Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 134 (2025) 157–162

www.elsevier.com/locate/procedia

Proceedings of the 58th CIRP Conference on Manufacturing Systems 2025

Identification of machine parameters in battery cell production: A comparison of different methods and approaches for mapping parameters

Wilhelm Jaspers[a*], Arno Schmetz[a], David Roth[a], Bandik Becker[a], Achim Kampker[a,b]

[a]Fraunhofer Research Institution for Battery Cell Production FFB, Bergiusstr. 8, 48165 Münster, Germany
[b]Production Engineering of E-Mobility Components, RWTH Aachen University, Bohr 12, 52072 Aachen, Germany

* Corresponding author. E-mail address: wilhelm.jaspers@ffb.fraunhofer.de

Abstract

Optimization of battery cell production processes is increasingly based on digital use cases that can reduce costs and improve quality. However, these requires accurate and semantically understandable data that is human- and machine-readable and stored in standardized formats. Machine manufacturers often use unique or proprietary parameter standards and naming conventions, making data integration difficult. Linking machine data to semantic database standards requires expert knowledge and significant manual effort. Automating this process could reduce the workload enormously.
This paper evaluates approaches for automating the mapping and comparison of machine parameter sets within the context of battery cell production. Using real production data and semantic modelling from multiple partners, different methods were applied to link individual parameters. The outcomes were assessed with specific metrics, revealing that Large Language Models (LLMs) show significant promise due to their flexibility and effectiveness in handling complex semantic comparisons.

Peer-review under responsibility of the scientific committee of the International Programme committee of the 58th CIRP Conference on Manufacturing Systems

Keywords: battery cell production; identification; machine learning

1. Motivation

The battery industry is currently undergoing exponential growth, which is unparalleled in the history of industrial markets. A comparison with the semiconductor industry reveals that it took 40 years for this sector to reach a market value of 550 billion euros. The battery industry, by contrast, is currently undergoing an even more impressive growth curve. It is estimated that by 2030, over 200 new battery cell factories will be constructed worldwide, representing an opening rate of approximately 2.5 factories per month. These rapid developments present both significant business opportunities and considerable challenges for manufacturers. [1]

One of the principal methods for surmounting these obstacles is the implementation of digital use cases that have the potential to facilitate substantial cost savings in the production of battery cells. The results of cost modeling indicate that the implementation of specific digital applications, such as predictive quality assurance and material flow simulations, can result in cost reductions of up to 0.8% in production. This equates to an annual potential saving of approximately 30 million US dollars per plant. Considering the intensifying competition and mounting price pressure in the battery industry, the deployment of these measures is of paramount importance for the long-term competitiveness of manufacturers. [2]

Nevertheless, the successful implementation of digital use cases hinges on the availability and usability of machine data. It is essential that this data be made accessible via suitable interfaces to enable efficient processing and integration into

Peer-review under responsibility of the scientific committee of the International Programme committee of the 58th CIRP Conference on Manufacturing Systems
10.1016/j.procir.2025.03.022

digital systems. A considerable number of existing machines, particularly those of an older generation, frequently lack the requisite interfaces, which renders the process of digitalization more challenging and costly.[3]

This represents a significant obstacle, particularly within an industry where machinery and systems originate from disparate industrial sectors and geographical regions. In particular, Asian manufacturers, who are dominant in the field of cell production, frequently utilize proprietary solutions that do not facilitate the establishment of a common standard for machine parameters. [4]

The automated mapping of machine parameters to internal information models represents a promising solution for overcoming the challenges. This process not only facilitates the comprehensibility and semantic description of the data, but also minimizes the effort and expertise required for mapping. Consequently, it facilitates the effective implementation of digital applications in battery cell production, thereby contributing to the optimization of production processes.

This paper presents an analysis and evaluation of various methods for enabling automated mapping, with a focus on their application in machine integration at Fraunhofer FFB and the “ENLARGE” research project. The objective is to facilitate productivity and competitiveness gains in this rapidly expanding industry.

2. Background

In the process of digitally integrating systems into existing systems, it is first necessary to identify the individual system parameters. It is only after the data generated in the plant has been mapped to the internal systems that it can be used profitably in production. Chapter 2.1 introduces the fundamental principles of mapping, while chapter 2.2 offers an overview of potential approaches for automating this process.

2.1. Mapping in digitalization

Purposeful data analysis requires an understanding of what the generated data represents: for sensor data generated in manufacturing machines additional information can be important, e.g. measured characteristic and position of the measuring device, the corresponding measured variable, unit and value range [5].

Equally important is the identification of data that is promising for profitable use. The creation of a "parameter map" of battery cell production with relevant and semantically clearly named parameters and their qualities is part of the “ENLARGE” research project. The designations of this “parameter map” shall then be used in the data and information modeling of the internal systems and digital solutions. Mapping refers to the assignment of the generated data to the parameters of the “parameter map” [6]. Only through this is the data consistently understandable and usable. Data mapping as part of data integration implies a time-consuming effort in many cases [7]. In our example it implies an effort for translation and transformation, as proprietary names for parameters may be different or unclear and other units or value ranges may be used.

2.2. Methods for mapping

Regarding the task of mapping, different methods and approaches exist currently. They differ in their basic concepts and types of application. This includes the inputs and contexts to be used. Such inputs for the mapping may consist of data, information or knowledge. Data is the representation of models and entities, while information is the result of a processing, providing meaning to data. Lastly, the knowledge describes results of cognitive processes, like associations and reasoning [8]. For the mapping task, data is the name and value for the given source, while information denote the context like which kind of machine or process the source is. In that case, knowledge describes the domain context and associations, like typical types of sensors in such machines and (in)plausible values.

Based on this, approaches for the mapping task can be clustered into different classes of approaches. When using only the names of the fields, no information or knowledge is used. Therefore, these approaches are pure syntactical approaches and are easy to use. In contrast to the syntactical approaches, semantical approaches leverage the benefits of the context. Simple semantical approaches use information in addition to the syntactical aspects to perform the mapping. Finally, the complex semantical approaches also use the knowledge of associations and the domain accordingly.

For the syntactical approaches, usually a distance metric is used to determine the syntactical distance from a given string to candidate. Using the minimum number of edits in terms of insertions, deletions and substitutions of characters, the Levenshtein distance provides an extension of the hamming distance being capable of strings of different lengths in quadratic time complexity [9]. Interpreting the given strings as vectors (e.g., by term frequencies), the cosine similarity can be applied. Based on the vectors, this method calculates the alignment of the vectors retrieving a score between -1 and 1, where the higher value indicates high mathematical alignment, representing similarity of the strings [10]. Current research deals also with the semantic addition of the cosine similarity for larger texts [11]. The Jaccard similarity also interprets the text as vectors, where the proportion of the terms in comparison to unique terms is calculated. Based on this proportion, two strings are compared, emphasizing the shared content between strings [12]. Splitting the texts into fixed-length sequences or contiguous elements is the key of the N-Gram matching approach, where the overlaps of the n-grams are calculated. This method is widely used for tasks like keyword matching. The fixed length is given by n, which denotes the name of the approach [13]. While the classical N-Gram matching uses word-level n-grams, extensions for character-level or other name-specifics exist [14]. When dealing with matching tasks, where a normalized similarity indication is needed, fuzzy

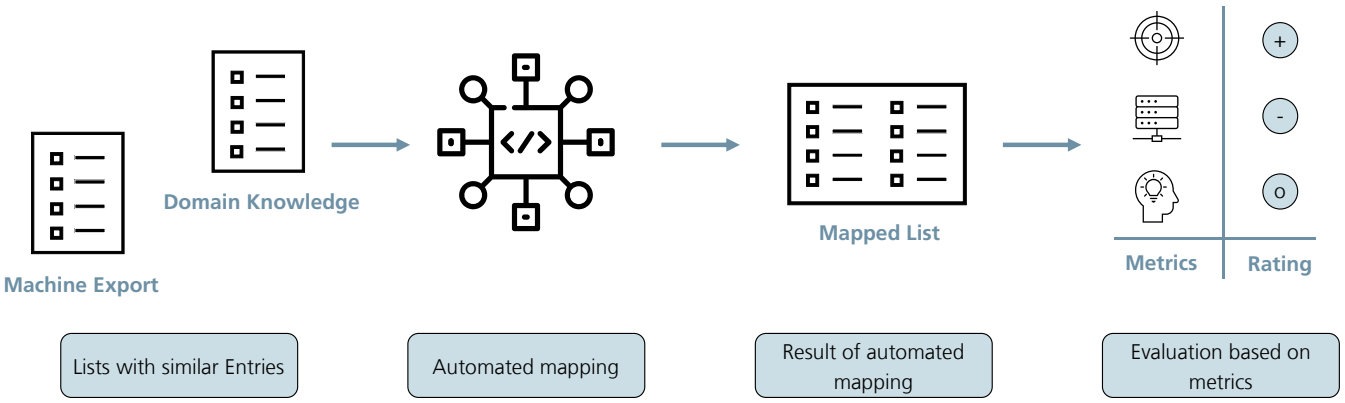

Fig 1. Procedure within the evaluation process.

matching allows the application of different metrics for distances (e.g., Levenshtein) and given a threshold, allows deciding on a matching. Therefore, fuzzy matching is more like an incorporation method to the approaches above [15].

While a variety of syntactical approaches exist, explicit simple semantic approaches cannot be found directly. A pure syntactical approach has to perform the matching and distance calculation for each existing parameter name. Using the context (information), the list can be reduced significantly to the set of relevant parameters. Therefore, simple semantic approaches mainly consist of syntactical approaches, where information is used to filter the possibilities beforehand. Further, the information about the likelihood of name lengths or patterns can help to adapt and parameterize the syntactical approaches accordingly. In consequence, we consider simple syntactical approaches as the parameterized and pre-filtered syntactical approaches.

For the complex semantical matchings, current research trends deal especially with the usage of large language models (LLM), but also other methods exist. The approach of Word Embeddings interprets text as a continuous vector space, in which the words are positioned. Using the knowledge about meaning and context of the words, their semantical similarity is used to position semantically similar words closer together in the vector space. Based on this vector space, then syntactical similarity metrics like the cosine similarity can be used for the overall similarity calculation [16]. Still, while using semantics, the Word embeddings may show shortcomings in cases of word with different meanings (polysemy) by the context (e.g., ‘die’ has different context for ‘slot die’ and ‘straight as die’). To overcome such limitations, BERT Embeddings use the bidirectional encoders to check on the context per word in the embeddings and adjust the positioning accordingly [17]. Regarding the LLM-based approaches, the LLM Embedding approach used the same concept as the Word Embedding approach, but for the positioning of the words, a complete LLM is used for the indication of contextual und semantical similarity. By this, deeper semantic understanding and handling of complex structures is addressed [18]. Another LLM-based approach is the method of LLM Few-Shot Learning for the matching. In this case, a trained LLM is presented with the exact matching task including examples. The LLM does not need a large-scale set of labeled test data of the task but uses the LLM-training to generalize from the given test cases during inference.

Another LLM-based approach by Microsoft and Siemens uses the context like the other LLM-based approaches and leverages the OpenAI (GPT 4) capabilities to also incorporate the technical documentation (provided for example as PDF) to perform a possible matching for onboarding of legacy equipment [19]. A completely different concept is the interpretation of the actual values observed and looking for similar expected signals from knowledge. This approach allows applications even with no names provided for the signals of the source but requires a series of data from the source to perform the mapping. Since this paper deals with matching for onboardings, where in-production time series data is not available, this approach is left out in the following.

Finally, the baseline approach is the manual task of mapping, where data engineering and production experts jointly search for possible mappings by expert knowledge and intuition. As this heavily relies on specifics of the chosen experts, this baseline is almost impossible to quantify in terms of performance. Based on expert interviews and experience, this baseline can be considered a major investment of time and resources without clear guarantees on the quality, rendering it the baseline.

3. Definition of the evaluation metrics and the evaluation process

Before a meaningful technical analysis of the different methods can be carried out, metrics for evaluating the methods must first be established. The degrees of fulfilment of the metrics then enable an evaluation and a comparison of the different methods with each other. To further increase comparability, a good basis of input parameters must be created, which is explained in more detail in chapter 3.2. Finally, the basic structure of the evaluation should be precisely defined.

3.1. Planning the evaluation

In addition to the evaluation metrics and a prepared parameter input, it is equally important to plan the execution carefully. The exact process of the evaluation is shown in Figure 1. The list of metrics prepared in 3.3 are used as input for the individual methods. The methods then perform a mapping and create a mapped list or matrix with a coefficient of correspondence. This matrix can then be aligned with the sample solution and evaluated regarding the metrics defined in 3.2. To ensure that all methods can be compared, they are programmed identically on a virtual machine using Python, whereby the reading of the tables for the input and the creation of the matrix as a result is structured identically. During the execution, the individual methods are tested several times with the same input parameters to minimize individual outliers and a falsification of the result.

The first two steps of the battery cell production process are used for the implementation. The data from the parameter list in 3.3 are thus filtered out specifically for mixing and coating. Similarly, excerpts of the machine parameters from two systems are used for each process step. This means that different machines as well as identical machines are compared with other parameter lists and information models.

3.2. Evaluation Metrics

Three categories were developed when setting up the metrics: *Efficiency*, *Performance*, and *Subjective Assessment*. These categories each include two precise evaluation metrics. A list of all metrics with a short description is shown in Table 1.

Table 1. Categories and Metrics for the evaluation.

	Category	Metric	Description
Cat. 1	Efficiency	Precision	The proportion of relevant instances among the retrieved instances
		Accuracy	The proportion of correctly identified pairs among all possible pairs
Cat. 2	Performance	Complexity	The behavior when scaling to more complex strings and larger lists
		Resource Consumption	The computation time and memory usage
Cat. 3	Subjective Assessment	Explainability	The ease with which the results of the method can be interpreted and understood
		Implementation Effort	The degree of difficulty in implementing and integrating the method into existing systems

The first category, *Efficiency*, includes the two metrics *Precision*, which is reflected by the proportion of relevant instances among the retrieved instances, and the metric *Accuracy*, which is reflected by the proportion of correctly identified pairs among all possible pairs. This category ensures that the methods also perform the correct mappings and thus lead to a correct result. *Accuracy* is measured as the proportion of correctly classified instances of the total number and is calculated using formula (1). TP (True Positives) and TN (True Negatives) represent the correct classifications, while FP (False Positives) and FN (False Negatives) are the incorrect classifications.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Precision, on the other hand, assesses the accuracy of positive predictions and is calculated as stated in formula (2).

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

As the focus of this paper is the comparison of different methods, the results of these metrics are normalized with the formula (3). This normalization gives the method with the best results 100 %. All others can then be evaluated based on these results, which makes comparison easier.

$$X_{norm} = \frac{X-X_{min}}{X_{max}-X_{min}} * 100 \quad (3)$$

In the second category, *Performance*, the two metrics *Complexity* and *Resource Consumption* are evaluated. In the context of the *Complexity* metric, the method is evaluated in terms of the inherent intricacy of the task at hand. This involves analyzing how the method behaves when scaling to more complex strings and larger lists. In the *Resource Consumption* metric, the execution of the method is evaluated in terms of its computation time and memory usage. To do this, the time from the starting point to the result is measured and the memory used is recorded. The longer a method takes to process, the worse the evaluation.

In the third category, *Subjective Assessment*, the two metrics of *Explainability* and *Implementation Effort* are considered. For the results and the mapping of the methods to be accepted, it is essential that the results are understood. The *Explainability* metric considers the ease with which the results of the method can be interpreted and understood. In addition to user acceptance, it is equally important to integrate the method into the existing IT architecture. The *Implementation Effort* metric assesses the degree of difficulty in implementing and integrating the method into existing systems. This metric also includes the necessary framework conditions and how flexible the setup is. For example, on-premises vs. cloud connection.

Since, in contrast to the first four metrics, there are no fixed numbers and results for these metrics, they are evaluated as subjective metrics within the project.

In this paper, our primary focus is on the scientific perspective of evaluation, making the result of the methods particularly significant. Therefore, the category *Efficiency* is rated higher in the evaluation than the other two categories. However, to gain a general overview of the methods and an assessment to select the most suitable method for the use case, the other two categories are also essential and are therefore included.

Table 2. Evaluation of methods for automated mapping according to the given metrics.

(-) = *Insufficient* (o) = *Adequate* (+) = *Excellent*

	Metric	Levenshtein Distance	Cosine Similarity	Jaccard Similarity	N-Gram Matching	Fuzzy Matching	Word Embeddings	BERT Embedding	LLM Embedding	LLM Few Shot	GPT 4
Cat. 1	Precision	o	-	o	-	o	-	-	o	-	+
	Accuracy	o	o	-	o	o	-	-	o	-	+
Cat. 2	Complexity	+	+	+	+	+	-	-	-	-	o
	Resource Consumption	+	+	+	+	+	o	o	o	o	-
Cat. 3	Explainability	+	o	+	+	o	-	-	-	-	-
	Implementation Effort	+	+	+	+	+	+	+	o	+	-

3.3. Data preprocessing

The data used for evaluation is based on four production machines for battery cell production – two (extrusion) mixing machines and two coating machines.

They were mapped against defined parameters of a parameter map which is a result of the research project "ENLARGE". These defined parameter lists include the most relevant parameters for process understanding and optimization as well as product quality. Based on expert domain knowledge the battery cell production was split into process steps. A list of parameters was defined for the specific process steps. For this evaluation the lists for the process steps mixing and coating were used.

The four existing machines - extruder and coater - provide process parameters with naming conventions based on the respective machine supplier. The machine parameters were mapped to the standardized parameter lists in the following way. For each machine all available parameters were read out via OPC UA. Subsequently, the parameters were categorized to filter for only process parameters, thereby excluding alarm values. This was achieved through the collaboration of experts for the specific process step and the considered machine in collaboration with experts for the data engineering and IT-systems, who worked out the mapping manually. Each standardized parameter was checked for an available counterpart available in the OPC UA server of the machine.

The resulting solution sheet has the available mappings for each standardized parameter and can be used for the evaluation of the automated mapping.

4. Results

After the individual methods have been implemented in separate Python scripts and the evaluation has been conducted following the process in chapter 3.1, the results are presented in detail in Table 2. Each method is evaluated based on six metrics, rated at varying levels: *Insufficient* (-), *Adequate* (o), and *Excellent* (+). Since both mathematically calculated metrics and subjective ones are evaluated, translation at these levels makes sense for comparability.

As advanced in chapter 2.2, it is confirmed that especially syntactical approaches perform relatively well in the categories of Performance and Subjective assessment. These approaches have been used for many different tasks for years. Moreover, these approaches are well integrated into several libraries, which simplifies the use of the methods. As a result, the *Implementation Effort* is particularly low, and the understandability of the individual methods is usually good. Since these approaches perform a purely syntactic comparison of the individual descriptions, the *Resource Consumption* is relatively low as well. These methods also handle scaling up the datasets well since managing multiple queries does not entail potential extra effort. In contrast to the very good results, these methods perform poorly to mediocre in the *Efficiency* category. The methods only do well with relatively obvious matches. However, the language has a significant influence here. For example, the Levenshtein distance achieves 33.6 % in both the precision and accuracy metrics. By comparison, the Jaccard similarity has a precision of 46.2 %, while it only achieves 19 % accuracy, which also reflects the general range of these methods for both metrics. These results suggest that the performance of the above methods, based on the normalization, is significantly worse compared to the best methodology. When different languages are used (e.g., parameter descriptions in German and machine exports in English), these methods perform poorly compared to the semantic methods.

In addition to this advantage of semantic models, simple semantic models performed relatively poorly in the evaluation. These models are much more complex and process much more in the background compared to purely syntactic methods. Therefore, it is more difficult to successfully handle such complex tasks with many datasets. Scaling up is thus almost impossible with simple semantic models. Likewise, these approaches require high computing power and offer little insight into how individual decisions were made.

However, the barriers to entry into semantic models are relatively low. These models can be quickly and easily deployed using simple libraries. More complex semantic

models like GPT 4 are the exception. These are comprehensive and require some time to familiarize oneself. Still, in comparison with the other models, they show the best results in the *Efficiency* category. Especially in this area, the addition of more context information can positively influence the result. The difference between the semantic methods becomes particularly clear in this context. While the LLM Few Shot method achieves the worst results (0 %), GPT 4 achieves the best values in the metrics precision and accuracy and thus sets the benchmark of 100 % for the evaluation of all other methods.

As this paper particularly considers this *Efficiency* category under the research aspect, the complex semantic methods seem the most promising overall for the mapping task when onboarding new facilities in battery cell manufacturing. For this reason, the particularities and challenges of implementing complex semantic models like LLMs in the context of the evaluation carried out here are highlighted in more detail in the following.

LLMs offer deep semantic analysis, capturing complex semantics beyond simple string matching. Moreover, their flexibility allows use in various Natural Language Processing (NLP) tasks. These include sentiment analysis, named entity recognition, and text summarization, rendering LLMs extremely valuable in the NLP field. [20]

The use of LLMs in complex tasks, while promising, presents certain challenges. High and potentially prohibitive computational requirements, significant cost implications particularly in cloud environments and the complexity of interpreting results stemming from intricate internal representations are notable difficulties. Thus, leveraging LLMs comes with its set of complexities.

5. Summary and Outlook

In this paper, the mapping challenge during the onboarding phase in the production machines for battery cell production was focused. To tackle this challenge, multiple approaches for the technical solution targeting automation were evaluated.

Using a set of evaluation metrics and prototypical implementations of the approaches, an evaluation took place. This evaluation shows that no approach is fulfilling all the requirements and therefore no all-suiting solution is available. The challenge of mapping remains a relevant field of action in the future.

Based on these results, further research on approaches, as well as combinations and multi-approach strategies should be investigated. Also, the usage of context information and knowledge from the actual production could enable new approaches leveraging the semantical models accordingly. This information may include, for example, aliases, detailed descriptions of parameters, translations or parameter-specific properties such as tolerances. By training LLMs with such information and the mapped data sets, the *efficiency* of automated onboarding can be greatly increased. Further, entangling the different signals might be an interesting approach to investigate. In this, the signals are not observed and evaluated separately, but in the complete group of all signals, allowing probabilities and mapping information of one signal to influence the mapping of other signals constructively.

Acknowledgements

Funding reference: The project “ENLARGE” is funded by the German Federal Ministry of Education and Research. Grant number: 03XP0539F.

References

[1] Grandl, G., et al.: Battery Manufacturing 2030: Collaborating at Warp Speed. Porsche Consulting & VDMA 2024.
[2] Krauß, J., et. al.: Potentials of Digitalization in Battery Cell Manufacturing, Fraunhofer FFB & Accenture Industry X, 2024.
[3] Gönnheimer, P., et al.: Interoperable System for Automated Extraction and Identification of Machine Control Data in Brownfield Production, 2023
[4] Roadmap Batterie-Produktionsmittel 2023, VDMA, 2023
[5] Compton M., et al. (2012): The SSN ontology of the W3C semantic sensor network incubator group. Journal of Web Semantics, Volume 17, 25-32.
[6] Research project “ENLARGE”, 2024: https://www.enlarge-projekt.de/
[7] Kruse, S. & Papotti, P. & Naumann, F. (2015): Estimating Data Integration and Cleaning Effort. In International Conference on Extending Database Technology, pages 61–72.
[8] Chen, M.; Ebert, D.; Hagen, H.; Laramee, R. S.; Van Liere, R.; Ma, K.-L.; Ribarsky, W.; Scheuermann, G.; Silver, D.: Data, information, and knowledge in visualization. In: IEEE computer graphics and applications 29 (2008), Nr. 1, pp. 12–19
[9] Yujian, L., & Bo, L. (2007). A normalized Levenshtein distance metric. IEEE transactions on pattern analysis and machine intelligence, 29(6), 1091-1095.
[10] Gunawan, D., Sembiring, C. A., & Budiman, M. A. (2018, March). The implementation of cosine similarity to calculate text relevance between two documents. In Journal of physics: conference series (Vol. 978, p. 012120). IOP Publishing.
[11] Rahutomo, F., Kitasuka, T., & Aritsugi, M. (2012, October). Semantic cosine similarity. In The 7th international student conference on advanced science and technology ICAST (Vol. 4, No. 1, p. 1). South Korea: University of Seoul.
[12] Niwattanakul, S., Singthongchai, J., Naenudorn, E., & Wanapu, S. (2013, March). Using of Jaccard coefficient for keywords similarity. In Proceedings of the international multiconference of engineers and computer scientists (Vol. 1, No. 6, pp. 380-384).
[13] Cavnar, W. B., & Trenkle, J. M. (1994, April). N-gram-based text categorization. In Proceedings of SDAIR-94, 3rd annual symposium on document analysis and information retrieval (Vol. 161175, p. 14).
[14] Salah, A. H., Al-Sanabani, M., Hadwan, M., & Al-Hagery, M. A. An Improved N-gram Distance for Names Matching.
[15] Kalyanathaya, K. P., Akila, D., & Suseendren, G. (2019). A fuzzy approach to approximate string matching for text retrieval in NLP. J. Comput. Inf. Syst. USA, 15(3), 26-32.
[16] Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) (pp. 1532-1543).
[17] Rajaee, S., & Pilehvar, M. T. (2021). An isotropy analysis in the multilingual BERT embedding space. arXiv preprint arXiv:2110.04504.
[18] Keraghel, I., Morbieu, S., & Nadif, M. (2024, April). Beyond words: a comparative analysis of LLM embeddings for effective clustering. In International Symposium on Intelligent Data Analysis (pp. 205-216). Cham: Springer Nature Switzerland.
[19] Barnstedt, E. (2024). Automatic Asset onboarding with OpenAI. Presentation at building IoT 2024 conference (Munich).
[20] Devlin, J., Chang, M., Lee, K. & Toutanova, K. (2019, June). Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Vol. 1, pp. 4171-4186).