

# Mapping Discourses in Social Work Research-A Topic Modeling Approach

Markus Eckl<sup>1,\*</sup>, Christian Ghanem<sup>2</sup> and Florian Spensberger<sup>3</sup>

<sup>1</sup>Department of Digital Humanities, University of Passau, Hans-Kapfnger-Straße 14 d, 94036 Passau, Germany

<sup>2</sup>Faculty for Social Sciences, Nuremberg Institute of Technology, Keßlerplatz 12, 90489 Nuremberg, Germany

<sup>3</sup>German Youth Institute, Nockherstraße 2, 81541 München, Germany

## Abstract

Social work researchers are increasingly required to process and use digital information appropriately. The aim of this scientometric study is to demonstrate how to use big data to analyze social work research questions. The study uses topic modeling to automatically extract topics from 42,001 abstracts retrieved from the Social Work Research Database. Some of these topics are used to illustrate the method's potential and to shed light on the core semantic structure of disciplinary social work. Some topics, such as child protection and addiction prevention, have become a focus in the social work literature, while other areas have received progressively less attention. Against the background of the outlined problems regarding the interpretation of automatically generated results, analysis methods like topic modeling should be understood as complementary to the more classical research approaches.

## Publication History:

Received: July 05, 2021

Accepted: July 07, 2021

Published: July 09, 2021

## Keywords:

Big data, Data science,  
Scientometrics

## Mapping Discourses in Social Work Research: A Topic Modeling Approach<sup>1</sup>

Empirical research is playing an increasingly important role in social work. Some scholars claim that this shift toward empirical research has been observed in the United States since the 1970s [2] and in German-speaking countries since the 1980s [3] due to the increased production of empirical knowledge. It is reasonable to assume that the advancing digitization has the potential to strengthen social work's path already taken toward an empirically-based discipline. The social transformations associated with the spread of digital information and communication technologies [4] impact not only social work addressees, professionals, and organizations [5] but also disciplinary social work, as new social phenomena, research questions, and research methodological approaches are emerging. In addition, new research methods and specially generated data sets can be used to address existing questions and bring new insights to light [6]. Although the effects of digitization on professional social work are being discussed extensively [7,8], the effects on disciplinary social work have received minimal attention. Nonetheless, the American Academy for Social Work and Social Welfare declares questions about the processing and appropriate use of the growing body of digital information as a "grand challenge" of social work [9], thereby directing researcher's attention to the challenge of undertaking the new tasks associated with disciplinary social work.

In this paper, the authors will first examine quantitative text analysis in the context of big data<sup>2</sup> and discuss its relevance to social work research. Thereafter, a research project is presented in which 42,001 scientific abstracts of international social work are analyzed. Central topics within this data corpus are identified and can be used for further analyses via the analytical method called topic modeling, which is a specific form of quantitative text analysis. Based on this research project, the added value of quantitative text analysis and its limitations are shown. Finally, the general potential of applied research methods and the extent to which they can be applicable to the context of social work research will be discussed.

<sup>1</sup>This is a translated and modified (with the permission of the publisher) version of a chapter that was recently published in a German handbook on social work and digitalization [1].

<sup>2</sup>The definition of "big data" differs depending on the discipline. We make use of a pragmatic understanding and thus refer to a scope of data that can be compared with conventional methods of analysis and that can no longer be processed; people would need years to simply read the respective texts [10].

## Quantitative text analysis and its relevance to social work research

Texts are probably the most common data basis for research projects in social work. Although some studies use quantitative text analysis [11], texts in the social work context are typically analyzed using qualitative research methods. For many of the research questions examined in the realm of social work, this seems to be useful, for example, when reconstructing the latent contexts of meaning for what is said or written. However, such approaches are limited in situations in which the research questions require the analysis of large databases. Presumably, quantitative methods of analysis will become more important due to the rapidly increasing production of digital texts (e.g., social media data, blogs, case files in organizations, and scientific publications), which enable the possibility of new research questions that can no longer be meaningfully analyzed and answered using the more classical approaches. Even a superficial analysis of texts with the help of common qualitative and common quantitative methods is impossible, for example, if one wants to review topics that are discussed in more than 40,000 scientific publications on social work. It is thus likely that the use of quantitative text analysis will become more important, especially in cases in which human resources will no longer be sufficient to process the amount of text.

As a consequence of increasing data production and its large mass, questions regarding how to approach texts are changing. For example, it is no longer possible to deduce contexts of meaning, as is done when using hermeneutic approaches. Rather, the goal is to structure an unstructured text by means of computer-assisted procedures facilitated by algorithms [12]. One can ask which words occur particularly frequently in a text, and it is a matter of ordering, editing, searching, and sorting the text [13]. The methods of doing so become more complex if the connections between different words and topics are also considered. Such analyses, described in the following section, have rarely been represented in social work, although other

**Corresponding Author:** Markus Eckl, Department of Digital Humanities, University of Passau, Hans-Kapfnger-Straße 14 d, 94036 Passau, Germany; E-mail: [markus.eckl@uni-passau.de](mailto:markus.eckl@uni-passau.de)

**Citation:** Eckl M, Ghanem C, Spensberger F (2021) Mapping Discourses in Social Work Research-A Topic Modeling Approach. Int J Global Soc Work 4: 118. doi: <https://doi.org/10.15344/ijgsw/2021/118>

**Copyright:** © 2021 Eckl et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

methods of analyzing large volumes of text, such as word frequency analyses, are increasingly attracting attention [14,15].

**How to apply quantitative text analysis: an exemplary study**

The application of quantitative text analysis using the example of topic modeling is described below. The aim is to show how quantitative text analysis can be applied to examine topics ranging from the formulation of questions to data analysis, based on a concrete research project on the analysis of disciplinary discourses in social work.

**Study objective**

As previously mentioned, an increasing number of digital texts are being produced in social work. Thus, in recent decades, new digital journal and book formats have been created, and progressively more digital publications are appearing [16]. It is no longer possible to take note of all the new publications, even within one’s own field of research. Furthermore, it is hardly possible to determine which topics are discussed regularly or appear in numerous publications. Consequently, the objective of this research project is to investigate and describe the development of the topics in disciplinary social work discourses. At the methodological level, this research provides insights into how algorithms can be used to extract the topics that are discussed within social work and how automated methods used in the discipline of digital humanities, among others, can be useful in the social work context.

**Methods**

**Data basis for the study**

The data basis, which is the bibliographic data set “Social Work Research Database” [16], consists of metadata from 42,001 articles and their abstracts (5,466,792 words). These data originate from social work research published in English-language journals between 1991 and 2015. The abstracts and their corresponding years of publication are used to obtain information about the topics covered by the corpus and how the most central of these have developed over time.

**Preparing the text for analyses**

The text must be prepared before a concrete method can be used to extract topics from the corpus. The programming language Python was used for this purpose [17], as well as for all further analyses<sup>3</sup>. First, using a process called “tokenizing,” the text was segmented into analyzable units. In general, a token can be a sentence, a word, or several words. In this study, a token consisted of one word. Second, punctuation marks, spaces, and commas were deleted, as they have no meaning in computer linguistics. Numbers and special characters (e.g., page numbers and statistical results) were also deleted. In addition, all words consisting of fewer than three letters were deleted, as such words (e.g., “is” and “a”) typically have no relevant semantic meaning. For the same reason, individual words defined as so-called “stopwords” (e.g., “are,” “where,” and “who”) were deleted. Certain “scientific stopwords” (e.g., “journal,” “discussed,” and “focus”) were also deleted, as such words appear frequently in abstracts, but they provided no relevant information in the context of the present study.

<sup>3</sup>The Python package *gensim*, in which the LDA algorithm is already implemented, was used for the analyses and the package *nlTK* (Natural Language Toolkit) for text cleaning and processing.

This process of data preparation had to be adapted to the research question and data basis. For example, some numbers, very short words, or abbreviations may be important for certain research questions. One possible research project that could require the inclusion of very short words would be an analysis of political discourses—for example, election campaigns on online media. In such cases, abbreviations such as “EU” (European Union) may provide important semantic information.

In addition to creating tokens and removing irrelevant content, it is important to ensure that all tokens (words) are either capitalized or lowercase, as a computer recognizes the same word, once capitalized and once lowercase, as two different words.

The next step in the data preparation process is called “lemmatization.” This is a process in linguistics and specifically in computer linguistics, in which different derivatives of a word are combined; for example, “walk,” “walks,” and “walking” become simply “walk.”

**Creating a document-term matrix to describe the frequencies of terms**

Once all of these cleanups have been implemented, the remaining words of each abstract or documents are transferred to a document-term matrix, in which each document has its own row (Table 1). Each column contains one word, and the cells show the absolute frequency of a word in a document. However, there are also so-called “bigrams” - that is, two connected words that often appear together in texts (e.g., “health\_care” and “evidence\_based”). The results of the exploratory analysis of the data set have shown that the most interpretable results are obtained when these connected words occur at least 20 times in the data corpus. The identified “bigrams” were entered into the document-term matrix as independent terms.

Document 1: “I go shopping because I want to . . .”					
Document 2: “I am going out for dinner.”					
Document n: . . .					
	me	go	shopping	eat	because
Document 1	2	1	1	0	1
Document 2	1	1	0	1	0
Document n	-	-	-	-	-
. . .	-	-	-	-	-

Table 1: Document-Term Matrix.

**Identifying topics based on algorithms: topic modeling**

How can topics from texts be determined with the help of algorithms? Topic modeling is a suitable and frequently used method for finding topics in texts, especially in the context of the digital humanities [18]. Latent Dirichlet Allocation (LDA) by Blei et al. [19] is especially suitable for this purpose, as researchers using this statistical procedure do not have to subjectively determine or specify topics; rather, the probability of certain topics is determined based on the words that occur together in the documents. LDA is based on a complex statistical model that allows for the investigation of semantic structures in the text. LDA will now be described with a focus on how it works, as well as its application, rather than on the presentation of its mathematical basics.

A basic assumption of LDA is that a document contains only a certain number of topics. The topics are, in turn, determined by the words that are contained in the document. If words occur together in two different documents, the documents are determined to be similar. With each additional document that is viewed, the algorithm determines which document the newly added one corresponds to in terms of the words. In cases in which there are extensive overlaps, the probability that certain words can be assigned to a topic increase. Hence, there is a certain probability distribution that a document consists of certain topics. A second probability distribution refers to the likelihood that a topic consists of certain words. For example, the topic "systems theory" can be determined by the words "Luhmann," "system," "autopoiesis," "sense," and "function". Only the words with the highest probability of representing a topic are used to determine a topic. If, for example, a word appears alongside many different words in numerous documents, the probability that it can represent a certain topic appropriately is reduced (an example of such a word is "and"), as indicated by the weak co-occurrence (i.e., a low weighting for the respective word). If, however, a word occurs often along with the same words, the probability that it can represent a certain topic increases (an example for such co-occurring words are "Luhmann" and "system"), and this would indicate a strong co-occurrence. Therefore, a word must appear along with the same words in several documents for it to receive a high weighting.

Several important parameters of the LDA model are decisive with regard to how well a topic can be interpreted. One such parameter is the number of topics that must be determined by the researchers. Although other parameters also influence the algorithm, the authors have deliberately chosen to discuss only the number of topics here; otherwise, the mathematical explanations would have exceeded the scope of the current paper. Interested parties are referred to the article by Blei et al. [19], which discusses the basic parameters.

The number of topics can be determined in several ways. One option, which was applied in this study, is to measure topic coherence [20]. The basic idea is to determine the quality of the topics based on the number and weighting of the co-occurrences of words in the documents. Fifteen models, each with a different number of topics, were calculated for this study. The best model yielded 30 topics<sup>4</sup>. Note that the "best" model is not necessarily the easiest to interpret (i.e., a high degree of coherence does not always go hand in hand with an accurate human interpretation of the topics). Therefore, different models must be calculated and the interpretability of the extracted topics assessed.

## Results and Discussion

### Possibilities and limits of topic modeling based on selected results

It is beyond the scope of this paper to present all 30 topics; thus, selected topics are used to illustrate the potentials and limitations of topic modeling. For those who want to examine the code and results more closely, we released an IPython notebook<sup>5</sup>.

Table 2 shows seven identified topics. The second column contains 10 of the highest weighted words for each topic. Hence, the word lists have a high probability of reflecting a particular topic in the

disciplinary discourse of social work. Subsequently, researchers need to find a generic term for the words that belong together (i.e., inductive category formation). Table 2 shows that most words within a topic are semantically close to each other. One can also see which subjects are discussed together within a topic. For example, the terms of the topic "professionalism" show that associated publications discuss subjects such as "professional decision-making" and "evidence-based practice." At the same time, it becomes apparent that these subjects are often discussed from an ethical point of view ("ethics"). Similarly, the terms of the topic "addiction prevention" show that subjects related to drug use ("alcohol" and "drug") in children and adolescents ("adolescent" and "youth") are often discussed in relation to the subjects of prevention ("prevention" and "risk"). The terms "behavior" and "abuse" are also frequently used in these publications, which could indicate an individual perspective on the subjects of addiction and addiction prevention. However, this also reveals a limitation regarding the interpretability of the results, because it is theoretically conceivable that an individualizing perspective on addiction subjects or the increase in preventive approaches could be criticized in corresponding publications. Hence, interpretations must consider that the algorithm is incapable of recognizing negations. In such cases, it would be useful to supplement the results of the LDA topic modeling with a qualitative analysis of a sample of corresponding publications to learn more about the qualitative meaning of the identified terms [21]. A triangulation of the results with a qualitative analysis of the data could also provide clues regarding how to interpret the individual terms of the topics that do not initially have a clear semantic connection to the remaining terms. For example, one could examine whether the term "system" in the topic "child protection" refers to systemic practices within this field of social work. Such connections cannot be recognized by topic modeling.

The topic "racism & gender" is note worthy in that the algorithm linked two subjects. This means that a direct connection is established between these subjects due to their mutual appearance in the publications. It might be hypothesized that critical social work approaches on issues of oppression, exclusion, and social inequality have been identified in one topic. The representative words of this topic ("american" and "african\_american") might refer to a limitation that results from the data basis. As explained by Eckl et al. [22], it is reasonable to assume that publications from the United States are proportionally overrepresented in the Social Work Research Database. Hence, the results have to be interpreted against the background of the relative exclusion of disciplinary discourses from other countries, and the identified topics might not reflect the semantic structure of social work research from other countries in which publishing in English-language journals is not the norm.

In a further analytical step, the identified topics can be presented with regard to their development to show which ones were discussed and how extensively. The generated graphics show time on the x-axis and a standardized topic-document weighting on the y-axis (see Figure 1-3). For this purpose, the topic weights are normalized for each document. This means that all values for each topic in a certain period (e.g., a year) are totaled and then divided by the sum of all weights in the respective periods. In the normalized data, the sum of all the normalized weights for each period is one. This way, the share of each topic in the total weighting of each period can be depicted. If a topic increases in a corpus, it also accounts for a larger share of the total weight [23]. This normalization of the weightings is used primarily when the topics are examined in relation to each other.

<sup>4</sup>For a detailed description, please see the link to the IPython notebook in footnote 5.

<sup>5</sup>[https://github.com/MarkusEckl/Eckl\\_Ghanem\\_Handbuch\\_Digitalisierung\\_und\\_Soziale\\_Arbeit](https://github.com/MarkusEckl/Eckl_Ghanem_Handbuch_Digitalisierung_und_Soziale_Arbeit) (retrieved 15.03.2021)

Topic	Representative terms of the topic
child protection	child, welfare, child welfare, protection, worker, child protection, service, case, system, practice
addiction prevention	youth, risk, adolescent, substance, behavior, drug, abuse, factor, prevention, alcohol
qualitative research	experience, interview, qualitative, participant, theme, life, conducted, depth, using, perspective
quantitative research	scale, assessment, factor, used, sample, instrument, validity, using, item, test
racism & gender	american, african, african_american, black, gender, white, ethnic, racial, race, male
domestic violence	woman, violence, abuse, sexual, domestic, victim, partner, domestic violence, physical, survivor
professionalism	practice, worker, decision, evidence, based, making, practitioner, ethical, professional, knowledge

Table 2: Selected topics and their representative terms.

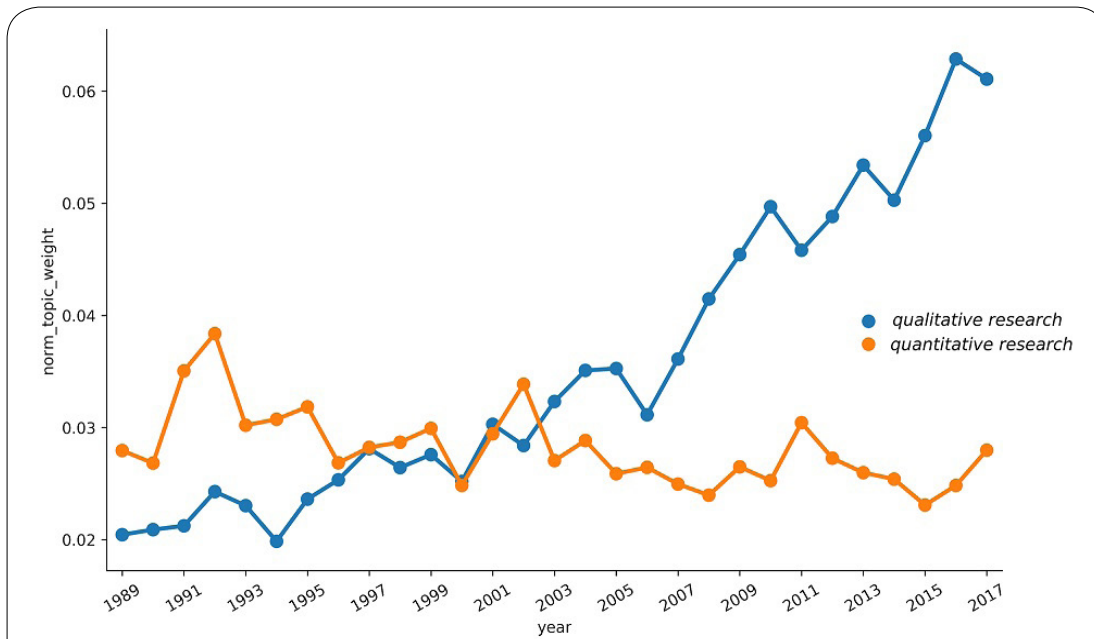


Figure 1: Development of the topics 'qualitative research' and 'quantitative research'.

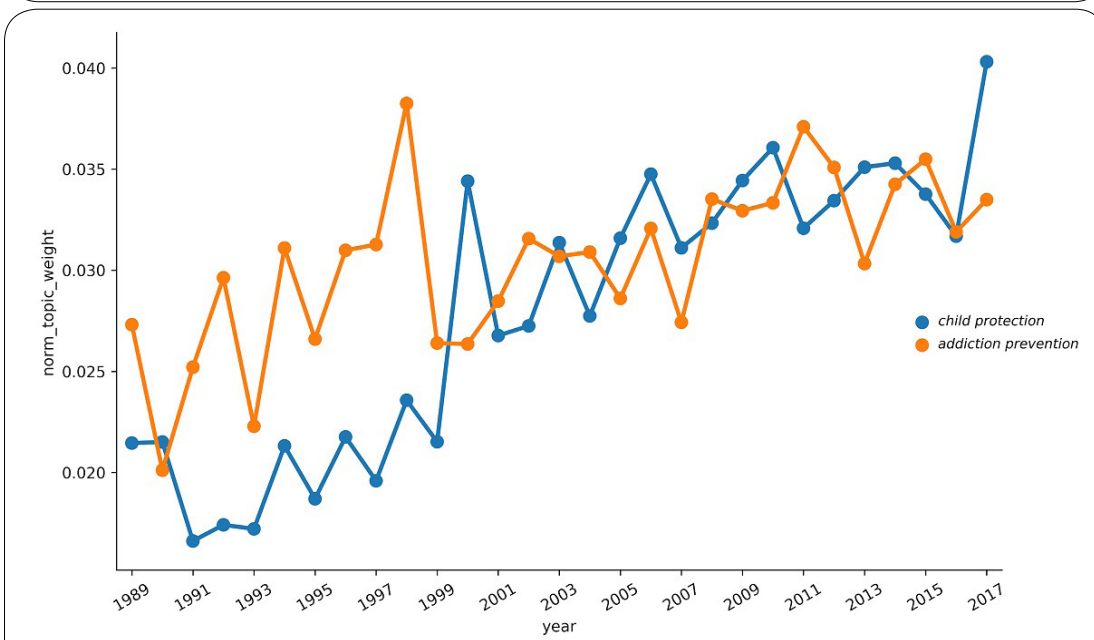


Figure 2: Development of the topics 'child protection' and 'addiction prevention'.



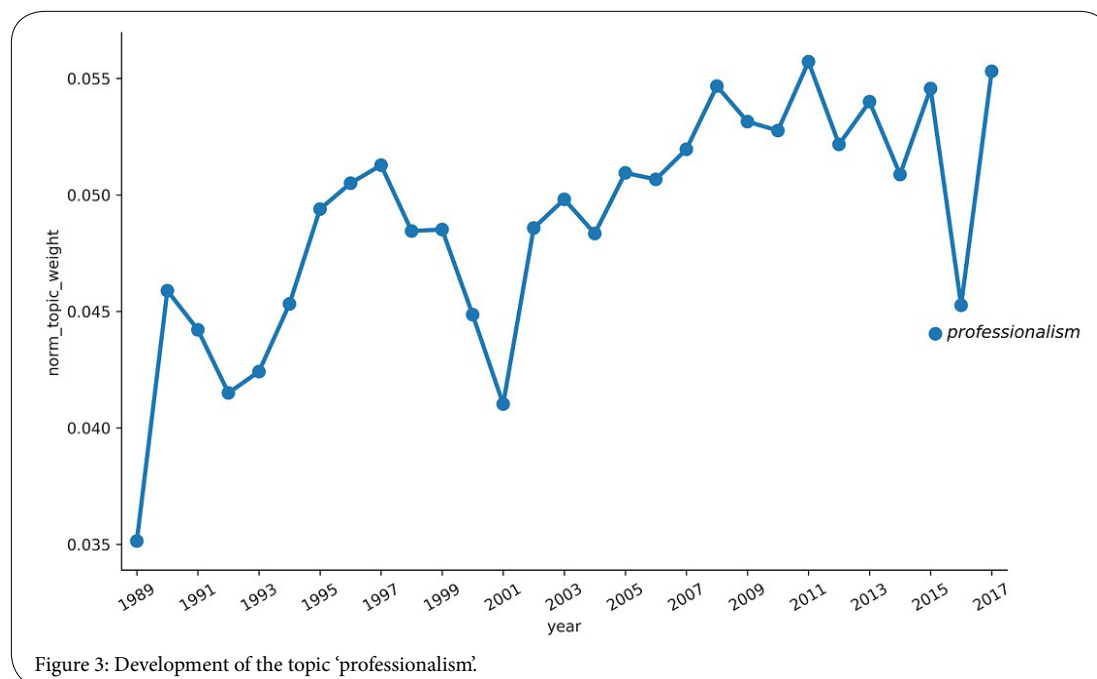


Figure 3: Development of the topic 'professionalism'.

Figures 1 to 3 show the selected topics and their occurrences in publications between 1989 and 2015. Figure 1 compares the occurrences of the topics “qualitative research” and “quantitative research.” The results show that the normalized relative frequency of publications on quantitative research remains relatively constant, whereas the relative frequency of the topic qualitative research increases constantly. In particular, since 2007, a vast increase in qualitative research topics can be observed.

A similar increase can be observed in the topics “child protection” and “addiction prevention” (Figure 2). In particular, child protection topics have been increasingly discussed since the turn of the millennium. However, addiction prevention topics, mostly regarding young people, have steadily gained in importance (Table 2). A possible explanation for this increase could be derived from research that shows that the topics of safety orientation and risk management and consequently the demand for preventive work are increasingly shaping social work orientation [24-26].

Figure 3 shows the appearance of the topic “professionalism.” Accordingly, subjects related to professionalism in social work seem to have gained in importance compared to other topics that have been discussed. Interestingly, such subjects are often associated with evidence-based practice (Table 2). This points to an aforementioned limitation—namely, that publications from some countries, such as the United States, are proportionally overrepresented in the Social Work Research Database [22]. The increase in the number of occurrences of this topic could be partly explained by the extensive discussions regarding evidence-based practice in disciplinary social work, especially in the United States. If this topic represents the discourse on evidence-based practice, this would be in line with the findings that show an increase in this discourse at the turn of the millennium [11].

#### Discussion of topic modeling as a research method in social work

The “grand challenges” mentioned in the introduction [9] not only emphasize the diverse transformation processes in the social sector but also offer concrete strategies for coping with them. Coulton

et al. [9] argue that social work should be considered an active actor in computational social science at the disciplinary level and that respective tasks must be solved in an interdisciplinary way via collaborations between social work and disciplines such as computer science. This way, specific research methods that can capture new phenomena associated with increasing digitization become more accessible for social work scholars. One method of addressing the overwhelming amount of data was presented exemplarily in this article. Topic modeling can be used to identify certain patterns and semantic structures in large amounts of text, thereby assisting to structure these data sets thematically. This method can also be applied to identify terms used by social workers in organizational documentation systems to document cases in social work practice. For example, researchers could ask whether central reference categories (such as “risk” and “security”) change over time in order to interpret child welfare or probation service cases.

#### Limitations of Topic Modeling

Although digitization offers tremendous potential and new possibilities for research, equally significant challenges also exist. For example, vast amounts of data can lead to exaggerated generalization tendencies and claims of truth [8]. However, these are not permissible from a methodological point of view, due to the data basis and the context within which the data were created. The present study can be criticized in this respect because the analysis is based exclusively on the abstracts of journal articles and does not include full texts, print media, and publications in languages other than English. As a result, countries in which journal publications are not widespread or are published primarily in non-English languages are significantly underrepresented in the analysis [22]. Accordingly, the data themselves do not permit valid statements about reality. The meaning and significance of digital data can be understood only if they are embedded in their sociohistorical contexts [27]. The results of automated analysis methods are therefore highly dependent on the researchers’ knowledge of the genesis of the data and the research subject in general. Hence, the final step of context-sensitive interpretation can be understood as a constitutive component of the research process.

This illustrates the importance of the interpretation of data and results. For example, the results of the topic “professionalization” are rather meaningless if not interpreted against the background of other scientific resources, such as knowledge regarding ideas related to the concept of professionalization in social work and to the question of how corresponding discussions have developed over time. Results based on large amounts of quantitatively analyzed data may suggest an objectivity that is further limited by various subjective decisions made during data processing. For example, when using topic modeling, one must decide how many topics the algorithm is supposed to identify. Even if the measure of coherence provides support for decision-making, one must consider the fact that this decision influences the results. In addition, when analyzing vast amounts of data, thorough and necessary data cleansing is possible only to a limited extent. Furthermore, results are influenced by decisions regarding which terms are defined as “scientific stop words” and how often related terms have to appear for them to be considered “bigrams.”

Nevertheless, methods such as topic modeling enable researchers to answer new research questions based on previously unmanageable data sets. Social work researchers must consider how such research approaches can be meaningfully applied to their own research subjects, given the increasing fragmentation and expansion of digital information.

## Conclusion

1. It can be assumed that the handling of big data will increase in importance for social work, and this study might facilitate social work scholars’ application of topic modeling.
2. However, against the background of the outlined problems regarding the interpretation of automatically generated results, analytical methods such as topic modeling should be understood as complementary to the classical research approaches. On the one hand, topic modeling can identify similarities in and generate information about the content of documents.
3. On the other hand, as recently demonstrated by Rodriguez and Storer [20], qualitative analysis can contribute to not only a deeper understanding of the results as compared to quantitative analyses, but also to their validation.
4. Validity is major quality criterion of measuring instruments. Validation means to reason about or test the question, whether a certain measure actually reflects a construct of interest. The use of topic modeling implies processes which may lead to shortcomings in terms of content (e.g., determining the number of topics which a given algorithm is supposed to identify). Qualitative analysis methods can be used to address these content-based shortcomings, thereby contributing to the validity of results generated by topic modeling.
5. Complementary qualitative analysis could also contribute to explain the evolution of certain topics relevant in social work research and get an idea of social or psychological determinants of specific developments.

## Competing Interests

The authors declare that they have no competing interests.

## References

1. Eckl M, Ghanem C (2020) Big data, Textanalyse und Forschung in der Sozialen Arbeit. *Handbuch Soziale Arbeit und Digitalisierung* Beltz Juventa, Germany.

2. Okpych NJ, Yu JLH (2014) A historical analysis of evidence-based practice in social work: The unfinished journey toward an empirically grounded profession. *Social Service Review* 88: 3-58.
3. Rauschenbach T, Züchner I (2012) Theorie der Sozialen Arbeit. In: Thole W (Ed.) *Grundriss Soziale Arbeit. Ein einführendes Handbuch*. VS Verlag für Sozialwissenschaften, Germany.
4. Brynjolfsson E, McAfee A (2016) *The second machine age: Wie die nächste digitale Revolution unser aller Leben verändern wird* (6th edition). Kulmbach: Börsenmedien, Germany.
5. Kutscher N, Ley T, Seelmeyer U (2014) *Mediatisierte Lebens- und Arbeitswelten*. *Blätter Der Wohlfahrtspflege* 161: 87-90.
6. Jannidis F, Kohle H, Rehbein M (2017) *Digital humanities: Eine Einführung*. Stuttgart: J. B. Metzler. Germany.
7. Gillingham P (2019) Developments in electronic information systems in social welfare agencies: From simple to complex. *British Journal of Social Work* 49: 135-146.
8. Schneider D, Seelmeyer U (2018) Der Einfluss der Algorithmen. *Sozial Extra* 42: 21-24.
9. Coulton CJ, Goerge R, Putnam-Hornstein E, Haan B (2015) Harnessing big data for social good: A grand challenge for social work.
10. Graham S, Milligan I, Weingart SB (2016) *Exploring big historical data: The historian’s macroscope*. London: Imperial College Press, England.
11. Ghanem C, Lawson TR, Pankofer S, Maragkos M, Kollar I, et al. (2017) The Diffusion of Evidence-Based Practice: Reviewing the Evidence-Based Practice Networks in the United States and German-Speaking Countries. *J Evid Inf Soc Work* 14: 86-118.
12. Stulpe A, Lemke M (2016) Blended reading. In: Lemke M, Wiedemann G (Eds.), *Text mining in den Sozialwissenschaften: Grundlagen und Anwendungen zwischen qualitativer und quantitativer Diskursanalyse*. Springer VS, Germany.
13. Hockey S (2000) *Electronic texts in the humanities: Principles and practice*. Oxford University Press, England.
14. Coulthard B, Taylor B, Mallett J, Katsikopou K (2018) Natural language processing of social workers’ court evidence statements. Lecture at the European Conference of Social Work Research, Edinburgh, Scotland.
15. Helbig C, Roeske A (2020) Konsequenzen der Digitalisierung für Studium/ Ausbildung, Fort- und Weiterbildung von Fachkräften. *Handbuch Soziale Arbeit und Digitalisierung*, Beltz Juventa, Germany.
16. Perron BE, Victor BG, Hodge DR, Salas-Wright CP, Vaughn MG, et al. (2016) Laying the foundations for scientometric research: A data science approach. *Research on Social Work Practice* 27: 802-812.
17. Van Rossum G, Drake F L (2009) *Python 3 Reference Manual*. Scotts Valley, United States.
18. McFarland DA, Ramage D, Chuang J, Heer J, Manning CD, et al. (2013) Differentiating language usage through topic models. *Poetics* 41: 607-625.
19. Blei DM, Ng AY, Jordan MJ (2003) Latent dirichlet allocation. *The Journal of Machine Learning*.
20. Mimno D, Wallach H, Talley E, Leenders M, McCallum A, et al. (2011) Optimizing semantic coherence in topic models. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.
21. Rodriguez MY, Storer H (2020) A computational social science perspective on qualitative data exploration: Using topic models for the descriptive analysis of social media data. *Journal of Technology in Human Services* 38: 54-86.
22. Eckl M, Ghanem C, Löwenstein H (2019) The evolution of social work from disconnected groups to a scientific community: A social network analysis. *The British Journal of Social Work* 49: 428-447.
23. Wieringa J (2017) Ways to compute topics over time.
24. Dollinger B (2014) Soziale Arbeit als Realisierung protektiver Sicherheitspolitiken. *Zeitschrift für Sozialpädagogik* 12: 296-314.
25. Harrikari T (2013) Securitized childhood: Childhood and youth in Finnish crime prevention programmes. *Youth Justice* 13: 57-72.
26. Peeters R (2015) The price of prevention: The preventative turn in crim epology and its consequences for the role of the state. *Punishment & Society* 17: 163-183.
27. Weyer J, Delisle M, Kappler K, Kiehl M, Merz C, et al. (2018) Big data in soziologischer Perspektive. *Big Data und Gesellschaft: Eine multidisziplinäre Annäherung*, Springer VS, Germany 69-149.