**Standard number:** a character type that consists of the standard code, followed by a space, a sequential number, a hyphen, and a four-digit publication year. This is located on the front page of the standard document.

**International standard classification number (ICS):** a character type used for international classification, providing a standardized reference for the document's subject matter.

**Chinese standard classification number (CCS):** a character type that categorizes the standard within the Chinese regulatory framework.

**Standard level:** indicates the hierarchy of the standard, such as GB (National Standard), DL (Industry Standard), T (Group Standard), Q (Enterprise Standard), and DB (Local Standard).

**Release date:** presented in YYYY-MM-DD format, indicates when the standard was officially published.

**Implementation or trial date:** also in YYYY-MM-DD format, specifies when the standard becomes effective or enters a trial phase.

**Text language:** identifies the language of the document, such as Chinese or English.

**Standard category:** classifies the standard into types such as terminology, symbols, classification, testing, specifications, codes of practice, and guidelines.

**Uniform book number:** a character type that serves as a unique identifier for the publication.

### Semantic information
This category captures the content-related aspects of the standard, providing insights into its purpose and scope.

**Chinese standard name:** the official title of the standard in Chinese, presented as free text.

**English standard name:** the official title of the standard in English, also presented as free text.

**Main revisions or changes:** a free text description of significant revisions or changes made to the standard, with different entries separated by semicolons.

**Scope:** defines the applicability of the standard, detailing what is covered and the contexts in which it is relevant.

### Correlated information
This category includes references and relationships that connect the standard to other documents and frameworks.

**Drafting rules:** refers to the standards that guided the drafting process, indicated by their standard numbers. This information is typically found in the preface of the document.

**Replaced standards:** lists the numbers and names of standards that this document replaces, with multiple entries separated by semicolons.

**Normative reference documents:** free text that includes standard numbers and descriptions of documents that are considered normative, with multiple items separated by semicolons.

**Cited literature:** free text that may include standard numbers and names of cited documents, with multiple entries separated by semicolons. This is typically found on the last page and copyright page of the document.

### Property information
This category encompasses details regarding the ownership and responsibility for the standard.

**Publishing organization:** the entity responsible for publishing the standard, ensuring its formal distribution and recognition, typically presented as free text.

**Submitted by:** the organization that submitted the standard for review and publication, which may include government bodies or industry groups, with multiple entities separated by commas.

**Subordinate unit:** any subsidiary or subordinate units under the main organization that contributed to the standard, listed in free text.

**Drafting unit:** the specific group or department within an organization tasked with drafting the standard document, presented as free text.

**Implementing unit:** the entity responsible for the enforcement and practical application of the standard, detailed in free text.

**Drafter(s):** individuals involved in drafting the standard, with multiple names separated by commas to acknowledge all contributors.

**Publisher:** the organization or entity that physically produces and disseminates the final version of the standard document, typically indicated as free text.

## Prompt-based extracting bibliographic metadata by LLMs
Taking the abilities of LLMs, we design a two-stage prompt-strategy to extract the bibliographic metadata from the standards documents in a simple-yet-powerful zero-code manner.

## Stage 1: state grid-oriented instruction construct

The first stage of our method is the '*State Grid-Oriented Instruction Construct*', where we meticulously design and implement a set of instructions aimed at guiding the Large Language Model (LLM) in extracting metadata from power grid standard documents. This stage is crucial as it sets the groundwork for how the LLM interacts with the documents and identifies the necessary metadata elements. Through these instructions, we effectively prompt the LLM to perform the task of metadata extraction, ensuring that it focuses on the relevant details and follows a structured approach to gather the required information.

### Identity instruction

This process is to prompt the LLM as a meta-data extraction expert. Then tell the LLM what to do next. The instruction is as follows: *now you are a metadata extraction expert. I will provide the PDF to be extracted and the extracted metadata items. Please help me extract the relevant content.*

### Metadata instruction

Two alternative prompting strategies are designed to extract the metadata.

**Multiple instructions (MI)** extract items based on the different locations of metadata within the standard document.

1) Prompts for metadata items in the homepage of standard files:

> *please output the following content in sequence: Standard number, International standard classification number, Chinese standard classification number, Standard level (GL-national standard, DL-industry standard, T-group standard, Q-enterprise standard, DB-local standard), Chinese standard name, English standard name, Release date, Implementation or trial date, Release organization, Text language, Standard category.*

2) For metadata items in the preface of standard files:

> *please output the following content in order: Drafting rules, Replacement standards, Main revisions or changes, Proposing unit, Responsible unit, Drafting unit, Implementing unit, Drafter.*

3) For metadata items in the body part:

> *please input the following content in order: Specified scope, Applicable scope, Inapplicable scope, Normative reference documents (multiple items separated by ;).*

4) For metadata items in the copyright page of standard files:

> *Please input the following content in order: References (multiple items separated by ;) , Publisher, Uniform book number.*

**Single instruction (SI)** simply extracts all metadata without location hints as following:

> *output the following content in sequence: Standard number, International standard classification number, Chinese standard classification number, Standard level (GL-national standard, DL-industry standard, T-group standard, Q-enterprise standard, DB-local standard), Chinese standard name, English standard name, Release date, Implementation or trial date, Release organization, Text language, Standard category, Drafting rules, Replacement standards, Main revisions or changes, Proposing unit, Responsible unit, Drafting unit, Implementing unit, Drafter, Specified scope, Applicable scope, Inapplicable scope, Normative reference documents (multiple items*

*separated by ;), Please input the following content in order: References (multiple items separated by ;) , Publisher, Uniform book number.*

## Stage 2: trustworthiness estimation (TE)

Following the metadata extraction facilitated by the State Grid-Oriented Instructions, the second stage of our method is the *'Trustworthiness Estimation.'* In this phase, we evaluate the credibility of the metadata extracted by the LLM to enhance the reliability of the model's results. This involves assessing the consistency and accuracy of the extracted data, ensuring that the final output is not only comprehensive but also trustworthy. By implementing a trustworthiness estimation mechanism, we aim to mitigate the inherent variability in LLM responses and select the most reliable metadata, thereby ensuring the highest quality of information for our users.

It is a common phenomenon that LLM's Randomness, i.e. if you prompt an LLM with the same instruction multiple times, it may give different answers. To overcome this challenge, we propose a trustworthiness estimation mechanism to make our instruction method more reliable.

We using the instructions to prompt LLM k times. For each unique meta-data answer, we define its trustworthiness as:

$$t = \frac{h}{k}$$

where h denotes the appearing time of this answer. Lastly, we choose the most trustworthy i-th answer as final answer, where

$$i = \mathrm{argmax}(T(i)).$$

# Experimental study and result

## Experiment setting

**Dataset**: we conducted the experiment on 96 state grid PDF samples. All these 96 samples are treated as testing data. As the LLMs do not need to be trained or fine-tuned.

**Evaluation metric**: we use accuracy as the evaluation metric. For each sample, the accuracy is calculated as the ratio of right metadata items in total items. The final results are defined as the average performance on all testing samples.

**Baselines**: four baselines are considered, GPT-4, Kimi, GLM, and LLaMA (7B).

**Settings**: we directly use the official API of each LLM to conduct our experiment.

## Result

### Main results

Our extensive testing of various LLMs using both single and multiple instruction methods (with TE by default) has yielded promising results shown in Table 1. All models have achieved an average accuracy of over 70%, validating the effectiveness of our instructional approach in extracting metadata. Notably, the GPT-4 model has demonstrated exceptional performance with the highest accuracy of 84% by multiple instructions, affirming its robustness as a backbone model for the task. This superior performance is likely due to GPT-4's advanced natural language processing capabilities, which allow it to better comprehend and respond to complex instructions.

Moreover, a consistent trend emerges, revealing that multiple instructions consistently outperform single instruction across all models. This trend suggests that the additional context provided by multiple instructions enables the models to generate more precise and detailed responses, thereby enhancing the overall accuracy of metadata extraction.

|  | LLaMA | GLM | Kimi | GPT-4 |
|---|---|---|---|---|
| SI without TE | 0.68 | 0.72 | 0.69 | 0.73 |
| SI | 0.72 | 0.75 | 0.74 | 0.77 |
| MI without TE | 0.73 | 0.77 | 0.76 | 0.82 |
| MI | 0.79 | 0.82 | 0.82 | 0.84 |

**Table 1.** Collation of experimental results

### Comparisons between multiple and single instructions

Figure 3 give a case study comparing the differences between the two instructions. This visual representation vividly captures the nuances in the responses generated by each method. Single instruction, while efficient for basic tasks, tend to yield only abbreviated metadata, providing a quick snapshot but lacking in depth. Conversely, multiple instructions, which are more elaborate and structured, result in a richer output that includes both the abbreviated forms and the full names of metadata elements. This comprehensive approach not only enhances the granularity of the data but also significantly improves the contextual understanding of the metadata, thereby facilitating more informed decision-making processes.
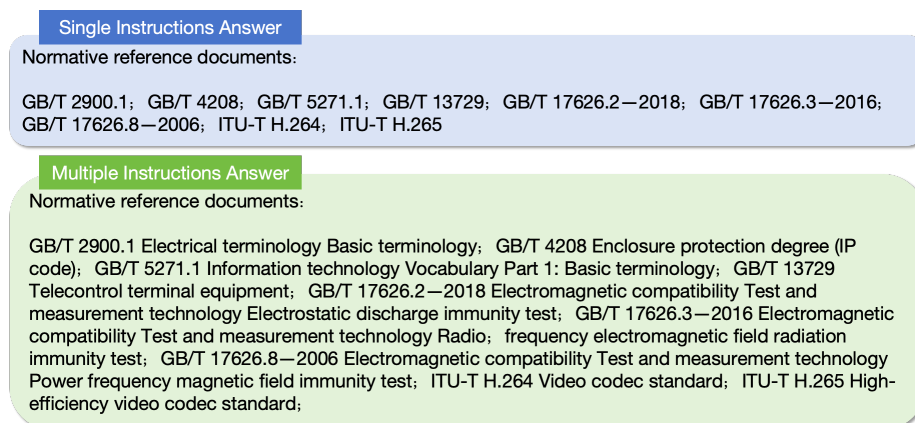


**Single Instructions Answer**

Normative reference documents:

GB/T 2900.1; GB/T 4208; GB/T 5271.1; GB/T 13729; GB/T 17626.2—2018; GB/T 17626.3—2016; GB/T 17626.8—2006; ITU-T H.264; ITU-T H.265

**Multiple Instructions Answer**

Normative reference documents:

GB/T 2900.1 Electrical terminology Basic terminology; GB/T 4208 Enclosure protection degree (IP code); GB/T 5271.1 Information technology Vocabulary Part 1: Basic terminology; GB/T 13729 Telecontrol terminal equipment; GB/T 17626.2—2018 Electromagnetic compatibility Test and measurement technology Electrostatic discharge immunity test; GB/T 17626.3—2016 Electromagnetic compatibility Test and measurement technology Radio; frequency electromagnetic field radiation immunity test; GB/T 17626.8—2006 Electromagnetic compatibility Test and measurement technology Power frequency magnetic field immunity test; ITU-T H.264 Video codec standard; ITU-T H.265 High-efficiency video codec standard;

**Figure 3.** Comparison of metadata extraction responses: single vs. multiple instructions

This observation underscores a critical point that the depth and breadth of instructions have a direct impact on the quality of metadata extraction. By leveraging multiple instructions, we are essentially equipping the LLMs with a more detailed roadmap, enabling them to navigate the complexities of document structures and extract metadata that is not just accurate but also contextually rich.

### Ablation results

To further validate the significance of our Trustworthiness Estimation (TE) mechanism, we conducted an ablation study. The results, as indicated in the table, reveal a notable decrease in performance when the TE mechanism is not utilized. This decline underscores the pivotal role of the TE mechanism in bolstering the reliability of the LLMs. By incorporating TE, we are effectively mitigating the inherent variability in LLM responses, ensuring that the extracted metadata is not only accurate but also consistent across different instances.

## Discussion and conclusion

Our study presents a comprehensive framework tailored for the power system industry, focusing on the systematic extraction of bibliographic metadata from standard documents with LLMs, contributing a significant advancement of document management and knowledge extraction in vertical fields like power grid.

## Theoretical and practical implications

The theoretical underpinning of this research is profound, marking a pioneering effort in applying LLMs to metadata extraction within the power industry. We propose the first trial of a comprehensive and unified conceptualization of bibliographic metadata for powergrid standard documents. On the other hand, the exploration of the simple-yet-powerful prompt-based strategy with zero code manner in our study transcends the conventional boundaries of AI application, offering a fresh lens through which to view the integration of artificial intelligence with industry-specific processes. By situating our study within the operational context of power systems, we contribute to the broader discourse on AI's role in enhancing operational efficiency and data integrity.

From a practical standpoint, this study underscores the transformative potential of LLMs in streamlining operational workflows and bolstering data management practices within the power industry (Schilling-Wilhelmi et al.,2024; De Santis et al.,2024). Compared with traditional knowledge extraction methods with high cost and complicated realization, such as regular expressions (RE; Chapman et al., 2017) and machine learning models (Fanni et al.,2023; Kang et al., 2020; Chowdhary& Chowdhary, 2020), we demonstrate a universal, convenient but effective pipeline with LLMs. The ability of LLMs to parse through complex documents and extract critical metadata with high accuracy not only enhances the speed of information retrieval but also mitigates the risk of human error. In addition, it also provides a new but more user-friendly interaction mode with natural language, instead of complicated RE grammar rules and machine learning algorithms traditionally leveraged in metadata extraction.

## Limitations and future work

Of course, there are also limitations in our study. One of the limitation of our current study is the modest size of the dataset involved in the experiments, which, while representative of standard documents in the power industry, may not encapsulate the full diversity of documents encountered in real-world applications. In the future work we will continue to investigate the feasibility and robustness of our model by the expansion of dataset to encompass a more comprehensive and heterogeneous collection of documents. Additionally, as we only explored the simple zero-shot prompt strategy to interact with LLMs, the model's performance, while basically satisfactory, may not fully account for the complexities and variability in language use across different documents and contents. Future work will investigate the prompts with few-shots to help LLMs better understand the tasks and also explore the integration of the retrieval-augmented generation (RAG) model to further improve the accuracy in metadata extraction.

## Acknowledgements

## About the authors

**Guowei Chen** works at State Grid Fujian Electric Power Co.,Ltd, Fuzhou, China. His research interests are technology management and knowledge services, and can be contacted at chen_guowei@fj.sgcc.com.cn

**Wei Xie** works at State Grid Fujian Electric Power Research Institute, Fuzhou, China. His research interest is artificial intelligence in electric power and can be contacted at xiewei3896@163.com

**Yanan Liu** is an Editor at Yingda Media Investment Group Co., Ltd. Her research interest is digitalization of electric power standards, and she can be contacted at 743113090@qq.com.

**Xiaoqun Yuan** is Associate Professor in School of Information Management, Wuhan University, China. He received their Ph.D. from Huazhong University of Science and Technology. His research interests are information resource management and knowledge services and can be contacted at yuan20030308@whu.edu.cn

**Liang Zhao** is Associate Professor in School of Information Mangement, Wuhan University, China. She received her Ph.D. from Tsinghua University. Her research interests are data mining and knowledge services. She is the corresponding author of the paper and can be contacted at liangzhao@whu.edu.cn

# References

Baca, M. (Ed.). (2016). *Introduction to metadata*. Getty Publications.

Bartz, C., Yang, H., & Meinel, C. (2017). STN-OCR: A single neural network for text detection and text recognition. *arxiv preprint arxiv*:1707.08831.10.48550/arXiv.1707.08831

Büchter, R. B., Weise, A., & Pieper, D. (2020). Development, testing and use of data extraction forms in systematic reviews: a review of methodological guidance. BMC medical research methodology, 20, 1-14. 10.1186/s12874-020-01143-3

Chapman, C., Wang, P., & Stolee, K. T. (2017). Exploring regular expression comprehension. In 2017 *32nd IEEE/ACM International Conference on Automated Software Engineering* (ASE),  405-416. 10.5555/3155562.3155616

China Power. (2024, July 19). State Grid lays a solid foundation for the construction of a new type of power system.

Chowdhary, K., & Chowdhary, K. R. (2020). Natural language processing. *Fundamentals of artificial intelligence*, 603-649. 10.1007/978-81-322-3972-7

Dagdelen, J., Dunn, A., Lee, S., Walker, N., Rosen, A. S., Ceder, G., ... & Jain, A. (2024). Structured information extraction from scientific text with large language models. Nature Communications, 15(1), 1418. 10.1038/s41467-024-45563-x

De Santis, A., Balduini, M., De Santis, F., Proia, A., Leo, A., Brambilla, M., & Della Valle, E. (2024). Integrating Large Language Models and Knowledge Graphs for Extraction and Validation of Textual Test Data. *arxiv preprint arxiv*:2408.01700. 10.1007/978-3-031-77847-6

Fanni, S. C., Febi, M., Aghakhanyan, G., & Neri, E. (2023). Natural language processing. In *Introduction to Artificial Intelligence* (pp. 87-99). Cham: Springer International Publishing. 10.1007/978-3-031-25928-9

Gal, M. S., & Rubinfeld, D. L. (2019). Data standardization. *NYUL Rev.*, 94, 737.

He, Y. (2020). Research on text detection and recognition based on OCR recognition technology. In 2020 *IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE)*, 132-140. 10.1109/ICISCAE51034.2020.9236870

Kabongo, S., & D'Souza, J. (2024). Instruction Finetuning for Leaderboard Generation from Empirical AI Research. arxiv preprint arxiv:2408.10141. 10.48550/arXiv.2408.10141

Kang, Y., Cai, Z., Tan, C. W., Huang, Q., & Liu, H. (2020). Natural language processing (NLP) in management research: A literature review. *Journal of Management Analytics*, 7(2), 139-172. 10.1080/23270012.2020.1756939

Kapoor, A., Gulli, A., Pal, S., & Chollet, F. (2022). *Deep Learning with TensorFlow and Keras: Build and deploy supervised, unsupervised, deep, and reinforcement learning models*. Packt Publishing Ltd.

Karthick, K., Ravindrakumar, K. B., Francis, R., & Ilankannan, S. (2019). Steps involved in text recognition and recent research in OCR; a study. *International Journal of Recent Technology and Engineering*, 8(1), 2277-3878.

Liu, J., Kong, Y., & Peng, G. (2022, June). Interpreting the Development of Information Security Industry from Standards. In *International Conference on Human-Computer Interaction* (pp. 372-391). Cham: Springer International Publishing. 10.1007/978-3-031-05463-1

Papaluca, A., Krefl, D., Rodriguez, S. M., Lensky, A., & Suominen, H. (2023). Zero-and few-shots knowledge graph triplet extraction with large language models. arxiv preprint arxiv:2312.01954. 10.48550/arXiv.2312.01954

Proctor, M., Fusco, M., Vacchi, E., & Sottara, D. (2019). Rule modularity and execution control enhancements for a Java-based rule engine. In 2019 *IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering* (AIKE) , 89-96. 10.1109/AIKE.2019.00023

Reyes, J., Mula, J., & Díaz-Madroñero, M. (2023). Development of a conceptual model for lean supply chain planning in industry 4.0: multidimensional analysis for operations management. *Production Planning & Control*, 34(12), 1209-1224. 10.1080/09537287.2021.1993373

Riley, J. (2017). Understanding metadata. *Washington DC, United States: National Information Standards Organization* ([http://www.niso.org/publications/press/UnderstandingMetadata.pdf](http://www.niso.org/publications/press/UnderstandingMetadata.pdf)), 23, 7-10.

Schilling-Wilhelmi, M., Ríos-García, M., Shabih, S., Gil, M. V., Miret, S., Koch, C. T., ... & Jablonka, K. M. (2024). From Text to Insight: Large Language Models for Materials Science Data Extraction. *arxiv preprint arxiv:2407.16867. 10.48550/arXiv.2407.16867*

Shao, W., Zhang, R., Ji, P., Fan, D., Hu, Y., Yan, X., ... & Chen, L. (2024). Astronomical knowledge entity extraction in astrophysics journal articles via large language models. Research in Astronomy and Astrophysics, 24(6), 065012.

Singh, P., & Manure, A. (2019). *Learn TensorFlow 2.0: Implement Machine Learning and Deep Learning Models with Python*. Apress.

Sivarajkumar, S., Kelley, M., Samolyk-Mazzanti, A., Visweswaran, S., & Wang, Y. (2024). An empirical evaluation of prompting strategies for large language models in zero-shot clinical natural language processing: algorithm development and validation study. JMIR Medical Informatics, 12, e55318. 10.2196/55318

Stevens, E., Antiga, L., & Viehmann, T. (2020). *Deep learning with PyTorch*. Manning Publications.

Xu, D., Chen, W., Peng, W., Zhang, C., Xu, T., Zhao, X., ... & Chen, E. (2023). Large language models for generative information extraction: a survey. arxiv preprint arxiv:2312.17617. 10.1007/s11704-024-40555-y

Xue, L., Zhang, D., Dong, Y., & Tang, J. (2024). AutoRE: Document-Level Relation Extraction with Large Language Models. In Proceedings of the 62nd Annual Meeting of the Association for

Computational Linguistics (Volume 3: System Demonstrations) , 211-220.
10.48550/arXiv.2403.14888

Zeng, M. L., & Qin, J. (2020). *Metadata*. American Library Association.

Zhang, J., Ullah, N., & Babbar, R. (2024). Zero-shot learning over large output spaces: utilizing indirect knowledge extraction from large language models. arxiv preprint arxiv:2406.09288.