Design and implementation of a text mining-based tool to support scoping reviews

Filippo Chiarello*

B4DS Research Group,
School of Engineering,
University of Pisa,
Largo Lucio Lazzarino, 56122 – Pisa, Italy
Email: filippo.chiarello@unipi.it
*Corresponding author

Luca Gastaldi

Department of Management, Economics and Industrial Engineering,
Politecnico di Milano,
Via Lambruschini 4b, Building 26b, 20156 – Milan, Italy
Email: luca.gastaldi@polimi.it

Antonella Martini

B4DS Research Group,
School of Engineering,
University of Pisa,
Largo Lucio Lazzarino, 56122 – Pisa, Italy
Email: antonella.martini@unipi.it

Abstract: Among literature reviews, scoping review is a relatively new approach that is increasingly gaining popularity since it helps researchers in defining emerging and multidisciplinary fields. While artificial intelligence for text processing can help researchers in this sense, we still lack clear procedures and tools to improve the reviewing process. Following a design science approach, in this article we propose a novel tool based on natural language processing (NLP) to support scoping review and to visualise its results. The tool (NLP4Scoping) is implemented using open-source software and is made available for reuse on GitHub. Each phase for its proper application is described focusing on the nascent literature stream on innovation management in digital ecosystems.

Keywords: scoping review; natural language processing; NLP; topic modelling; open source.

Reference to this paper should be made as follows: Chiarello, F., Gastaldi, L. and Martini, A. (2023) 'Design and implementation of a text mining-based tool to support scoping reviews', *Int. J. Technology Management*, Vol. 91, Nos. 3/4, pp.147–161.

Biographical notes: Filippo Chiarello received his PhD in Engineering Management from the School of Engineering, University of Pisa, Italy, in 2019. He is currently an Assistant Professor with the School of Engineering, University of Pisa. His research focuses on natural language processing techniques for studying technological and HR-related phenomena.

Luca Gastaldi is an Associate Professor at the School of Management of Politecnico di Milano where he teaches 'organisation design', 'business process management' and 'leadership and innovation'. His research interests are in the fields of digital innovation, smart working and design thinking, with a peculiar emphasis on public entities. He is a board member of CINet, an international research network on continuous innovation. He is also the Director of 'Digital Agenda', 'Digital Identity', 'Design Thinking for Business' and 'Smart City' Observatories at Politecnico di Milano. Since November 2021, he serves as an expert on digital topics in the Technical Secretariat of the Presidency of the Council of Ministers of the Italian Government for the realisation of the Italian National Recovery and Resilience Plan.

Antonella Martini is a Full Professor of Business Economics and Strategic Analysis at the School of Engineering, University of Pisa. She is also the President of CIMEA, the Centre for Academic Mobility and Equivalence. Her research interests include organisational ambidexterity, competitive intelligence, and strategic foresight. She has published in, among others: *California Management Review*, *Technological Forecasting and Social Change*, *Long Range Planning*, *Journal of Business Research* and *International Journal of Management Reviews*.

1 Introduction

Among literature reviews, scoping reviews is a relatively new approach that is increasingly gaining visibility since it helps researchers in defining emerging and multidisciplinary fields (Munn et al., 2018). A scoping review has the aim of determining the scope of a body of literature on a given topic, to give an indication of the volume of documents available as well as an overview of its focus. It is particularly useful in the following cases (Munn et al., 2018):

- identify and classify available evidence in the field
- clarify key concepts in the field, and give definitions
- list the used research approaches
- identify knowledge gaps.

In all the other cases, (e.g., guide decision making, confirm current practices, inform for future research or investigate conflicting results), researchers can rely on other approaches for literature review.

Currently there exists little guidance for researchers regarding

1 how to carry on a scoping review
2 how to use recent technologies such as natural language processing (NLP) to support this process.

As for point (1), scoping reviews (Munn et al., 2018) follow a defined and structured approach to determine the coverage of a field of study, indicating information such as the volume of literature available in the field, its focus, how research is conducted, and which are the main gaps to be filled (Armstrong et al., 2011). Even if scoping reviews are newer with respect to systematic literature reviews, they increasingly represent a common approach for mapping broad topics. Furthermore, because of variability in their adoption, there exists a need for a methodological standardisation (Pham et al., 2014).

One of the most used frameworks to conduct scoping reviews has been proposed by Arksey and O'Malley (2005). The approach includes the following five phases:

1 identifying the research question

2 identifying relevant studies

3 study selection

4 charting the data

5 collating, summarising, and reporting the results.

As reported by Pham et al. (2014), although Arksey and O'Malley (2005) framework is the most frequently used, about 50% of the published reviews did not use any framework.

As for point (2), the increased creation of knowledge, and the complexity of its structure, calls for supporting tools to help researchers in reviewing the scope of a literature domain. This process is creating interesting opportunities for artificial intelligence, to support scientists in their activity and make them more efficient and reproducible. AI tools, among other things, enable the reduction of authors' efforts in time-consuming and repetitive tasks. In this way, researchers can dedicate more time to the creativity-intensive tasks, such as interpretation, intuition, and contextualisation (Tsafnat et al., 2014).

In several fields, artificial intelligence is starting to change traditional research methods (Wagner et al., 2021). Literature review activities stand out in this setting because they work with vast, quickly expanding amounts of partially organised data. In general, AI may speed up parts of the literature review process, but AI utilisation in this context is still in its early stages of development.

Recently, Wagner et al. (2021) have developed a framework that classifies AI tools for supporting literature review, considering each step of the review process. These steps are:

- *Problem formulation:* identify the research questions, central concepts (Templier and Paré, 2018), or specify the research gap to verify (Müller-Bloch and Kranz, 2015).
- *Search:* collect the relevant literature using different search methods, such as database or citations searches, table-of-content scans, and complementary manual searches (Templier and Paré, 2018).
- *Screen:* filter for relevant papers, based on some pre-defined of relevancy for the topic (Harrison et al., 2020).
- *Quality assessment:* to check studies for methodological consistency (Higgins and Green, 2008; Kitchenham and Charters, 2007).

- *Data extraction:* identify relevant of data (qualitative and quantitative) and collect them for further analysis (Jonnalagadda et al., 2015).
- *Data analysis and interpretation:* can take various forms depending on the type of review but aims at synthetising the results.

Despite the work of conceptualisation done by Wagner et al. (2021), there still exists a lack of contributions that shows how to use AI tools to analyse the textual content of the papers, and that arrive to a re-usable output for other researchers.

Recently, NLP has proven to be an effecting AI tool to support literature reviews (Chiarello et al., 2021; Kobayashi et al., 2017; Belingheri et al., 2021).

The purpose of this article is to propose a novel NLP tool to support scoping review. The tool, named NLP4Scoping, is implemented using open-source software and is made available for reuse on GitHub. Each phase for its proper application is described focusing on the nascent literature stream on innovation management in digital ecosystems.

The article is structured as follows. In Section 2, we discuss the requirement of this tool, under the lens of design science. In Section 3, we disclose the design of NLP4Scoping tool and explain its functionalities in a real-world scenario for the analysis of the literature on digital ecosystems.

2 Requirement analysis

Considering the state of the art, it emerges the need for a tool to support researchers in the implementation of scoping reviews. A wide literature exists on the process to be adopted to design it and the main contribution comes from information systems (IS) discipline. Hevner et al. (2004) argues that, in order to apply this knowledge to solve real world problems, it is important to rely on the paradigm of design science.

The design-science paradigm is a problem-solving paradigm and has its roots in engineering and the sciences of the artificial (Simon, 1988). Its goal is to create innovations: define the ideas, practices, technical capabilities, and products through which the analysis, design, implementation, management, and use of IS can be effectively and efficiently accomplished. Hevner et al. (2004) provide a set of guidelines which help IS researchers conduct, evaluate and present design-science research and to define the requirements to guide the development of the artefact. In the case of NLP4Scoping tool, the requirements have been identified following these guidelines.

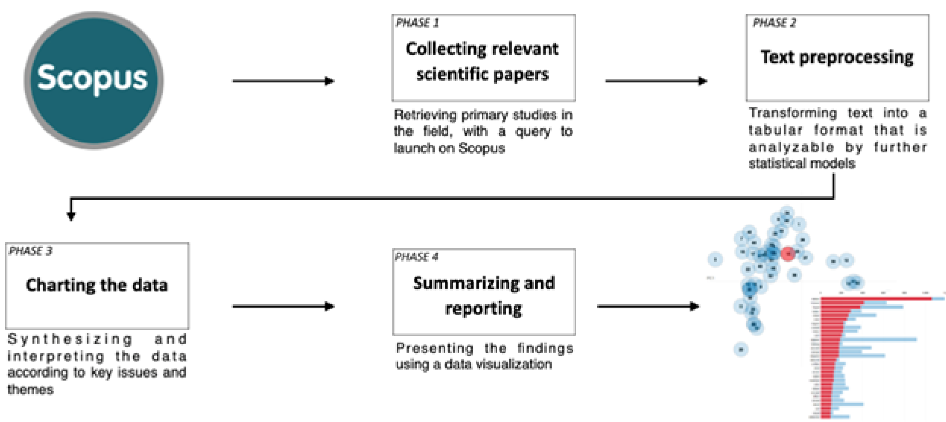
The set of requirements are:

- *Usability:* design science research must create a viable artefact in the form of a construct a model, a method or, in our case, a software. The software must adapt to users' skills and knowledge's.
- *Relevancy:* the tool must solve a real need emerging from the state of the art. In our case, the application of the tool must let the user to gain new knowledge and to support the scoping review process.
- *Novelty:* the tool must solve an unsolved problem or solving a known problem in a more effective or efficient manner. Thanks to the mix of NLP and interactive visualisation of results, our tools is novel with respect to the state of the art in

scientific literature analysis and has different goals with respect to widely used tools such as Scopus Search Analytics[1], or WoS viewer.

- *Scientifical solidity:* the tool must rely on scientific evidence and practices. For this reason, we relied on the theory of latent Dirichlet allocation (LDA) (Blei et al., 2003) and on the evaluation of the quality of unsupervised learning process (Deveaud et al., 2014) to provide the user with quantitatively validated results.
- *Communicability:* the results must be presented effectively to technology-oriented audience (research) as to managerial audience. For this reason, we describe the tool in the present paper.
- *Accessibility:* the tool has been developed using only open-source software. Also, we have released the tool on git-hub, with further guidelines for its application.
- *Elasticity:* the tool can take as input data coming from different scientific articles database. In the present paper we consider Scopus, but similar results can be achieved using other sources such as WoS, or private databases.

Figure 1 Workflow (see online version for colours)



3 Designing the tool

We followed a slightly modified version of a widely used framework in the context of scoping reviews (Arksey and O'Malley, 2005). The four stages (Figure 1) are described in further detail in the next sections. Each phase for its proper application is described focusing on the nascent literature stream on innovation management in digital ecosystems.

3.1 Collecting relevant scientific papers

The goal of the first stage is to identify and collect articles that discuss innovation in digital ecosystems. First, we collected all the scientific papers (journal and conference papers), published between 2000 and 2020 and containing the sequent string in the title, abstract or keywords:

‘innovation AND (ecosystem OR platform) AND (digital OR digitalisation OR ‘artificial intelligence’ OR ‘big data’ OR ‘data science’ OR blockchain OR ‘cloud computing’ OR ‘augmented reality’ OR ‘virtual reality’ OR ‘Embedded Systems’)’.

As it is evident, the query is the intersection of three set of documents: innovation related documents, ecosystems related documents and digital technologies related documents. For this last set, we selected the technological keywords considering previous literature on the topic (Chiarello et al., 2018).

We decided to search the documents in the Scopus database. This approach, executed in July 2021, led to the collection of 1,115 scientific papers. The search was constrained to the last 20 years since even if DS has sub-fields, such as AI, that have roots in the early 1960s (Minsky, 1961), the specific field has emerged only in the early 2000s (Cleveland, 2001).

3.2 Text pre-processing

Text pre-processing aims at transforming text into a tabular format that is analysable by further statistical models. In the present section, we describe the pre-processing chain we applied to the title and abstract of the selected papers. These steps have been performed using the R package udpipe (Straka and Straková, 2017).

3.2.1 Data preparation

We prepared the data following these steps:

- *Tokenisation:* splitting the text in tokens (single words).
- *Speech tagging:* tagging words considering their morphological role and morphosyntactic context; for instance, if the token is a determiner, the next token is a noun or an adjective with very high confidence.
- *Lemmatisation:* determining the dictionary root of a word; the output allows finding if two words share the same root, despite their surface differences; lemmatisation was preferred over stemming considering the interpretability of the output; in fact, lemmatisation attempts to return the lemma or dictionary form of a word, while stemming returns the stem, (e.g., both methods may return ‘work’ for ‘working’, but only lemmatisation correctly outputs ‘good’ from ‘better’).

3.2.2 N-grams extraction

An n-gram (or multi-word) is a sequence of words that has a meaning that is different from the single words (e.g., ‘credit card’, ‘machine learning’). Identifying multiword is crucial for scientific literature analysis since most of the technical jargon is made of multi-words. Multi-word extraction is usually implemented with statistics and linguistic rules, thus using the statistical properties of n-grams, or machine learning approaches (Newman et al., 2012). However, in this paper we rely on keywords identified by Scopus to take multiword in texts, as accomplished by Mazzei et al. (2021).

3.2.3 Feature selection

The entire text corpus comprised 250,704 words and multi-words. We applied a series of words unification and removal steps, to merge words with similar meaning and to select only the relevant words for each abstract (i.e., features). These steps have proven to be effective in several similar works (Chiarello et al., 2019; Cascini et al., 2013).

We thus removed:

- Sparse terms occurring in less than 0.1% of all documents.
- Common terms occurring in more than 10% of all documents.
- List of domain-related stop-words (e.g., abstract, paper, research).
- Short words, composed of less than three characters.
- Every part of speech that does not belong to nouns and adjectives.

At this stage, the entire text corpus comprised 54,215 words and multi-words. Finally, output data were structured as a document-term matrix comprised of 1,115 documents and 3,537 (unique) terms. This matrix is the input of the following topic modelling step.

3.3 Charting the data

Following this, we implemented a charting phase, where data were collected in a database using the statistical software RStudio. The database is composed of 1,115 scientific papers and has the following structure: reference number, Scopus ID, authors, year of publication, name of the journal or of the conference, title, paper type (journal or conference paper), number of citations and topic probability vector. To identify the relevant topics within the paper set, we use the LDA (Blei et al., 2003). The primary assumption behind LDA is that the topics have a sparse Dirichlet prior distribution (Ng et al, 2011). This assumption is supported for abstracts of papers, since these documents are likely to cover only a small set of topics. This method has already been proven to work in texts as scientific abstracts (Amami et al., 2016).

LDA computes two relevant values: alpha and beta. Given a set of D documents represented by T different tokens (words and multi-words) and chosen K as the number of topics, these two values can be defined as:

- *Alpha:* maps the topics on the documents; it indicates for each of the K topics, which is the probability that a document belongs to it; thus, a K-dimensional vector is computed for each document.
- *Beta:* maps the words on the topics; given a topic, indicates for each of the T tokens which is the probability that the considered topic contains each token; thus, a T-dimensional vector is computed for each topic.

The ideal number of clusters K is unknown a priori, as is the case with many clustering tasks, and must be calculated either by some measure judging the grouping's quality or through domain knowledge. In the first case, the Kullback divergence of the salient distributions derived from the LDA model's factorisation matrices (Arun et al., 2010), the density-based approach to maximise intra-cluster similarity and minimise inter-cluster dissimilarity (Cao et al., 2009), the Jensen-Shannon divergence of the topic distribution

(Deveaud et al., 2014), and the maximum likelihood estimator (Griffiths and Steyvers, 2004) were all tested. There isn't a single statistic that works in every situation. As a result, all four criteria stated above were employed, and K was chosen as the point at which all, or almost all, of the metrics agreed.

3.4 Summarising and reporting

Topic modelling algorithms gives a rich output in terms of topics-documents relations and topics-words relations.

The first group opens to a series a visualisation and analysis of the results, where topics can be linked to meta-information on the paper, such as year if publication and journal. This can give insight on how the topics are distributed over time, and how different sources are more (or less) focused on specific topics. To this aim, static graphs such as bar-plots and scatter plot are used.

Topics-words relations are useful to have a focused view on the content of the topic. The visualisation is useful to understand the content of each topic and how they are related to each other. Since the content is particularly reach, to summaries this report we used a Shiny-based interactive interface for exploring the output from LDA topic models (Sievert and Shirley, 2014).

We used an open-source software, LDAvis (Sievert and Shirley, 2014), to interactively explore the results of the scoping review. LDAvis is an interactive and web-based visualisation tool built using a combination of R and D3 that provides a bird-eye view of the whole results of the topic model, showing similarities and dissimilarities between the topics. Also, it makes it possible to explore the terms most highly associated with each individual topic. The visualisation system allows users to flexibly explore topic-term relationships using relevance to better understand a fitted LDA model.

4 The visualisation tool

The visualisation tool (illustrated in Figure 2) has two basic zones of analysis.

The left panel presents a global view of the topic model. Here the topics as circles in two-dimensional space where the position of the circles are determined by computing the distance between each topic, and then by using multidimensional scaling to project the inter-topic distances onto two dimensions, as is done in Chuang et al. (2012). The area of the circle encodes topics' prevalence. Also, the topics are numbered in decreasing order of prevalence.

The right panel of the visualisation tool shows a bar chart that represent the individual terms. The terms are at the top-30 in terms of saliency, and are the most useful for interpreting the currently selected topic on the left. This visualisation can help users in understanding the meaning of each topic, supporting the process of labelling it with the most representative words. Each overlaid set of bars represents both the corpus-wide frequency of a given term as well as its topic-specific frequency, as in Chuang et al. (2012).

Figure 2 The layout of LDAvis (see online version for colours)

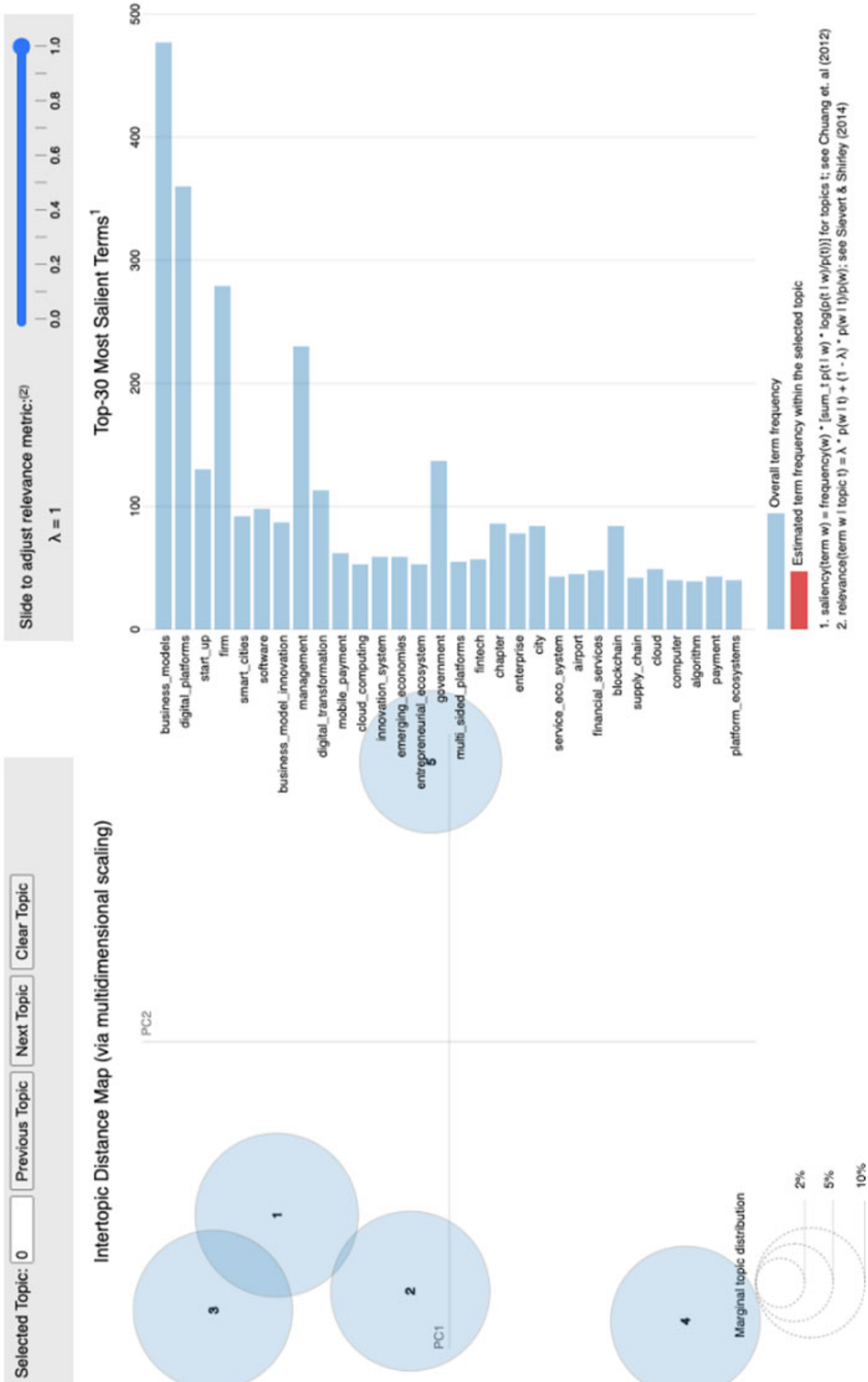


Figure 3 The layout of LDAvis (topic 1 selected) (see online version for colours)

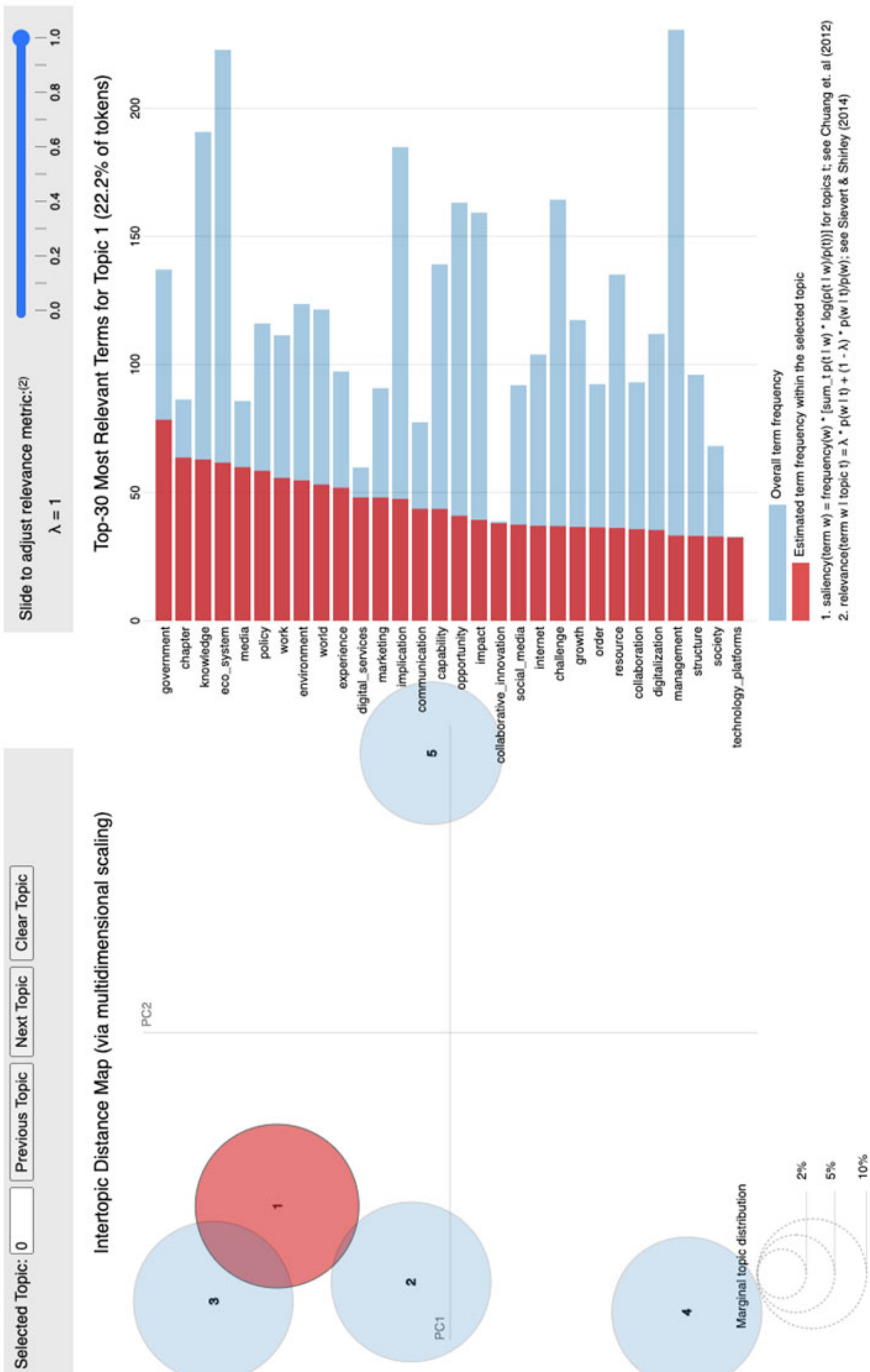
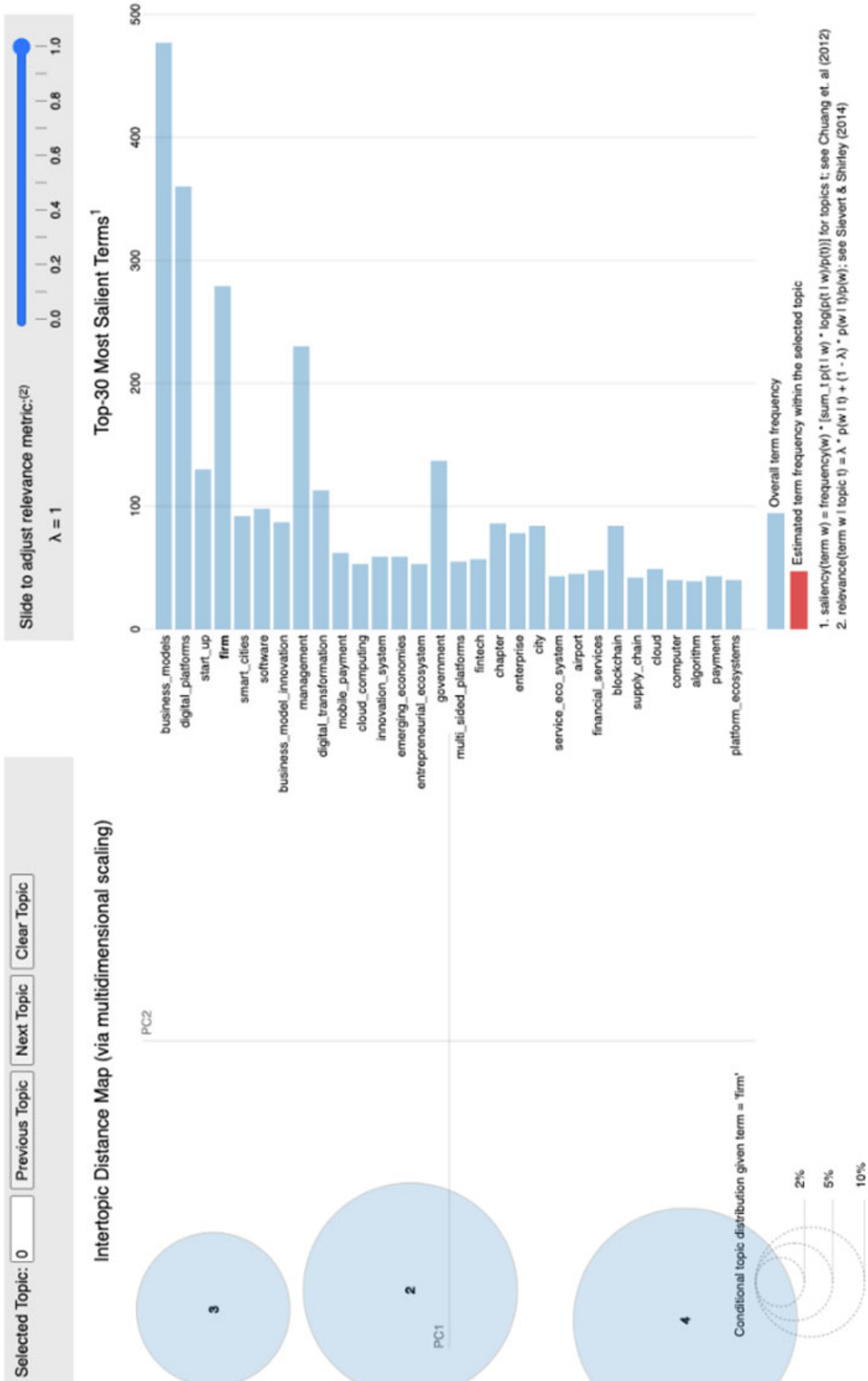


Figure 4 The layout of LDAvis (the word 'firm' selected) (see online version for colours)



The left and right panels are linked. If the user selects a topic on the left, the tool adapts to reveal the most useful terms on the right. Also, if the user selects a term on the right, the left panel shows the conditional distribution over topics for the selected term.

As said, a first important functionality of the LDAvis is to give the possibility, once a topic is selected, to explore the most relevant terms that are related to that specific topic. In Figure 3, topic 1 is selected, and its 30 most relevant terms populate the bar chart to the right (ranked in order of relevance from top to bottom). The widths of the blue bars represent the corpus-wide frequencies of each term; the widths of the red bars represent the topic-specific frequencies of each term. By comparing the widths of the red and grey bars for a given term, users can quickly understand whether a term is highly relevant to the selected topic because of its lift (a high ratio of red to grey), or its probability (absolute width of red). The top three most relevant terms in Figure 3 are 'government', 'chapter', and 'knowledge'. Also, a slider allows to change the value of delta, which can alter the rankings.

On the left panel, two visual features provide a global perspective of the topics. First, the areas of the circles are proportional to the relative prevalence of the topics in the corpus. In the five-topic model fit to the data, the first three topics comprise about the 30%, of the corpus, and all contain common, non-specific terms. In addition to visualising topic prevalence, the left pane shows inter-topic differences.

The second core feature of LDAvis is the ability to select a term on the bar chart to reveal its distribution over topics. This distribution is visualised by altering the dimensions of the circles such that they are proportional to the term-specific frequencies. This allows the user to verify if the multidimensional scaling of topics has faithfully clustered similar topics (i.e., containing similar words) in two-dimensional space. For example, in Figure 4, the term 'firm' is selected. As it can be seen, the left side of the left-panel contains all the topics that talks about 'firm'. This means that the left area of the two-dimensional space is linked to this concept. Upon inspection, this group of topics can be interpreted broadly as a discussion of companies and business. This verifies, to some extent, their placement, via multidimensional scaling, into the same two-dimensional region.

Finally, it is worth mentioning that the output of the visualisation tool can be easily shared by creating an HTML page. We make available our results, to let other researchers explore the tool and our results (https://bl.ocks.org/FilippoChiarello/raw/f0d070021b7e4218e072b46d8536e918/?raw=true).

5 Conclusions

The purpose of the presented article is to propose NLP4Scoping, a novel tool to support the scoping review process thanks to the use of NLP techniques and interactive data visualisation tools. The requirements of the tools have been developed following a design science approach. Each phase for the proper application of the tool has been described, to guide researchers in the application of the tool in any field of research. Also, we used the literature of innovation in digital ecosystem as a context of application. The code and the tool are made available on GitHub for reuse (https://github.com/FilippoChiarello/scientific-paper-analysis).

The work presents some limitations that can be mitigated by future research. First, we are aware that more advanced and modern NLP techniques could lead to better classification of the results. For instance, other approaches for topic modelling (e.g., Peinelt et al., 2020) can be tested. Also, future activities can focus on the measuring and improving the usability of the user interface, using approaches such as A/B testing to understand if and which different user interfaces and data visualisation of the tool can improve its usability.

Finally, the tool can show more information about the results, mixing the topic modelling with the metadata of the paper. In this way, the user can have access to relevant information, such as the number of papers published for each topic in each year, the distribution of topics on the scientific journals and conferences, the most important authors and universities per topic and the geographical distribution of topics. Interestingly, Elsevier's Scopus API (used in the present paper), offers this information. Researchers and computer scientist wanting to add this information to the tool, have the information already available, that only need to be integrated with the one presented in this paper.

To conclude, some steps of the process, (e.g., paper retrieval, topic validation) need still an important manual effort. Research can focus on automatising these steps to make the process even more efficient. In this way, experts will have more time to work on the most value-added activities (design of the research questions and interpretation of the results).

References

Amami, M., Pasi, G., Stella, F. and Faiz, R. (2016) 'An LDA-based approach to scientific paper recommendation', Paper presented at the *International Conference on Applications of Natural Language to Information Systems*, June.

Arksey, H. and O'Malley, L. (2005) 'Scoping studies: towards a methodological framework', *International Journal of Social Research Methodologies*, Vol. 8, No. 1, pp.19–32.

Armstrong, R., Hall, B.J., Doyle, J. and Waters, E. (2011) 'Scoping the scope of a Cochrane review', *Journal of Public Health*, Vol. 33, No. 1, pp.147–150.

Arun, R., Suresh, V., Madhavan, C.V. and Murthy, M.N. (2010) 'On finding the natural number of topics with latent Dirichlet allocation: some observations', Paper presented at *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, June.

Belingheri, P., Chiarello, F., Fronzetti Colladon, A. and Rovelli, P. (2021) 'Twenty years of gender equality research: a scoping review based on a new semantic indicator', *Plos One*, Vol. 16, No. 9, pp.1–27.

Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003) 'Latent Dirichlet allocation', *Journal of Machine Learning Research*, Vol. 3, No. 2, pp.993–1022.

Cao, J., Xia, T., Li, J., Zhang, Y. and Tang, S. (2009) 'A density-based method for adaptive LDA model selection', *Neurocomputing*, Vol. 72, No. 7, pp.1775–1781.

Cascini, G., Fantoni, G. and Montagna, F.F.M.P. (2013) 'Situating needs and requirements in the FBS framework', *Design Studies*, Vol. 34, No. 5, pp.636–662.

Chiarello, F., Belingheri, P. and Fantoni, G. (2021) 'Data science for engineering design: state of the art and future directions', *Computers in Industry*, Vol. 129, No. 1,.1–32.

Chiarello, F., Melluso, N., Bonaccorsi, A. and Fantoni, G. (2019) 'A text mining-based map of engineering design: topics and their trajectories over time', in *Proceedings of the Design Society: International Conference on Engineering Design*, Vol. 1, No. 1, pp.2765–2774.

Chiarello, F., Trivelli, L., Bonaccorsi, A. and Fantoni, G. (2018) 'Extracting and mapping industry 4.0 technologies using Wikipedia', *Computers in Industry*, Vol. 100, No. 1, pp.244–257.

Chuang, J., Manning, C.D. and Heer, J. (2012) 'Termite: visualization techniques for assessing textual topic models', in *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pp.74–77.

Cleveland, W.S. (2001) 'Data science: an action plan for expanding the technical areas of the field of statistics', *International Statistical Review*, Vol. 69, No. 1, pp.21–26.

Deveaud, R., San Juan, E. and Bellot, P. (2014) 'Accurate and effective latent concept modeling for ad hoc information retrieval', *Document numérique*, Vol. 17, No. 1, pp.61–84.

Griffiths, T.L. and Steyvers, M. (2004) 'Finding scientific topics', *Proceedings of the National Academy of Sciences*, Vol. 101, No. 1, pp.5228–5235.

Harrison, H., Griffin, S.J., Kuhn, I. et al. (2020) 'Software tools to support title and abstract screening for systematic reviews in healthcare: an evaluation', *BMC Medical Research Methodology*, Vol. 20, No. 7, pp.1–12.

Hevner, A.R., March, S.T., Park, J. and Ram, S. (2004) 'Design science in information systems research', *MIS Quarterly*, Vol. 28, No. 1, pp.75–105.

Higgins, J.P.T. and Green, S. (2008) *Cochrane Handbook for Systematic Reviews of Interventions*, John Wiley & Sons, Ltd., Chichester, UK.

Jonnalagadda, S.R., Goyal, P. and Huffman, M.D. (2015) 'Automating data extraction in systematic reviews: a systematic review', *Systematic Reviews*, Vol. 4, No. 1, pp.78–102.

Kitchenham, B.A. and Charters, S. (2007) *Guidelines for Performing Systematic Literature Reviews in Software Engineering*, EBSE Technical Report.

Kobayashi, V.B., Mol, S.T., Berkers, H.A. et al. (2017) 'Text mining in organizational research', *Organizational Research Methods*, Vol. 21, No. 3, pp.733–765.

Mazzei, D., Chiarello, F. and Fantoni, G. (2021) 'Analyzing social robotics research with natural language processing techniques', *Cognitive Computation*, Vol. 13, No. 2, pp.308–321.

Minsky, M. (1961) 'Steps toward artificial intelligence', *Proceedings of IRE*, Vol. 49, No. 1, pp.8–30.

Müller-Bloch, C. and Kranz, J. (2015) 'A framework for rigorously identifying research gaps in qualitative literature reviews', Presented at the *International Conference on Information Systems*, 16 December.

Munn, Z., Peters, M.D., Stern, C., Tufanaru, C., McArthur, A. and Aromataris, E. (2018) 'Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach', *BMC Medical Research Methodologies*, Vol. 18, No. 1, pp.76–143.

Newman, D., Koilada, N., Lau, J.H. and Baldwin, T. (2012) 'Bayesian text segmentation for index term identification and keyphrase extraction', in *Proceedings of COLING 2012*, Vol. 1, No. 1, pp.2077–2092.

Peinelt, N., Nguyen, D. and Liakata, M. (2020) 'tBERT: topic models and BERT joining forces for semantic similarity detection', in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Vol. 1, No. 1, pp.7047–7055.

Pham, M.T., Rajić, A., Greig, J.D., Sargeant, J.M., Papadopoulos, A. and McEwen, S.A. (2014) 'A scoping review of scoping reviews: advancing the approach and enhancing the consistency', *Research Synthesis Methods*, Vol. 5, No. 4, pp.371–385.

Sievert, C. and Shirley, K. (2014) 'LDAvis: a method for visualizing and interpreting topics', Presented at the *Workshop on Interactive Language Learning, Visualization, and Interfaces*, June.

Simon, H.A. (1988) 'The science of design: creating the artificial', *Design Issues*, Vol. 1, No. 1, pp.67–82.

Straka, M. and Straková, J. (2017) 'Tokenizing, pos tagging, lemmatizing and parsing ud 2.0 with udpipe', Paper presented at the *CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencie*s, July.

Templier, M. and Paré, G. (2018) 'Transparency in literature reviews: an assessment of reporting practices across review types and genres in top IS journals', *European Journal of Information Systems*, Vol. 27, No. 5, pp.503–550.

Tsafnat, G., Glasziou, P. and Choong, M.K. (2014) 'Systematic review automation technologies', *Systematic Reviews*, Vol. 3, No. 1, pp.1–15.

Wagner, G., Lukyanenko, R. and Paré, G. (2021) 'Artificial intelligence and the conduct of literature reviews', *Journal of Information Technology*, Vol. 2, No. 1, pp.68–96.

Notes

1 https://blog.scopus.com/topics/analysis-tools (accessed November 2021).