A Comparative Analysis of Historical Culinary Recipes Using Topic Modeling and Data Visualization

Maria Teresa Artese[0000-0001-8453-0807], Isabella Gagliardi[0000-0002-3706-0993]

IMATI-MI National Research Council, Via A. Corti 12, 20133, Milan, Italy
artese@mi.imati.cnr.it, gagliardi@mi.imati.cnr.it

Abstract. The proposed method involves analyzing historical culinary recipes through topic modeling and data visualization. It utilizes an unsupervised pipeline, pre-trained language models, and visualization techniques. The method was thoroughly tested on English-language cookbooks and provides information on the evolution of culinary practices.

Keywords: Topic Modelling, Historical Culinary Recipes, Clustering, Data Visualization, Large Language Models (LLM).

1 Introduction

Culinary traditions are an important cultural heritage that reflects the interplay of culture, society, and environment throughout human history. Recipes have been passed down through generations, preserving the essence of diverse cuisines and offering insights into the evolution of culinary practices. In the digital age, vast repositories of historical culinary texts provide an unprecedented opportunity for in-depth exploration of the culinary heritage of societies worldwide.

This paper presents a new method for analyzing historical culinary recipes using topic modelling and data visualization. Our aim is to uncover the complex culinary knowledge contained in historical texts and identify patterns, trends, and connections that have developed over centuries of culinary evolution. The goal of this work is to provide integrated tools for topic modelling and data visualization with pre-trained Large Language Models (LLM) to facilitate the comparative analysis and interpretation of historical culinary texts, and to highlight how new recipes differ from older ones. The effectiveness of this approach and its usefulness for users stem from several key features.

Unsupervised pipeline: once a few hyperparameters and transformation models have been optimized, the pipeline is completely unsupervised. This means that the system can autonomously process and analyze data without requiring manual intervention.

Pre-trained language models enhance the reliability and robustness of the system, even with a relatively small dataset consisting of just a few hundred entries. Additionally, pre-trained models can handle a wide range of vocabulary, including uncommon or historical terms, mitigating concerns regarding Out of Vocabulary (OOV) terms, which are common in old recipes or historical books.

Digital Presentation and Preservation of Cultural and Scientific Heritage. Conference Proceedings. Vol. 14, Sofia, Bulgaria: Institute of Mathematics and Informatics – BAS, 2024. ISSN: 1314-4006, eISSN: 2535-0366

Visualization through Graphs and Word Clouds: The presentation of data through graphs and word clouds allows for rapid interpretation by a wider range of users. This visual representation simplifies the comprehension of intricate data structures and patterns, improving user engagement and system usability.

Highly adaptable to different datasets and languages: The use of UMAP, HDBSCAN and pre-trained language models significantly simplifies the adaptation of the prototype to other contexts and languages.

The pipeline has been built from scratch, integrating dimensionality reduction (McInnes et al., 2020) and clustering algorithms (HDBscan used here) (Campello et al., 2013; McInnes et al., 2017) to evaluate a variety of hyperparameters. Afterwards, it underwent thorough testing using publicly available English-language cookbooks. This approach ensured the pipeline's robustness and effectiveness for different applications.

In the following sections, we present related works and describe the methodology used to analyze historical culinary recipes, including data collection, preprocessing, topic modeling, and visualization techniques. After presenting the experimentation, the paper briefly discusses the preliminary results. Future work involves expanding the method to encompass broader perspectives within the digital humanities.

2 Related Works

This paper focuses on several key topics in natural language processing and machine learning. Topic modeling techniques, such as Latent Semantic Indexing (LSI), probabilistic latent semantic indexing (pLSI), and Latent Dirichlet Allocation (LDA), have been used to uncover hidden themes within text data since the 1990s. LDA has been used extensively for topic modelling. Non-negative Matrix Factorization factorizes the term-document matrix into two lower-dimensional matrices representing the topics and their distributions over documents. The development of language models has given rise to new tools and methods. Bidirectional Encoder Representations from Transformers (BERT), the bidirectional unsupervised language representation model developed by Google (Devlin et al., 2019), has shown good result in topic extraction. BERTopic (Grootendorst, 2022) utilizes the BERT model to embed documents into high-dimensional vector space and then applies hierarchical clustering to identify topics. Top2Vec (Angelov, 2020) employs a combination of BERT for document embedding and UMAP (McInnes et al., 2020) for dimensionality reduction. It discovers topics based on dense areas in the document embedding space.

Bertopic, Top2Vec, and other topic modelling methods have achieved significant success due to their ease of use and high-quality results. These techniques have been applied in various domains, such as banking (Ogunleye et al., 2023), tourism, cultural policy setting, and crisis support (Sprenkamp et al., 2023). Topic modeling methods have been used in the tourism industry to analyze traveller reviews and identify popular destinations, attractions, and trends (Li, 2023). Additionally, these techniques have facilitated the analysis of cultural data in the realm of cultural policy setting (Wojciechowska et al., 2023). This enables policymakers to make informed decisions regarding cultural initiatives, funding allocations, and resource management. Twitter, with its extensive user base and real-time data stream, has served as a valuable source of data for evaluating and comparing different topic modeling models (Egger & Yu, 2022).

3 Our Approach

The goal of this study is to explore a straightforward and direct approach for comparing the semantic content of diverse textual datasets. The novelty of the pipeline lies in its processes and techniques that support the identification of key concepts within each dataset, their clustering, and the creation of visualizations to highlight similarities and differences between datasets. These processes usually involve several steps, including data reduction, data clustering, identification of the most significant terms globally and within each cluster, and selection of optimal pre-trained models and hyperparameters for models and methods.

3.1 Defining Features of Our Approach

UMAP (Uniform Manifold Approximation and Projection) (McInnes et al., 2020) is a dimensionality reduction technique used in machine learning and data analysis, competing with other popular methods such as PCA and t-SNE. UMAP uses manifold learning to transform high-dimensional data into a lower-dimensional space while preserving the local structure of the data. It uses a mixture of topographic and geometric techniques that effectively address nonlinear relationships within the data, and excels at handling high-dimensional data sets. The algorithm relies on two primary parameters: the number of dimensions and the distance metric used.

HDBSCAN (Campello et al., 2013; McInnes et al., 2017) is a clustering technique that groups similar data points together, regardless of their shape, density, or size. It does this by finding regions of high density in the data set and classifying points according to how close they are to these regions.

Unlike other clustering methods, HDBSCAN does not require the user to specify the size of the neighborhood to search or the number of clusters to include. Both the number and geometry of the clusters are determined automatically. In addition, HDBSCAN can recognize non-clustered points as noise. Therefore, it can be applied to data sets with significant density variations. The technique involves several adjustable parameters tailored to the dataset, such as the minimum cluster size and the minimum number of samples needed to form a dense region.

An advantage of employing UMAP and HDBSCAN over other clustering methods lies in their capability to manage nonlinear relationships among data points. This is especially crucial in natural language processing, where the interactions between words and sentences can be highly intricate and nonlinear.

Transformers and Large Language Models. Originally *proposed* by Vaswani et al. (2017), transformers represent a new neural network architecture designed to process sequential data, particularly text. Departing from the sequential processing constraints of traditional recurrent neural networks (RNNs) and long short-term memory networks (LSTMs), transformers provide an improvement that reduces the computational burden associated with capturing the long-range dependencies inherent in linguistic structures.

Large Language Models (LLMs), derived from transformer-based architectures, represent a turning point in NLP research. These models are trained on large corpora of text data using unsupervised or semi-supervised learning paradigms, giving them the ability to understand the intricacies of natural language. Prominent examples of LLMs include OpenAI's Generative Pre-trained Transformer (GPT) series, Google's Bidirectional Encoder Representations from Transformers (BERT), and Facebook AI's Robustly Optimized BERT Approach (RoBERTa), among others.

We used these distinctive components to create a system that can easily generate semantic multi-level graphs for archives. In that case, the system demonstrated the ability to automatically generate navigable graphs of multilingual archives, while providing users with user-friendly tools for discovering archives whose contents are unknown (Artese & Gagliardi, 2023; Gagliardi & Artese, 2023).

C-TF-IDF, or Class-based TF-IDF (Grootendorst, 2022), represents a variant of the conventional Term Frequency-Inverse Document Frequency (TF-IDF) approach utilized in information retrieval and text mining. TF-IDF serves as a numerical metric indicating the significance of a word within a document concerning a corpus of documents. It is widely employed for tasks such as document classification, clustering, and information retrieval. C-TF-IDF enhances this concept by incorporating the class (or category) of documents to which a particular document belongs. It computes TF-IDF scores based on the term frequency within a specific class of documents, rather than across the entire document collection.

3.2 Pipeline

The first operation is data pre-processing, which is the cleaning, correction, and normalization of the data in preparation for subsequent processing. In this context, a key step that influences the results of subsequent processes is the careful selection of terms for clustering, similarity assessment, and topic modeling. A number of automatic keyword extraction methods have been explored, ranging from inclusive approaches, such as considering all terms while excluding a list of stopwords, to more refined techniques focusing only on nouns or nouns accompanied by adjectives. In addition, Python libraries such as NLTK and SpaCy have been used for these efforts. Automatic keyword extraction algorithms such as RAKE (Rapid Automatic Keyword Extraction) (Rose et al., 2010) and TextRank (Mihalcea & Tarau, 2004) were also evaluated.

UMAP and HDBSCAN algorithms are used to cluster vectors derived from fine-tuned transformers. Several pre-trained models were evaluated and fine-tuned on the objects of interest to determine the most appropriate model for the given task. In addition, this process includes careful selection of UMAP and HDBSCAN hyperparameters to improve performance on each dataset. The result of this task is item clustering, which includes both the centroids and the collection of items within each cluster. Topic modeling works on the clustered items by identifying the most salient keywords within each cluster. This process draws on methods from information retrieval, using techniques such as weighting using TF-IDF for each class. Visual representations such as graphs and word clouds are used to present the results in a way that is easy to interpret. These visual aids not only facilitate comparison between datasets, but also enhance understanding of the content within each dataset. Furthermore, to ensure the effectiveness of the methods used, an initial evaluation of the results obtained by users was conducted as part of the validation process. The whole pipeline is depicted in Table 1.

Table 1. Pipeline of the proposed approach.

# Task 1: Dataset Preparation	
-	Preprocessing (possibly strip stopwords, accents, …)
-	Process data to extract items to be used
-	Output: items of interest
# Task 2: Items clustering	
-	Choice of transformers and pre-trained models
-	Fine tuning of pre-trained Bert-like models to obtain the vectors
-	Choice of hyperparameters for UMAP and HDBSCAN
-	Output: centroids of clustered items, and elements of each cluster
# Task 3 Topic Modelling and Data Visualization	
-	Computation of C-TFIDF for the elements of each cluster
-	Creation of word clouds and topic modeling visualization
-	Output: topic modeling visualization for comparison of datasets
-	Preliminary evaluation of the results with domain experts and web users

4 The Experimentation

4.1 The Dataset Preparation

English-language cookbooks from Archive.org and other public repositories were used in this experiment to examine the evolution of ingredients, cooking techniques, and food combinations across different time periods. The analysis of these cookbooks aimed to gain insights into how culinary practices have changed over time, including shifts in ingredient preferences, innovations in cooking methods, and the emergence of new food pairings. Several methods of keyword extraction were used, including:

- All 1-gram excluding stopwords, top 1-gram, two-grams, three-grams, and 1-to-3 grams. The algorithms split the text considering all 1, 2 and 3 grams and then take the most frequent ones: this causes the average number of keyword or keyphrases to be rather higher than the original number of words.

- TF-IDF scored words: TF-IDF serves as a metric to measure the importance of a term within a document relative to a larger collection of documents. Terms with higher TF-IDF values are considered more important.
- TextRank is an unsupervised graph-based algorithm inspired by PageRank and is primarily used for keyword and keyphrase extraction. TextRank considers text units (words or phrases) as vertices in a graph and computes the importance of each unit based on its connectivity to other units in the text.
- RAKE is a domain-independent, unsupervised technique for extracting keywords and keyphrases from text documents. It relies on simple heuristics, including word frequency, co-occurrence, and the presence of stop words, to identify candidate keywords.
- Noun phrases often encapsulate meaningful information within a sentence or document. Noun phrase extraction techniques focus on identifying and extracting these phrases as keywords. We have used both home-made methods as well as methods based on Python packages (SpaCy, Pattern and NLTK).

The experiments conducted in this study focused on analyzing recipe books from both the last century and the current century. These books represented a variety of culinary styles, including traditional Italian cuisine and healthy cooking. Table 2 provides an initial overview of the data. Each book page is considered as a document, to be clustered. This table includes details such as, the average number of terms found per page, and the total number of terms. These counts are based on all words present in the recipes, excluding common stop words. In addition, we also present counts for word combinations ranging from single words to three-word phrases (1 to 3 grams).

4.2 Clustering

We use pre-trained language modelling in conjunction with UMAP and HDBSCAN to cluster data. BERT and other Bert-like models have been used to transform text data into high-dimensional vectors that capture semantic meaning. We then apply UMAP to the vectors to obtain a lower dimensional space that is the input to HDBSCAN for grouping similar text. This approach has proven to be very effective in this context. Several tests were performed to evaluate the best value.

Only one parameter, n_neighbors, was evaluated for UMAP. This parameter determines how UMAP balances local and global data structures. The default value in the Python implementation is 15. Our tests assessed the following values: 20, 15, 10, and 5. For HDBSCAN, we evaluated several min_cluster_size and min_samples values. Min_cluster_size is set intuitively to the smallest cluster size for which a cluster should be considered. It is important to consider the value of min_samples in conjunction with the parameter being evaluated, as it provides insight into the level of conservatism in the grouping. The values tested for min_cluster_size (HDBSCAN) were 15, 10, and 5, while for min_samples (HDBSCAN) they were 15, 10, 5, and 1. Additionally, pre-trained transformer models were also evaluated.

Depending on the values of the hyperparameters and the methods used for term extraction in clustering and topic modeling, the number of clusters for the entire dataset can vary from 2 to 41. A low number of clusters can lead to a loss of detail and the merging of heterogeneous groups, while a high number of clusters can enable a precise discrimination of different aspects of the data. However, there is also a risk of losing the overview of individual datasets and may result in insignificant or overlapping clusters.

Evaluation using the Silhouette Score can help identify both the optimal number of clusters and the appropriate term extraction method. The silhouette score measures the consistency of clusters and the separation between them, providing insight into the quality of clustering by assessing how well each observation fits into its cluster relative to other clusters. In addition, the Calinski-Harabasz score serves as a measure of cluster cohesion versus inter-cluster separation. A higher score indicates greater intra-cluster cohesion and greater inter-cluster separation, suggesting better cluster structure.

Meta, Google, OpenAI, and other organizations have been developing and training various architectures and language models for various tasks. Every month (even every day), new models are introduced or updated, each with an expanding array of hyperparameters and trained in multiple languages. In this context, we used large language models specifically trained for clustering and classifying semantic similarity, focusing on the English language. We tested sentence-transformers/bert-base-nli-mean-tokens, sentence-transformers/all-MiniLM-L6-v2, sentence-transformers/bert-base-wikipedia-sections-mean-tokens. All models are taken from the Hugging Face model repository, which constantly updates them[1].

Table 2. Some information on the datasets used in the experimentation.

id	**orig_text terms**	**Avg orig_text**	**1-gram**	**Avg 1-gram**	**1-to-3 grams**	**Avg 1-to-3 grams**
1 - Healthy dinner	16531	183,6	12305	136,7	67330	748,1
2 – Family cookbook	18445	249,2	11453	154,7	62415	843,4
3 – Traditional Italian recipes	20503	320,3	16559	258,7	89184	1393,5
4 – Italian recipes - 1919	13571	135,7	11362	113,6	64208	642,08
5 – Modern Italian Recipes	32081	344,9	23075	248,1	126377	1358,8

The table 3 shows the results obtained by adjusting the hyperparameters, sorted by decreasing values of the Calinski-Harabasz score. Optimal results tend to cluster around 7-8 clusters. The figures in the following pages illustrate the results obtained with the parameter configuration specified in row 1: n_neighbors=10, min_cluster_size=5, using token extraction methods ranging from single words to three-word phrases (1 to 3 grams). The Large Language Model (LLM) used is Microsoft MiniLM-L6-v2.

[1] https://huggingface.co/models last consulted on 20th March 2024

4.3 Topic Modelling and Data Visualization

After clustering the documents, we extract the keywords for each cluster and create a word cloud visualization. To better understand the differences between datasets, we compare the words that are present in both clusters and those that are exclusive to one cluster and display them as word clouds. These visualizations are either based on simple absolute frequency or use the C-TF-IDF algorithm, which gives more weight to terms appearing within the cluster compared to the entire set of documents. This approach provides a more nuanced understanding of the semantic characteristics and differences between clusters, facilitating insightful analysis and interpretation.

Figure 1 illustrates the composition of clusters for each dataset, using the 10-10-5 parameters with 1 to 3 grams, in bold in Table 3. At first glance, it's clear that each dataset is predominantly characterized by one primary cluster. It's worth noting that datasets 3 and 4 both pertain to traditional Italian cuisine. They share cluster number 7, indicating similarity in their culinary themes. Conversely, the primary cluster for dataset 5 is cluster 1, delineating its distinct focus within the culinary domain. In addition, clusters 4 and 2 are the primary clusters for datasets 1 and 2, respectively.

Table 3. Number of clusters obtained as the values of n_neighbors, min_cluster_size and min_samples vary, sorted by decreasing values of the Calinski-Harabasz score.

Keyphrases	**n_neighbors**	**min_cluster_size**	**min_samples**	**Clusters_n**	**silhoutte**	**Calinski harabasz**
1-to-3_grams	**10**	**10**	**5**	**8**	**0,714**	**5767,777**
1-to-3_grams	10	10	1	9	0,673	5483,703
1-to-3_grams	10	5	1	11	0,663	4503,167
1-to-3_grams	10	5	5	10	0,653	4434,171
onebigram	15	10	1	8	0,591	4141,500
bigram	15	10	1	6	0,725	3973,755
1-to-3_grams	15	10	1	8	0,643	3937,788
top_unigram	15	10	1	7	0,670	3656,074

In the process of clustering and topic modeling, keywords or keyphrases from 1 to 3 grams are extracted using different methods as described in Section 4.1. Then, for each generated topic, the most significant or frequent tokens are selected to create word clouds for better visualization.

Figures 2 and 3 show the most significant terms in Topic 7, measured both by frequencies within the topic and using a C-TF-IDF based weighting method, which highlights the most significant terms in the topic relative to the entire dataset. In comparison, the C-TF-IDF version provides clearer insights, albeit with a smaller set of words used, facilitating a more complete understanding of the topic. The all-words version, on the other hand, provides a better overview of the terms used.

Figures 4 and 5 provide insight into the most prevalent terms specific to dataset 4 that do not occur in dataset 5 and vice versa. In dataset 4, prominent terms such as "butter," "veal," and "grated cheese" are observed, which are commonly found in traditional Italian cuisine. Conversely, in dataset 5, alongside abbreviations like "tablespoon" or "teaspoon" terms such as "heat" and "dough" are notable. These differences highlight distinct ingredient usage and cooking instructions between the two datasets.

5 Conclusions and Future Works

This work presents a prototype for comparing datasets using topic modeling, which is a method based on machine learning and pre-trained language models. The final goal is to offer a simple and unsupervised way to evaluate the overlap or distance of topics covered in different datasets. Users can explore and navigate the data in an engaging and intuitive way through a visual representation of the dataset's contents.

This prototype uses topic modeling of historical culinary recipes to identify the evolution of cuisines, cooking techniques, ingredient use, and cultural influences over time. Historians, anthropologists, and curious users have all praised the use of topic modeling on historical culinary recipes to illustrate the evolution of food cultures. Further experiments on a wider range of recipes and a more significant period of time are necessary to understand the socio-cultural dynamics that shape human diets and culinary practices. Experiment on other Cultural Heritage contexts is also foreseen.

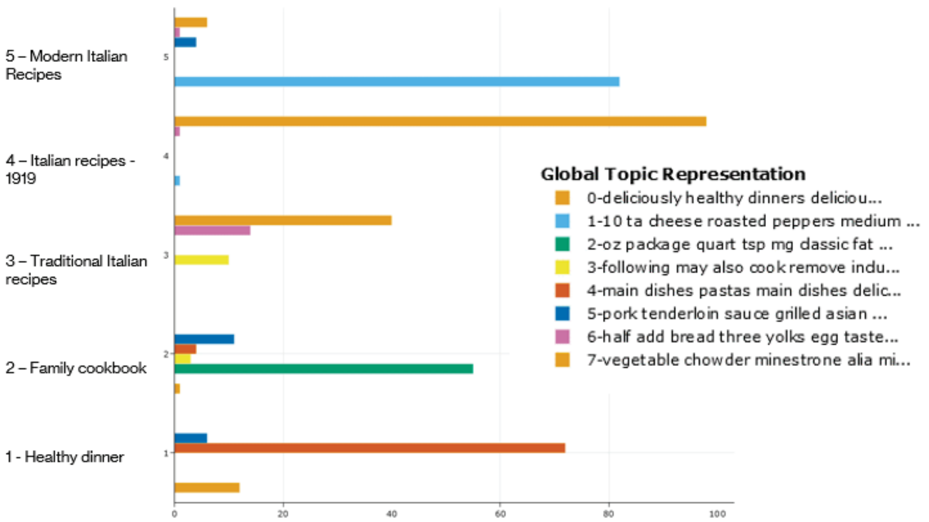
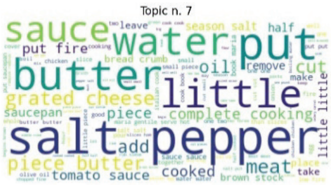


Fig. 1. Topic representation for each dataset.



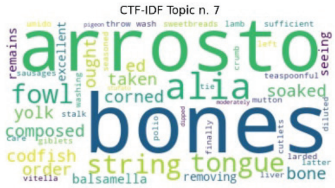


Fig. 2. and **3.** Topic 7 significant words, based on frequency (fig2) and C-TF-IDF weights (fig3).

The prototype received a positive qualitative assessment for both the produced clusters and overall word clouds, evaluated by topic and across different datasets. Future developments will compare datasets on diverse topics and languages. Further studies are needed to evaluate the entire pipeline comprehensively and quantitatively. Our goal is to broaden the testing of this prototype, which is currently limited to books, to encompass the entire archives. This expansion will enhance the efficiency and effectiveness of the system, aiding users in better understanding the contents of the archives.

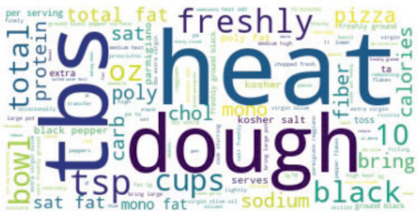


Fig. 4. and **Fig. 5.** Significant words of dataset 4 not in dataset 5 (Fig. 4) and vice versa (Fig. 5).

References

Angelov, D. (2020). *Top2Vec: Distributed Representations of Topics*. arXiv. http://arxiv.org/abs/2008.09470

Artese, M. T., & Gagliardi, I. (2023). Unsupervised Creation of Semantic Graphs to Navigate Intangible Cultural Heritage Using Transformers. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, *13*, 137–148.

Campello, R. J. G. B., Moulavi, D., & Sander, J. (2013). Density-Based Clustering Based on Hierarchical Density Estimates. In J. Pei, V. S. Tseng, L. Cao, H. Motoda, & G. Xu, *Advances in Knowledge Discovery and Data Mining* (Vol. 7819, pp. 160–172). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-37456-2_14

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* arXiv. https://doi.org/10.48550/arXiv.1810.04805

Egger, R., & Yu, J. (2022). A topic modeling comparison between lda, nmf, top2vec, and bertopic to demystify twitter posts. *Frontiers in sociology*, *7*, 886498. https://doi.org/10.3389/fsoc.2022.886498
Gagliardi, I., & Artese, M. T. (2023). Intuitive Semantic Graph Tool for Enhanced Archive Exploration. In *Proceedings of the 34th ACM Conference on Hypertext and Social Media (HT '23)* (Article 11, 1–3). Association for Computing Machinery, https://doi.org/10.1145/3603163.3609069
Grootendorst, M. (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure.* arXiv. http://arxiv.org/abs/2203.05794
Li, Y. (2023). *Insights from Tweets: Analysing Destination Topics and* Sentiments*, and Predicting Tourist Arrivals* [PhD Thesis, Durham University].
McInnes, L., Healy, J., & Astels, S. (2017). hdbscan: Hierarchical density based clustering. *J. Open Source Softw.*, *2*(11), 205. https://doi.org/0.21105/joss.00205
McInnes, L., Healy, J., & Melville, J. (2020). *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction.* arXiv. https://doi.org/10.48550/arXiv.1802.03426
Mihalcea, R., & Tarau, P. (2004). Textrank: Bringing order into text. In *Proceedings of the 2004 conference on empirical methods in natural language processing* (pp.404–411). https://aclanthology.org/W04-3252
Ogunleye, B., Maswera, T., Hirsch, L., Gaudoin, J., & Brunsdon, T. (2023). Comparison of topic modelling approaches in the banking context. *Applied Sciences*, *13*(2), 797. https://doi.org/10.3390/app13020797
Rose, S., Engel, D., Cramer, N., & Cowley, W. (2010). Automatic Keyword Extraction from Individual Documents. In M. W. Berry & J. Kogan, *Text Mining* (pp. 1–20). Wiley. https://doi.org/10.1002/9780470689646.ch1
Sprenkamp, K., Zavolokina, L., Angst, M., & Dolata, M. (2023). Data-Driven Governance in Crises: Topic Modelling for the Identification of Refugee Needs. In *Proceedings of the 24th Annual International Conference on Digital Government Research (dg.o '23)* (pp. 1–11). https://doi.org/10.1145/3598469.3598470
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In I. Guyon, U. von Luxburg, S. Bengio, H. M. Wallach, R. Fergus, S. V. N. Vishwanathan & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA* (pp. 5998--6008). https://doi.org/10.48550/arXiv.1706.03762
Wojciechowska, J., Sypniewski, M., Śmigielska, M., Kamiński, I., Wiśnios, E., Schreiber, H., & Pieliński, B. (2023). *Deep Dive into the Language of International Relations: NLP-based Analysis of UNESCO's Summary Records.* arXiv. https://doi.org/10.48550/arXiv.2307.16573

Received: March 15, 2024
Reviewed: April 16, 2024
Finally Accepted: May 10, 2024