IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 9, 2019, accepted May 1, 2019, date of publication May 20, 2019, date of current version June 3, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2917719

The Application of Artificial Intelligence Technologies as a Substitute for Reading and to Support and Enhance the Authoring of Scientific Review Articles

RÜDIGER BUCHKREMER, ALEXANDER DEMUND, STEFAN EBENER, FABIAN GAMPFER, DAVID JÄGERING, ANDREAS JÜRGENS, SEBASTIAN KLENKE, DOMINIK KRIMPMANN, JASMIN SCHMANK, MARKUS SPIEKERMANN, MICHAEL WAHLERS, AND MARKUS WIEPKE

Institute for IT Management and Digitization, FOM University of Applied Sciences, 40211 Düsseldorf, Germany

Corresponding author: Rüdiger Buchkremer (ruediger.buchkremer@fom.de)

This work was supported in part by the SAS Institute, Heidelberg, Germany, and in part by the RapidMiner GmbH, Dortmund, Germany.

ABSTRACT To gain a comprehensive overview of new scientific findings with the enormous, ever-increasing amount of published information, we apply a new combinatorial approach that complements the process of reading scientific articles by supplementing artificial intelligence technologies. We present a combinatorial approach, which we illustrate in the form of a ''double funnel of artificial intelligence.'' Our approach suggests to largely increase the amount of data at the beginning of the data collection process and to subsequently clean and enrich the data set in order to gain much more knowledge at the end of the procedure compared to a ''classical'' literature review. We use natural language processing and text visualization techniques to uncover findings that are generally unbeknown to the human reader due to the inability to process very large amounts of text. By illustrating the individual steps using practical examples taken from use cases, we demonstrate the merits of our approach. With our methodology, we are able to reproduce findings from ''regular'' review papers; however, we discover additional and new findings in different fields, such as data science or medicine. We also point out the limitations of our approach. Finally, we make suggestions as to how the methodology could be further developed.

INDEX TERMS Computational and artificial intelligence, document handling, fuzzy control, knowledge acquisition, pattern analysis, scientific publishing, text mining, text processing.

I. INTRODUCTION

The number of scholarly peer-reviewed English-language journals increases 5% to 6% every year and has reached approximately 33,100 active journal titles as of August 2018. If we take only these scientific journals alone, we need to consider a number of articles on the order of 3 million – per year [1]. This has previously been reported by Gu and Blackmore in 2016. They observed that the number of newly created journal titles lies in the range of 1.500 per year – with a rising tendency [2]. These figures demonstrate that it seems almost impossible to comprehend this enormous amount of information with classical human reading and learning processes. New methods for the evaluation of publications are essential so that researchers can identify relevant knowledge in their respective fields. In this context, natural language processing (NLP), one of the founding components of artificial intelligence, in combination with machine learning promises a bright future for computer-aided text analysis [3]. However, before we can apply these technologies to extract knowledge from a large body of literature, we must first examine how current review articles are produced. It is not uncommon for researchers to start the investigative process by simply stringing search terms together and inserting them into input boxes of various literature databases, rarely applying any operators or keyword techniques; the outcome

The associate editor coordinating the review of this manuscript and approving it for publication was Huazhu Fu.

of the simple search process, the body of literature, however, is often dissected using high-quality procedures, mostly applying manual and human cognitive methodologies. Vom Brocke *et al.* recently stated that many authors completely refrain from explaining how the articles on which the review article is based were put together [4]. For many readers and journal reviewers, it appears to be satisfactory that the authors possess expertise and, thus, know how many and which input articles to use. We do not want to question the esteem in which articles are valued, but we would like to suggest relatively simple means of improving the state-of-the-art manual search procedure. This paper focuses on this problem and describes a methodology that initially represents a bootstrapping [5] concept, a variance of reinforcement learning, which is being applied to texts. The proposed approach may generate an even larger set of input query terms, as well as a much larger literature corpus, compared to regular literature reviews. However, by applying artificial intelligence methodologies, it is possible to "read" the large sets of articles. As indicated above, we not only present a procedure to improve the compilation of relevant articles, we also apply text analytics and visualization techniques to improve the output, the extraction of knowledge from the corpus. The methodology described supersedes, to a certain degree, the classical generation of human literature reviews, especially for cases where the review focuses on "central topics" according to the review paper classification scheme of Fettke [6].

In this work, we first present the scientific background of the technologies applied here and present similar approaches that have already been published.

The phases of our STIRL (systemic taxonomy for information retrieval from literature) methodology are briefly discussed in section three of this paper and applied on the basis of individual case studies in section four where various optimization proposals are presented as well. The case studies demonstrate the individual advantage of different phases of the methodology. The concluding section five will point out opportunities, limitations and possibilities to further extend our approach in the future.

II. BACKGROUND AND RELATED WORK

By 2002, Webster and Watson had already stated that an "effective review creates a firm foundation for advancing knowledge." It facilitates a "theory development and closes areas where a plethora of research findings exists, and uncovers areas where research is needed" [7] In recent decades, researchers have established various methods to enhance a scientific literature analysis, and these methods can be classified into different groups. The traditional approach consists of manual reading, the extraction of important key findings and statements, their comparison and evaluation, and a summarizing of knowledge. These methods have been named "traditional literature reviews" [8], [9]. Due to the growing number of publications, however, the risk of ending up with an information bias has increased because the finding and viewing of the relevant amount of literature can no longer be guaranteed manually due to the quantity and the required time needed [10]–[12]. Driven by this problem, further technical approaches have been developed, especially in the area of medicine. The systematic analysis is often biased by the development of the individual author writing strategy. At the heart of this strategy lies the first definition of specific questions for which corresponding publications are already known to the authors. Because the search process is very demanding without modern tools, review teams have been formed to work through this process together [10], [13], [14]. Another approach based on both traditional and systematic methods is meta-analysis. In this case, it is not only essential to find and evaluate the literature, the setting of ranking criteria is often needed. This takes external (meta-) factors into consideration. For example, factors such as the number of reviews, the reputation of the journal and authors, and the references are being used. Furthermore, contents are also evaluated to check whether research gaps can be identified [15]–[17].

The support infrastructure to conduct literature research and evaluate results has developed further in recent years through the Internet in general, automated search algorithms and extensive literature databases, which, as already briefly mentioned, requires adjustments to classical retrieval methods. In the field of systematic analysis, a growing interest in advanced methods appears to exist because the challenge of forming teams to handle the complex search and evaluation is often not performed by one person alone [6], [13].

In the context of this paper, the systematic reviews are generated by a combination of tools and processes that largely belong to the field of natural language processing (NLP). It can help to generate an effective and successful literature review with manageable effort despite a very large number of publications. To achieve this goal, text mining subroutines such as latent Dirichlet allocation (LDA), text clustering or text bootstrapping are applied in combination with machine learning algorithms. Furthermore, the improved analysis is now made possible with the increased computing capacity required for automated text analysis [3], [18]–[22]. Fayyad *et al.* described the methodical approach Knowledge Discovery in Databases (KDD) in 1996 and stated that "it is only natural to turn to computational techniques to help us unearth meaningful patterns and structures from the massive volumes of data," as well as that "practical computational constraints place severe limits on the subspace that can be explored by a data-mining algorithm" [23]. The applied techniques from KDD, such as "data mining," "artificial neural networks" or "support vector machines," as well as the process steps "preparation," "modeling" and "postprocessing," can be found in the technical implementations of a systematic literature review (SLR), with the result that the approaches described by Arji *et al.* can be applied [24]. At first, it may appear that limitations from the past can be neglected due to technological advancements and cloud computing, but in comparison to the past, the amount of data to be analyzed has become more extensive [1], [2].

Current scientific articles and their metadata are listed in online databases of publishers and other search engines as part of the publishing process. Older scientific articles can be found in databases so that simple keyword search queries can lead to very large amounts of search results. To get mostly relevant results, users must optimize search parameters [25]. This optimization is mostly sophisticated and unintuitive such that suitable experience in the operation of search engines is required [26]. Moreover, there is a lack of structured approaches on how big data or artificial intelligence concepts can be leveraged in this space. Based on the work of vom Brocke *et al.* and their SLR process, the fuzzy-based search query generation of MacFarlane et al., as well as the text mining methods of Liu *et al.*, an approach will be presented in this paper wherein different online databases can be partially or fully searched via automated procedures and evaluated with respect to specific scientific topics [10], [18], [19].

Fayyad *et al.* used data and text mining as early as in the late 1990s to analyze extensive information packages and texts [23]. Today, text mining technologies are used to support systematic reviews across various research areas [14], [27], [28]. According to Yu and Menzies, however, it is a ''relatively simple task to find a few relevant papers for any particular research query.'' However, the problem is ''not to find a few papers, but instead to find the most relevant papers.'' They consider different methods, such as search-query optimization methods, reference-based procedures, supervised learning, semisupervised learning, unsupervised learning, and active learning, and they name it $FAST^2$ [29]. A fundamental problem of earlier models is the lack of recognition of the application in practice. With the model $FAST^2$, Yu and Menzies have stated themselves that the model has been tested only in the field of SLR of ''software development'' and that, in other fields, investigations are missing [29]. Although different models mentioned so far demonstrate how researchers could include the model in their investigations, concrete practical case studies or evidence of transferability to other scientific areas are rarely available [27]. Moreover, there is a lack of documented structured approaches that support researchers in applying text analytics for their reviews [30]. In this contribution, various phases are therefore discussed using practical case studies to facilitate their application in a variety of research areas.

The first step of a literature search is supported by information retrieval methods, and it is important to define the range and purpose by setting a search query range in advance [31]–[34]. Various truncation-, proximity- and Boolean operators can be used to define the relationship between semantic terms that have been enriched by various thesauri [28], [32]. Ontologies, as well as their enrichment (for example, via bootstrapping), have been extensively used for mapping relationships in various domains [35]–[38]; however, a combined approach by utilizing mining and learning procedures, as well as advanced information retrieval, is still lacking. We also would like to address an important intermediate step, namely, the generation of the literature corpus. Because various steps from a large set of possibilities lead to the corpus and many possibilities result from the corpus, we selected a model illustrated as an artificial intelligence double funnel (see Figure 5).

The collection of semantic terms to build a taxonomy or an ontology may generally be pursued via supervised (requires manual labeling and training), reinforcement learning, such as bootstrapping or unsupervised learning (using automated classification procedures) [3], [39], [40]. Relevant text clusters for semantic enrichment may be identified by using mathematical algorithms, such as k-means [41] or maximum entropy models [42], [43]. These methods have also been used to identify thematic areas within a text [44]. Within this paper, only abstracts of scientific papers are used to create a large body of literature. On the one hand, abstracts and keywords are, most of the time, available in many online databases without access restrictions; however, on the other hand, the abstracts provide the highest information density within a corresponding scientific publication [45], [46].

III. STIRL – GENERATION AND APPLICATION OF SYSTEMIC TAXONOMIES VIA INFORMATION RETRIEVAL AND SEMANTIC LEARNING

The STIRL methodology presented in this work (see Figure 1) leverages learnings from previous studies that apply the use of ''text mining'' and ''information retrieval'' to support systematic reviews: Felizardo *et al.* describe an approach that combines classical methods for the initial phases of a review with text mining support for later stages [47]. Xiao and Watson provide a current literature review typology framework for stand-alone (describe, test, extend, critique) as well as background reviews [17]. STIRL supports descriptive as well as critical reviews since all of them are denominated as ''include all.''

Thomas *et al.* suggest that the search and analysis phase of a literature review can be supported by text mining [27]. Therefore, we combine current literature review methodologies [8] with state-of-the-art text mining approaches [48], [49] into a process model, which we used for the analysis described in this paper - see Figure 1. The phases of STIRL are presented in detail within the following subsections.

A. INFORMATION RETRIEVAL

The collection of scientific articles represents a major ''side effect'' of any scientific research process, and it is often supported by classical information retrieval [34]. An aggregated literature set often comprises a vast amount of papers where the subject broadness is often defined by the research scope. For ''extend'' and ''test'' literature reviews [31], the authors thoroughly select high-quality articles in a magnitude of approximately two hundred articles by applying expert knowledge. The articles generally provide high overall information density with respect to the topic in question.

To mitigate the ''garbage-in-garbage-out'' effect, it is important to start the literature search with the most suitable

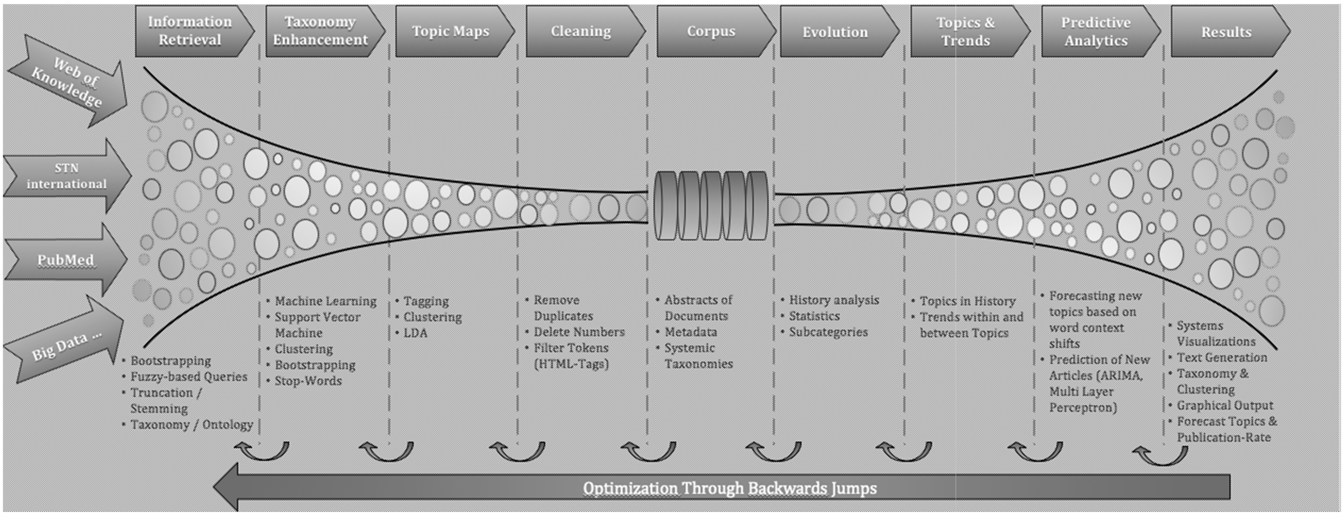


FIGURE 1. STIRL methodology – the artificial intelligence double funnel.

search algorithms. We would like to point out that "information retrieval" is not a new technology. It has been in use for decades but has mostly been used by trained information scientists or librarians. However, we extend the use of classical information retrieval by reinforcement learning, which produces significantly more hits than conventional online searching. In addition, we recommend the generation of text clouds, which allows for visualization of terms from the semantic proximity of known search terms. We suggest that the application of basic information retrieval algorithms should be used by every scientist because it is quite easy to understand and because it is becoming more and more indispensable due to the greatly increased number of scientific articles. We suggest focusing on Boolean, proximity and truncation operators since almost all literature portals offer their use.

Our approach applies a fuzzy-based query and aggregation procedure where the input vocabulary gets enhanced by a semisupervised learning procedure, which can also be regarded as bootstrapping since a small set of terms gets accumulated step by step using "precision" and "recall" as evaluation criteria [3].

Sometimes it might be essential to get a broader overview that covers more distant areas of the research subject – with our new approach, it is possible to extend the review process to get a "view beyond the horizon" [50]. Therefore, it is essential to formulate a comprehensive search query where several thousand hits may be obtained. A search query for "data science," for example, should also include "big data," and it may also include "artificial intelligence" since AI technologies are often related to big data procedures [35]. A search for "customer relationship management," for example, should also include the abbreviation "crm" as well as the search term "customer" in proximity to the abbreviation since "CRM" is often used for "certified reference material" [36]. A search query should therefore be enhanced by using proximity and truncation operators [34], as well as taxonomies such as MeSH or WordNet ([37], [38]. To search for medical diseases, we thus could include articles where the medical term has been intercepted by other terms ("bacterial infections" and "bacterial skin infections"). A search query to find articles about organization design applied to information technology could look as follows:

((Organi?ation* NEAR/3
(design OR structure) AND
(information NEAR/3 technology))

The methodologies mentioned in chapter II are generally selected to improve the precision of a search. At this point, however, high precision [39] is not yet essential if the primary objective is to obtain a view beyond the horizon related to a given subject. To break down the surplus of articles due to query extensions, subsequent analytical methods may be required. However, the magnitude of the surplus needs to be optimized via search query optimization steps [39]. Thus, by applying our STIRL methodology, we switch between recall and precision enhancing methods until we end up with the first comprehensive literature corpus.

B. TAXONOMY ENHANCEMENT

Moed and Halevi illustrate that the level of downloads within a given research subject is highest for review articles [51]. In accordance with our STIRL methodology, we suggest separately identifying and classifying a set of review papers because they may play an important role in generating systemic taxonomies or ontologies [17], [52]. Review articles help by defining borders and specifics of a given research field. Systematic reading and scanning of review articles supported by classification tools and learning procedures, as well as articles that are dedicated to generating taxonomies or ontologies, have played an important role in enhancing our given taxonomy, which will be demonstrated later in this article [53]. Finally, we expand our search queries through automatic classification of the literature corpus (for an overview of text extraction methodologies, see [18]). Thus, taxonomies taken from review articles and through automatic classification of a larger research framework will be combined. Throughout the full STIRL methodology, it is recommended to constantly monitor the precision / recall ratio that may require manual reading of abstracts or full articles.

In conclusion, it is important to dissect the corpus of documents into different topics. According to the subject area to be examined, there are different approaches in the literature that discuss an optimal number of topics. Through a normalized mutual information (nMI) process, it can be quantitatively compared to the output of clusters and used as an external validation source. This results in how close the individual clusters are to the underlying taxonomy [54].

In addition, pointwise mutual information (PMI) can be used to determine the correlation of topics. The coherence is defined as:

$$PMI\left(k\right) = \sum_{j=2}^{N} \sum_{i=1}^{j-1} log \frac{p\left(w_i, w_{ij}\right)}{p\left(w_i\right) p\left(w_j\right)}$$

The larger the PMI is, the more likely it is to imply that the topics k and p are coherent. As a rule, topics are formed with a number of 10 top words [55], [56].

Further methods for determining the best number of topics, which are partly based on the PMI, are known as the "Elbow Method" [57] or the "Davies Bouldin Index" [44], [58]. Due to numerous variations of those methods with their own advantages and disadvantages, we reference existing papers related to this topic. For example, Liu *et al.* compared 11 methods in 2010 [59].

The generated taxonomy is then applied to the full literature corpus. Precision / recall ratios [52] can provide an indication for evaluation of a given taxonomy. In case of a nonsatisfying evaluation, the taxonomy should be further improved. In accordance with state-of-the-art text mining methods, we suggest an iterative approach between various phases of STIRL. Both supervised and unsupervised learning techniques can be applied [60], [61]. Once the evaluation results are satisfactory with respect to the precision-to-recall ratio, the systemic taxonomy can be applied.

C. BOOTSTRAPPING

Designations and abbreviations disambiguate over time and vary according to the context of the topic. Just as CRM can stand for "customer relationship management," it can also be synonymous with "certified reference material," as we explained earlier. For diseases directly related to the term "cancer," we currently find approximately 5,000 terms in the catalog of the National Library of Medicine (MeSH catalog), and if different spellings are taken into account, there are even more. Even among experts, the scope of terminology or the degree of disambiguation is not always fully clear. Especially for amateurs in a field, it is a challenge to find the right search terms for diseases. However, to simplify this process, we use the know-how of experts that have published articles about it to define the search term scope. An initial set of articles found that way may provide further terms that are often located in close proximity to the terms already used. Together with the new terms, a subsequent search is performed until no new search terms can be found with reasonable effort. Similar to the small strap, which helps to pull up the entire boot, an initial search helps. The verification of the bootstrapping process should always be checked against the precision-to-recall ratio.

D. CORPUS

An important intermediate result of the STIRL model is the creation and quality assurance of the text corpus generated with the methods described above. It contains scientific articles that are to be used for further steps. The corpus may also contain meta-information, such as keywords or tagging information, provided by the authors or created by a comparison to known taxonomies, catalogs or ontologies. Time stamps should also be taken into account, for example, whether they contain the publication date or the data entry date. In addition, duplicates and formatting conflicts must be removed or harmonized. For further processing, numbers and special characters must be removed. This includes approaches for reducing a number of words to root forms, removing program code, such as HTML syntax, or applying stop word lists. The corpus should also be examined randomly for meaningfulness and gross errors. For this purpose, an expert should be consulted if the reader is not an expert themselves. The body of literature can also be dynamically maintained by employing alerts or monitoring. If the data sets vary dynamically, an audit trail to monitor changes is mandatory. Such a text corpus can also be considered as a collective reference containing the aggregated information on a subject. In addition, the corpus provides an easy way for researchers to exchange scientific research results to define a common starting point.

E. TOPIC MAPS, LATENT DIRICHLET ALLOCATION AND OTHER TOPIC-FINDING METHODOLOGIES

After the corpus has been created, contents and data quality need to be reviewed. This includes a rough sorting or clustering of the contents and can be performed using keywords or time intervals. Unsupervised clustering may be useful in many cases. For this purpose, we also use methodical approaches, which can be assigned to artificial intelligence. The ISO standard methodology "Topic Maps" may help with navigating via abstracted levels and through simplifying semantic structures - a document corpus can hereby easily be mapped and searched for patterns. The different topic maps vary depending on the number of articles they comprise. The semantic boundaries and distances between different topics can also be determined and visualized, as will be depicted at a later point within this article. The topic maps may be described by corresponding keywords and semantic relationships – a knowledge map could be the result of a topic map analysis. A well-known generic syntax for visualizing corresponding topics is OWL (web ontology language) [62]. Another important tool for identifying subtopics with smaller text passages is represented by a methodology called ''latent dirichlet allocation'' [63], [64].

F. EVOLUITON

Because most scientific topics develop over the course of several decades, a historical analysis is indispensable. This makes it relatively easy to determine when and where the topic originated. Because it is also natural for central topics to evolve semantically, system visualization can be helpful, as we will demonstrate later in this article with the topic ''Data Science.'' Some scientific statements are quickly forgotten or are only popular for a certain period of time. Through a historical analysis, the influence of classical media may also be noticed, and sometimes the thematic assessment of the nonscientific world is quite distinct from the scientific one. In that way, we were able to show that the level of managerial knowledge in the environment of enterprise architectures is quite different from that in science. Via time series analysis throughout the complete article set, historic events can be identified, new subcategories may be determined, and future predictions can be made [65].

G. TOPICS & TRENDS, PREDICTIVE ANALYTICS

The generated systemic taxonomy, as well as the historic analysis, can now be utilized in more detail for trend detection [66] or predictive analytics [67]. Because of increased attention, more articles will be published within a given time frame and may be defined as a trend. Various automatic and semiautomatic text mining procedures may be applied to detect trends [66]. Another aspect in the prediction can be the derivation of future interests. Different approaches and algorithms, such as ARIMA or ''Multi-Layer Perceptron,'' are available [68].

H. RANKING AND VISUALIZATION OF RESULTS

To investigate a field of research, it is not the thematic analysis alone that is important - very often, the overall ranking of topics, journals and authors within a corpus is also a matter of weighting. Which topics are published in journals with the highest impact factor? Which are the most important and most frequently cited authors? These are elements that are not apparent through simple reading but that can only be made visible through additional analyses or through experience and expertise. However, the use of the ''Journal Impact Factor'' alone is not sufficient, as Seglen has explained, since articles in lesser known journals can also have a high scientific value [69]. Furthermore, the recitation rate is also not always meaningful on its own, since recently published articles, for example, cannot yet have high citation scores. A large number of scientific publications receive no attention, as MacRoberts and MacRoberts explain [70]. Through the analysis of an extensive, systemic body of literature, neglected topics can be discovered. Here, for example, the k-means algorithm can be used to discover rare but potentially important topics [71]. Various types of metadata might be available depending on the source of contents within the corpus that would be suitable for weighting [72]. Sometimes, metadata are available only to a portion of the corpus – a separate analysis might be suitable, for example, the PubMed database is enriched by MeSH medical subject headings, which represents a taxonomy by itself. To compare distinctive metadata findings with other results, the mapping of metadata might be suitable, as we demonstrate later within this article where we map MeSH with the nomenclature of the World Health Organization (WHO). In addition to quantitative characteristics, such as the number of citations, trends should be checked critically. Other qualitative characteristics can be checked with, e.g., the ''C.R.A.P. Test'' [73]. The background of the journal in which the publication appeared, for example, helps to classify the exact context, just as the information regarding the peer reviewers of the respective issue gives an indication of the possible focus. The journal's ranking provides further insights. There are various scales, such as http://eigenfactor.org [74] or the h-index. Forums such as https://pubpeer.com/ help for checking whether there are already ongoing critical discussions about the article in the scientific community [1].

The more analysis criteria are used, the more confusing it becomes to derive the findings. For this reason, it is recommended to use a visualization technique. When selecting the form of presentation, it is recommended to examine advanced alternatives in addition to classical visualization forms (e.g., tables, text-clusters, line / bar / pie diagrams). Rekik *et al.*, for example, used wheel graphs in their work [75] to present extensive and complex meshed results. A comprehensive overview of different visualization types has been elaborated by Kucher *et al.* When they published their work in 2015, a total of 141 different types had already been identified and published on a web page with filter options to find the suitable methodology (http://textvis.lnu.se/). In December 2018, the web page contained 430 different types, each of which has its own advantages and disadvantages and must therefore be selected on the basis of the available data [76]. If complexity increases greatly, we recommend a system visualization as it is used, for example, in systems biology or medicine. In the following case studies, we will present an illustrative example [77].

IV. APPLICATION AND EVALUATION OF THE STIRL METHODOLOGY

A. QUALITATIVE CONTENT ANALYSIS TO GENERATE A TAXONOMY FOR AN INFORMATION TECHNOLOGY ORGANIZATION DESIGN FRAMEWORK

The first case study deals with the generation of a search query algorithm (information retrieval, bootstrapping). One of the major research goals of this project has been to provide generic guidance on how to design an information technology organization framework in an increasingly

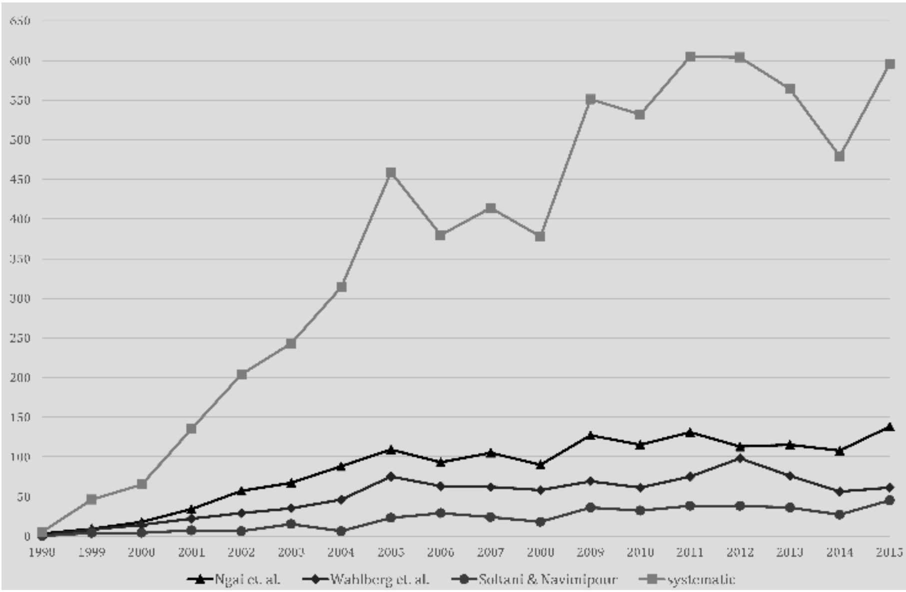


FIGURE 2. Number of articles per year applying the systemic taxonomy generated with STIRL compared to three published taxonomies.

digitized world. It is important to set up a search query that generates high precision at first to generate an initial document pool and taxonomy for bootstrapping. It is recommended to test various search algorithms using syntactic operators [32] by using, for example, Boolean [78], proximity and/or truncation operators (please note that retrieval algorithm syntax needs to be modified if you switch from one portal to another; NEAR/3 at "Web of Science" is equivalent to W/3 at "ScienceDirect"). The query string from above, when used in various popular literature database portals (e.g., AIS eLibrary, IEEE Xplore, ScienceDirect, Web of Science, Wiley Online Library and SpringerLink), revealed 354 unique articles with high precision. Taxonomy building (bootstrapping) is being performed via qualitative content analysis (QCA). QCA has been described as a flexible methodology where impressionistic, intuitive, interpretive analyses may be used – it is highly useful when analyzing full texts for bootstrapping, especially, if only a small number of initial papers are available [53]. 820 different text snippets have been assigned in this project with QCA from 51 selected high-quality articles to form 23 distinct and seven broad categories: Strategy, Structure, Sourcing, Processes, People, Governance and Information. [79] The initial query needs to be extended with the new taxonomy, and it leads to a set of retrieval queries to build the literature corpus.

B. TAXONOMY ENRICHMENT AND EVALUATION FOR CUSTOMER RELATIONSHIP MANAGEMENT (CRM)

Our second example demonstrates, in an impressive manner, how bootstrapping can extend the search index. Customer relationship management (CRM) has been discussed in the literature since the 1990s and still represents a broad topic that is highly popular in science [80]. Within four decades, a broad and dynamic field has emerged.

Due to the long time frame and the availability of various review articles [81]–[83], a small study investigated how a holistic taxonomy can be created by using cluster analysis and the subsequent merging of existing taxonomies with precision/recall quality assurance. For verification of the validity of the initial systematic taxonomy taken from review articles, a comparison of existing taxonomies has been carried out using a test corpus. The literature corpus is created and administered using the "SAS Content Categorization Studio" and comprised 1,206 articles using the database "Web of Science" only within the research area "Business Economics." To avoid falsifications by linguistic changes (e.g., American / British) or generalized terms (e.g., management), the taxonomies were cleaned up first. The combined systemic taxonomy taken from three review articles plus terms collected via bootstrapping achieved significantly higher hit rates on the test corpus than the existing comparative taxonomies taken from the individual review articles. A closer look reveals that the existing taxonomies represent, to a certain extent, the research focus of the respective authors. The systemic taxonomy, however, creates a generalized form that represents a higher degree of objectivity (see Figure 2).

As seen from the graph, the time-related analysis reveals that the existing and systemic taxonomies produce comparable amounts of papers at first (see Figure 2). However, starting from the year 2000, it becomes evident that the field has broadened and that the systemic taxonomy reveals a lot more hits. From 2009 onward, the stability in the number of hits is perceptible for all review article taxonomies, whereby

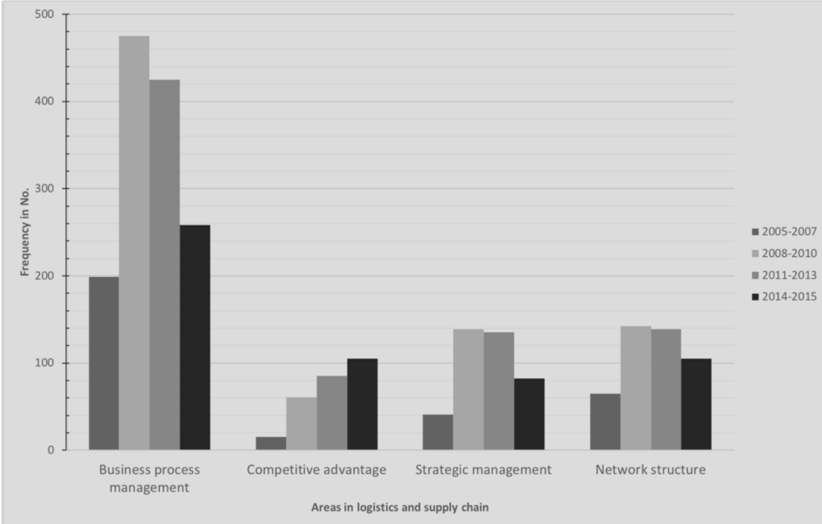


FIGURE 3. Development of four major topic areas in logistics and supply chain management.

the systemic taxonomy yields between 500-600 hits and the reference taxonomies between 40-200 hits.

In addition to our systemic taxonomy, other analyses have shown that the dynamics in the field of CRM have increased [84], [85]. For the year 2001, 20 relevant CRM subtopics have been identified, whereas in the year 2015, approximately 350 became recognizable. An article can be assigned to several subtopics due to different focal points, as mentioned earlier.

C. DETECTION OF MAJOR TRENDS IN LOGISTICS AND TREND INTERACTION ANALYSIS

Our methodology has been used to get an overview of recognizable external changes in the view of one major trend – demographic change. This may allow logistics dynamics concepts to be more sustainable and less sensitive towards future trends [83].

Because supply chain management (SCM) and logistics have often been addressed alongside each other in a broad array of journals [86], we conducted a parallel in-depth analysis [87]. Again, we utilized the data platform "Web of Science" because the contents of that database are often characterized by high-quality measures and the retrieval algorithms that can be formulated offer a relatively high degree of flexibility. A more sophisticated search is possible using database portals such as "STN International." We selected articles published between 2005 and 2015. Although logistics and SCM experts were present in this study, we thoroughly compared our results to already existing reviews; therefore, we focused on the list of relevant journals provided by McKinnon (McKinnon, 2013). To identify major trends, we conducted automatic classification procedures using a mining software package (SAS Enterprise Miner) and approved the validity of four identified main trend areas (business process management, competitive advantage, strategic management, network structure) - see Figure 3.

By connecting these selected trends to others from our taxonomy via semantic proximity classification analysis (for a methodological overview, see [88]), we arrived at the following results: sustainability, information integration, Internet of things, supply chain collaboration and coordination, and integration provide the highest combined visibility. Operational and performance topics, such as vehicle routing, reverse logistics, purchasing and distribution, skills/competences, and organizational learning, could be identified as well. A trend interaction analysis [89] reveals that the interdependency of logistics with collaboration and information integration plays an important role that may shortly become essential. This may influence strategic perspectives on human resources and knowledge management within logistics (see, for example, [90]).

D. TRENDS IN ENTERPRISE ARCHITECTURE MANAGEMENT

In the following analysis, we originally intended to focus on enterprise architecture because it differs from the pure information technology architecture. However, we noticed that many authors do not seem to be aware of these differences in nomenclature, especially in healthcare. Thus, we extended our taxonomy and chose "a view beyond the horizon" in the title of the article [50]. The discipline Enterprise Architecture (EA) was introduced in the late nineteen eighties and has since evolved into a well-known practice for managing information systems in alignment with business capabilities [91]. The rate of articles published in this field is still significantly above average compared to other science publications [2], and we expected to find interesting new dynamics. We initially searched for scientific publications in the following databases: IEEE Xplore, ScienceDirect, SpringerLink, Web of Science, and ACM Digital Library. The selection of portals is based on initial precision / recall studies that we perform with simple search algorithms. We did

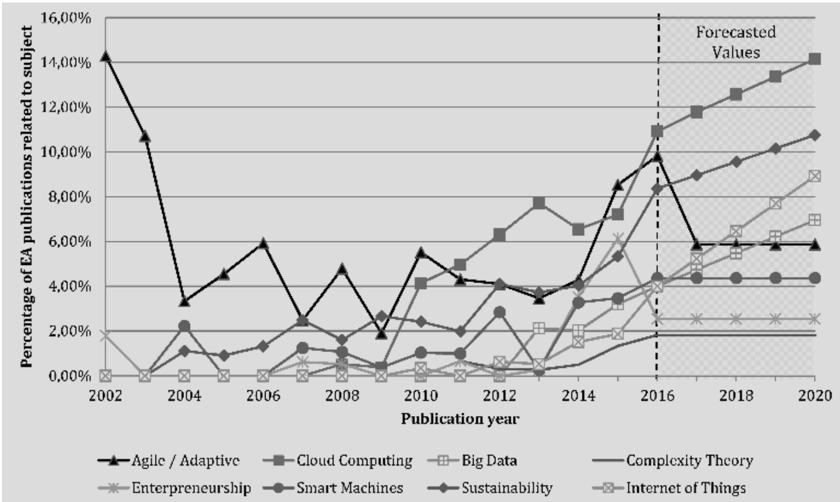


FIGURE 4. History & forecast of current research trends in scientific enterprise architecture literature.

not restrict our search to the term ''enterprise architecture'' since the naming of the discipline evolved with time. After approval of our search strategy and quality assurance of the results, we ended up with 3807 relevant publications since the beginning of this research area. We applied a semisupervised topic identification method [92] to identify past trends. In addition, we applied a fully unsupervised analysis with maximum entropy classifiers [93] to better learn and understand the dynamics of trends. A combination of the software products ''SAS Content Categorization Studio'' and ''RapidMiner'' was used for the text analysis. RapidMiner is available as ''open source'' and as a commercial application. We applied the educational license for this study. An educational license can be obtained from the RapidMiner website at https://rapidminer.com/.

We propose trends that may evolve or diminish in the future by applying the ARIMA models within the software product ''R''; however, from early-stage studies, we anticipate that it is possible to improve the models for prediction. We are currently elaborating genetic algorithms to improve our predictions [94], [95].

Our examination reveals that, for almost a decade after introduction of the new research field, publications focused mainly on subject understanding. Later, the focus shifted towards application and management of enterprise architectures [96]. We identified ''modeling'' as the third significant subcategory, and again, we observed increasing dynamics within this research field. For almost two decades, enterprise architecture management remained relatively silent with two semantic subcategories only, agility and sustainability. As of today, the field is quite broad and scientific importance is still given, and we propose that it will stay that way for several years [97]. The results of our analysis for current trends are depicted in Figure 4.

Agile and adaptive concepts were a major subject in the context of EA more than a decade ago - they are currently becoming more prominent again; also, we identified and approved smart machines, sustainability and ''complexity theory'' as major fields that will influence enterprise architecture management in the near future. These topics have not been identified in many earlier reviews [50].

E. SCIENTIFIC COVERAGE OF DISEASES VERSUS BOD (BURDEN OF DISEASE) STUDIES

We have selected the following study in this paper to demonstrate that, for a specific purpose, some steps can be skipped. The aim of the following study is to compare a large quantity of findings from the scientific literature world with a large quantity of findings from another discipline - the comparison to the burden imposed on humans by diseases. We do not need bootstrapping since we use all terms for all diseases that are available within the NLM's (National Library of Medicine) MeSH [98] (medical subject headings) thesaurus listings for diseases. To be exact, we searched for disease names listed in section C. However, we apply some proximity rules.

We want to investigate how the quantity of medical research publications related to diseases is congruent to the output of the WHO burden of disease studies [99]. The problem that we face appears to be the amount and degree of ambiguity of terms used to describe diseases [28]. For example, the MeSH catalog of the year 2017 [100] provides approximately 5000 different terms alone for ''cancer'' (or neoplasms, respectively). However, we managed to map the complete term set of sector C (diseases) of the MeSH thesaurus of 2015 into input queries to search for relevant articles – in addition, we applied a proximity operator and algorithm framework [101] to identify diseases where the sequence of words may vary (e.g., ''connective tissue neoplasms'' and ''neoplasms of the connected tissue'') or where the medical term has been intercepted by other terms (''bacterial infections'' and ''bacterial skin infections''). The search for articles was limited to both the type of publication

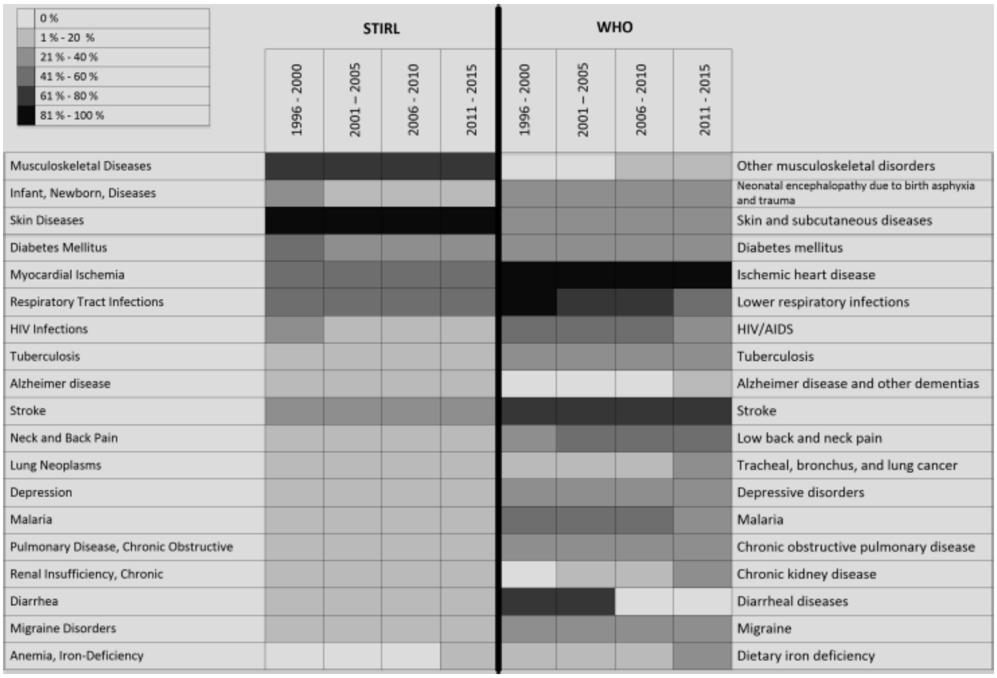


FIGURE 5. Heatmap of STIRL and GBD analysis.

of "Journal Article" and a selection of 10 general highly ranked (H-Index) medical journals that have exclusively been dedicated to major medical research events [72]. After data quality assurance, we analyzed 55024 journal articles, including metadata, keywords and abstracts covering approximately 30 years as taken from the PubMed database [45].

We analyze the results via a simple relevance ranking mechanism and compare the output to the WHO Global Burden of Disease [102] data for the top 20 disease categories that cause the largest number of burdens according to the WHO (see Figure 5). It is interesting to note that, only approximately ten years ago, Jensen *et al.* [28] pointed out that further developments in the area of literature analysis can be expected from integrating results with other data sources.

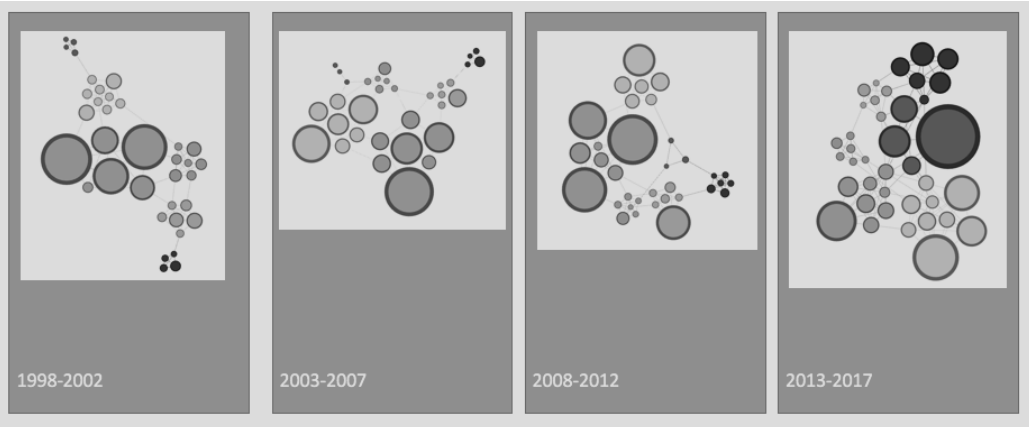
To integrate our results with the WHO database, we apply the MeSH terms as input queries and mapped to codes with the WHO disease classification scheme. We would like to point out that we were able to successfully use the MeSH thesaurus as an input query in contradiction to the findings of Atal *et al.* [102].

By comparing our results with the GBD output, we observe several similarities – the quantitative amount of research publications in major medical journals appears to correlate with the urgency of research in many areas (see Figure 5). However, we also identify some differences (differently shaded sequences, see below). For example, we found that diarrheal diseases that represent a major cause of deaths in the world are not equivalently mentioned in major research journals, while other diseases are covered much more frequently, for example, "HIV Infections" and "Congenital Abnormalities." Alzheimer's disease is underrepresented compared to the burden that is caused by the disease. For breast cancer, for example, we observed an increase in publications on the subject before the disease was listed in the GBD. It became popular before it was included in the WHO listings.

F. THE METAMORPHOSIS OF DATA SCIENCE

Our final case study in this context presents an example of using the entire workflow of the STIRL methodology. We analyze various scientific research topics that belong to "data science" – we regard this term as a universal technology and methodology that includes both "big data" and "artificial intelligence" (among others). The basics of artificial intelligence have already been presented by McCarthy *et al.* in 1956 [103]. Advances in machine learning have been published in the 1950s as well [104]. Both topics continued to develop – in the 1960s, similar topics such as "Library and Information Science" evolved [105], and "Business Intelligence" systems were mentioned as early as 1958 [106] and have been extended with what we know as data warehouse technology [107]. Now, there are also many articles available where "Big Data" [21] has been mentioned as the overwhelming technology. We have also noticed that authors report on interfaces between various disciplines (for example, see [3]) and that many authors try to illustrate comprehensive graphics in which several data science areas are visualized in an interlinked fashion. We analyze this interlinking and examine approx. 50,000 publications with the STIRL methodology to obtain a "clear" systemic visualization and to gain further insight regarding data science.

Figure 6 demonstrates that different areas such as "learning," "artificial intelligence," "business intelligence",



	1998-2002	2003-2007	2008-2012	2013-2017
AI	326	378	738	1202
Analytics	104	194	639	4124
BI	1253	2665	5058	5725
Big Data	35	47	193	6087
Information Systems	309	512	1656	1377
Learning	424	2213	2608	8314

FIGURE 6. The metamorphosis of data science.

''information systems,'' ''big data'' and ''analytics'' had originally been identified as separate topics. However, the intensity of the interconnecting processes is clearly increasing so that we can almost speak about a single conglomerate in the years of 2013-2017. Because all subject areas are encompassed in ''data science,'' we can observe a ''metamorphosis'' or a ''symbiogenesis'' of data science. We predict that, within the next ten years, the technologies will further merge into an all-embracing data science.

V. DISCUSSION AND CONCLUSION

We apply our methodology STIRL with varying mining methods to different literature database outputs. One major difference from ''classical'' review generation procedures is the taxonomy enrichment achieved both by bootstrapping and by applying natural language processing methods. We apply advanced information retrieval technologies by using proximity and truncation operators. Depending on the diversity of a topic, this may easily become much more complicated. However, if the search algorithms consist only of a few basic retrieval operators, it may supersede a manual selection process. In this paper, the artificial intelligence double funnel has only been applied to articles. However, because we consider it principally possible and useful to also examine other big data sources with the double funnel, we have also included ''big data'' in the chart (see Figure 1).

To compile our methodology, we proceeded step by step as illustrated by the case studies described in this article. For the investigations described in this article, we used various commercially available tools (i.e., SAS or RapidMiner). However, we also strive to map the entire process in one step using open-access tools. To check the general practicability of our approach, we utilize the STIRL procedure in teaching. Our students work largely independently on the various steps with their individual topics, and it appears to work for a variety of topics. Within the framework of bachelor and master theses, we further develop and synthesize the technical environment such that we currently operate a one-stop shop featuring open-source tools.

Basically, our current system works as follows. Once the user has entered specific search queries, the service automatically calls various APIs (Application Programming Interfaces), such as ScienceDirect's, and begins the STIRL process described above. Alternatively, users can also start in the middle of the process and upload a finished corpus and then choose between different analytic methods. The current status of our web service ranges from preprocessing, such as removing stop words or lowercase and uppercase letters, calculating the inverse document frequency, Porter-, Schneeball- or Lancaster stemming methods, creating clusters (user-defined or calculated), to final visualization in word clouds or graphical networks. The latter can connect to the interface of the ''Gephi'' program that has been used for the visualizations presented in this article. The user is able to view various graphs and tables in an overview.

To complement our procedure, we are currently developing methods with which we can generate simplified review articles from the graphics or the corpus using natural language generation (NLG). We hereby suggest using the following open-source building blocks:

The web service is a combination of different Python programs as backend services, such as Flask (1.02), flask-restplus (0.12.1) and Flask-SQLAlchemy (2.3.2), as well as Ionic CLI (4.10.3) and Ionic Framework (ionic-angular 3.9.2) for the user interface. For the text mining process, we have used libraries such as matplotlib (3.0.3), nltk (3.4), numpy (1.16.2), pandas (0.24.1), SpeechRecognition (3.6.3), textract (1.6.1), textstat (0.5.5) and wordcloud (1.4.0). For more information, please consult https://www.elsevier.com/solutions/sciencedirect/support/api, https://gephi.org/gephi/0.9.2/apidocs/, https://ionicframework.com/, https://angularjs.org/, https://www.python.org/.

In addition to our continuous technical improvement, we constantly apply quality assurance to obtain reliable artificial intelligence, for example, by testing the precision to recall ratio. We verify the usefulness of our outputs at various stages in the context of several divergent research areas. STIRL appears to be applicable to generate a scientific ''view beyond the horizon'' of any subject; we also believe that the methodology can easily be expanded to other media, such as images, sounds or video sequences; however, additional research is required to prove the general adequacy. We would like to point out that our methodology is designed to be suitable for highly developed topics. It is questionable, for example, whether this method is applicable to very current topics, such as blockchain, or to current developments in the field of quantum computing. Further research on the use of artificial intelligence methods on very contemporary or limited topics is definitely needed. To be able to use artificial intelligence meaningfully in this context, a large amount of data should be used. The research field under investigation needs to provide dynamics and many subtopics to yield meaningful results.

We therefore propose limiting the STIRL procedure to research areas that have been the subject of scientific publications for more than a decade and where several thousand publications are present. However, by analyzing large literature sets with information retrieval, text mining or other artificial intelligence technologies, it should also be possible to identify neglected areas of interest. To fully complete the procedure, it should also be possible to write the review article using text generation [108]. We are currently working on realizing this final step.

REFERENCES

[1] R. Johnson, A. Watkinson, and M. Mabe, ''The STM Report: An overview of scientific and scholarly publishing,'' Int. Assoc. Sci., Tech. Med., The Hague, The Netherlands, Tech. Rep., 2018. [Online]. Available: https://www.stm-assoc.org/2018_10_04_STM_Report_2018.pdf

[2] X. Gu and K. L. Blackmore, ''Recent trends in academic journal growth,'' *Scientometrics*, vol. 108, no. 2, pp. 693–716, 2016.

[3] A. Moreno and T. Redondo, ''Text analytics: The convergence of big data and artificial intelligence,'' *Int. J. Interact. Multimed. Artif. Intell.*, vol. 3, no. 6, pp. 57–64, 2016.

[4] J. vom Brocke, A. Simons, B. Niehaves, K. Riemer, and R. Plattfaut, ''Reconstructing the giant: On the importance of rigour in documenting the literature search process,'' in *Proc. 17th Eur. Conf. Inf. Syst.*, vol. 9, Jun. 2009, pp. 2206–2217.

[5] A. Gliozzo, C. Strapparava, and I. Dagan, ''Improving text categorization bootstrapping via unsupervised learning,'' *ACM Trans. Speech Lang. Process.*, vol. 6, no. 1, Oct. 2009, Art. no. 1.

[6] M. Tate, E. Furtmueller, J. Everman, and W. Bandara, ''Introduction to the special issue: The literature review in information systems,'' *Commun. Assoc. Inf. Syst.*, vol. 37, pp. 103–111, Aug. 2015.

[7] J. Webster and R. T. Watson, ''Analyzing the past to prepare for the future: Writing a literature review,'' *MIS Quart.*, vol. 26, no. 2, pp. 13–23, 2002.

[8] W. Bandara, E. Furtmueller, E. Gorbacheva, S. Miskon, and J. Beekhuyzen, ''Achieving rigor in literature reviews: Insights from qualitative data analysis and tool-support,'' *Commun. Assoc. Inf. Syst.*, vol. 37, no. 8, pp. 154–204, 2015.

[9] G. Schryen, ''Writing qualitative is literature reviews—Guidelines for synthesis, interpretation, and guidance of research,'' *Commun. Assoc. Inf. Syst.*, vol. 37, pp. 286–325, Aug. 2015.

[10] J. vom Brocke, A. Simons, K. Riemer, B. Niehaves, R. Plattfaut, and A. Cleven, ''Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research,'' *Commun. Assoc. Inf. Syst.*, vol. 37, no. 8, pp. 205–224, 2015.

[11] C. J. Pannucci and E. G. Wilkins, ''Identifying and avoiding bias in research,'' *Plastic Reconstruct. Surg.*, vol. 126, no. 2, pp. 619–625, 2010.

[12] B. Howard, J. T. Scott, M. Blubaugh, B. Roepke, C. Scheckel, and M. Vassar, ''Systematic review: Outcome reporting bias is a problem in high impact factor neurology journals,'' *PLoS ONE*, vol. 12, no. 7, 2017, Art. no. e0180986.

[13] C. Cooper, A. Booth, J. Varley-Campbell, N. Britten, and R. Garside, ''Defining the process to literature searching in systematic reviews: A literature review of guidance and supporting studies,'' *BMC Med. Res. Methodol.*, vol. 18, no. 1, p. 85, 2018.

[14] A. O'Mara-Eves, J. Thomas, J. McNaught, M. Miwa, and S. Ananiadou, ''Using text mining for study identification in systematic reviews: A systematic review of current approaches,'' *Syst. Rev.*, vol. 4, no. 1, p. 5, 2015.

[15] A. C. McKinnon, ''Starry-eyed: Journal rankings and the future of logistics research,'' *Int. J. Phys. Distrib. Logistics Manage.*, vol. 43, no. 1, pp. 6–17, 2013.

[16] J. M. Hines, H. R. Hungerford, and A. N. Tomera, ''Analysis and synthesis of research on responsible environmental behavior: A meta-analysis,'' *J. Environ. Educ.*, vol. 18, no. 2, pp. 1–8, 2010.

[17] Y. Xiao and M. Watson, ''Guidance on conducting a systematic literature review,'' *J. Planning Educ. Res.*, vol. 39, no. 1, pp. 93–112, Aug. 2017.

[18] D. Liu, Y. Li, and M. A. Thomas, ''A roadmap for natural language processing research in information systems,'' in *Proc. 50th Hawaii Int. Conf. Syst. Sci.*, 2017, pp. 1112–1121.

[19] A. MacFarlane and A. Tuson, ''Local search: A guide for the information retrieval practitioner,'' *Inf. Process. Manage.*, vol. 45, no. 1, pp. 159–174, 2009.

[20] V. Gupta and G. S. Lehal, ''A survey of text mining techniques and applications,'' *J. Emerg. Technol. Web Intell.*, vol. 1, no. 1, pp. 60–76, Aug. 2009.

[21] D. Che, M. Safran, and Z. Peng, ''From big data to big data mining: Challenges, issues, and opportunities,'' in *Proc. Int. Conf. Database Syst. Adv. Appl.*, 2013, pp. 1–15.

[22] H. Chen, R. R. H. L. Chiang, and V. C. V. Storey, ''Business intelligence and analytics: From big data to big impact,'' *MIS Quart.*, vol. 36, no. 4, pp. 1165–1188, 2012.

[23] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, ''From data mining to knowledge discovery in databases,'' *AI Mag.*, vol. 17, no. 3, pp. 37–54, 1996.

[24] G. Arji, R. Safdari, H. Rezaeizadeh, A. Abbassian, M. Mokhtaran, and M. H. Ayati, ''A systematic literature review and classification of knowledge discovery in traditional medicine,'' *Comput. Methods Programs Biomed.*, vol. 168, pp. 39–57, Jan. 2019.

[25] M. A. Andrade and P. Bork, ''Automated extraction of information in molecular biology,'' *FEBS Lett.*, vol. 476, nos. 1–2, pp. 12–17, 2000.

[26] J. C. Carver, E. Hassler, E. Hernandes, and N. A. Kraft, ''Identifying barriers to the systematic literature review process,'' in *Proc. ACM/IEEE Int. Symp. Empirical Softw. Eng. Meas.*, Oct. 2013, pp. 203–212.

[27] J. Thomas, J. McNaught, and S. Ananiadou, ''Applications of text mining within systematic reviews,'' *Res. Synth. Methods*, vol. 2, no. 1, pp. 1–14, 2011.

[28] L. J. Jensen, J. Saric, and P. Bork, "Literature mining for the biologist: From information retrieval to biological discovery," *Nature Rev. Genet.*, vol. 7, no. 2, pp. 119–129, Feb. 2006.
[29] Z. Yu and T. Menzies, "FAST^2: An intelligent assistant for finding relevant papers," *Expert Syst. Appl.*, vol. 120, pp. 57–71, Apr. 2019.
[30] U. Sivarajah, M. M. Kamal, Z. Irani, and V. Weerakkody, "Critical analysis of big data challenges and analytical methods," *J. Bus. Res.*, vol. 70, pp. 263–286, Jan. 2017.
[31] J. Niu and R. R. A. Issa, "Developing taxonomy for the domain ontology of construction contractual semantics: A case study on the AIA A201 document," *Adv. Eng. Inform.*, vol. 29, no. 3, pp. 472–482, 2015.
[32] B. R. Schatz, "Information retrieval in digital libraries: Bringing search to the net," *Science*, vol. 275, no. 5298, pp. 327–334, Jan. 1997.
[33] P. Gupta, K. Bali, R. E. Banchs, M. Choudhury, and P. Rosso, "Query expansion for mixed-script information retrieval," in *Proc. 37th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, vol. 14, 2014, pp. 677–686.
[34] S. Mukherjea, "Information retrieval and knowledge discovery utilising a biomedical Semantic Web," *Brief. Bioinf.*, vol. 6, no. 3, pp. 252–262, 2005.
[35] A. Kayed and R. M. Colomb, "Extracting ontological concepts for tendering conceptual structures," *Data Knowl. Eng.*, vol. 40, no. 1, pp. 71–89, 2002.
[36] R. Navigli and P. Velardi, "Ontology enrichment through automatic semantic annotation of on-line glossaries," in *Proc. Int. Conf. Knowl. Eng. Knowl. Manage.*, 2006, pp. 126–140.
[37] H. Wimmer and L. Zhou, "Word sense disambiguation for ontology learning," in *Proc. 19th Amer. Conf. Inf. Syst.*, 2013, pp. 1–10.
[38] M. Missikoff, P. Velardi, and P. Fabriani, "Text mining techniques to automatically enrich a domain ontology," *Appl. Intell.*, vol. 18, no. 3, pp. 323–340, 2003.
[39] J. R. Erskine, G. L. Peterson, B. E. Mullins, and M. R. Grimaila, "Developing cyberspace data understanding: Using CRISP-DM for host-based IDS feature mining," in *Proc. 6th Annu. Workshop Cyber Secur. Inf. Intell. Res. (CSIIRW)*, 2010, Art. no. 74.
[40] W. Gersten, R. Wirth, and D. Arndt, "Predictive modeling in automotive direct marketing: Tools, experiences and open issues," in *Proc. 6th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2000, pp. 398–406.
[41] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, Sep. 1999.
[42] T. Iwata, K. Saito, and T. Yamada, "Recommendation method for improving customer lifetime value," *IEEE Trans. Knowl. Data Eng.*, vol. 20, no. 9, pp. 1254–1263, Sep. 2008.
[43] O. L. Q. Montoya, L. F. Villa, S. Muñoz, A. C. R. Arenas, and M. Bastidas, "Information retrieval on documents methodology based on entropy filtering methodologies," *Int. J. Bus. Intell. Data Mining*, vol. 10, no. 3, pp. 280–296, 2015.
[44] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, no. 2, pp. 224–227, Apr. 1979.
[45] H. Chen and B. M. Sharp, "Content-rich biological network constructed by mining PubMed abstracts," *BMC Bioinf.*, vol. 5, no. 1, p. 147, 2004.
[46] J. Hartley, "Improving the clarity of journal abstracts in psychology: The case for structure," *Sci. Commun.*, vol. 24, no. 3, pp. 366–379, Mar. 2003.
[47] K. R. Felizardo, E. Y. Nakagawa, D. Feitosa, R. Minghim, and J. C. Maldonado, "An approach based on visual text mining to support categorization and classification in the systematic mapping," in *Proc. 14th Int. Conf. Eval. Assessment Softw. Eng. (EASE)*, 2010, pp. 1–10.
[48] W. Fan, L. Wallace, S. Rich, and Z. Zhang, "Tapping the power of text mining," *Commun. ACM*, vol. 49, no. 9, pp. 76–82, 2006.
[49] P. Chapman *et al.*, "CRISP-DM 1.0," CRISP-DM Consortium, Tech. Rep., 2000, p. 76.
[50] F. Gampfer, A. Jürgens, M. Müller, and R. Buchkremer, "Past, current and future trends in enterprise architecture—A view beyond the horizon," *Comput. Ind.*, vol. 100, pp. 70–84, Sep. 2018.
[51] H. F. Moed and G. Halevi, "On full text download and citation distributions in scientific-scholarly journals," *J. Assoc. Inf. Sci. Technol.*, vol. 67, no. 2, pp. 412–431, Feb. 2016.
[52] C. W. Belter, "Citation analysis as a literature search method for systematic reviews," *J. Assoc. Inf. Sci. Technol.*, vol. 67, no. 11, pp. 2766–2777, 2016.
[53] H.-F. Hsieh and S. E. Shannon, "Three approaches to qualitative content analysis," *Qualitative Health Res.*, vol. 15, no. 9, pp. 1277–1288, Nov. 2005.
[54] A. Banerjee and S. Basu, "Topic models over text streams: A study of batch and online unsupervised learning," in *Proc. SIAM Int. Conf. Data Mining*, 2007, vol. 14, no. 2, pp. 431–436.
[55] L. Yao *et al.*, "Incorporating knowledge graph embeddings into topic modeling," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 3119–3126.
[56] J. H. Lau, K. Grieser, D. Newman, and T. Baldwin, "Automatic labelling of topic models," in *Proc. 49th Annu. Meeting Assoc. Comput. Linguistics, Hum. Lang. Technol.*, 2011, pp. 1536–1545.
[57] A. O. Steinskog, J. F. Therkelsen, and B. Gambäck, "Twitter topic modeling by tweet aggregation," in *Proc. 21st Nordic Conf. Comput. Linguistics*, May 2017, pp. 77–86.
[58] S. Saitta, B. Raphael, and I. F. C. Smith, "A bounded index for cluster validity," in *Machine Learning and Data Mining in Pattern Recognition*. Berlin, Germany: Springer, 2007, pp. 174–187. doi: 10.1007/978-3-540-73499-4_14.
[59] Y. Liu, Z. Li, H. Xiong, X. Gao, and J. Wu, "Understanding of internal clustering validation measures," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Dec. 2010, pp. 911–916.
[60] N. N. Karanikolas, "Supervised learning for building stemmers," *J. Inf. Sci.*, vol. 41, no. 3, pp. 315–328, Mar. 2015.
[61] D. Erhan, Y. Bengio, A. Courville, P.-A. Manzagol, P. Vincent, and S. Bengio, "Why does unsupervised pre-training help deep learning?" *J. Mach. Learn. Res.*, vol. 11, pp. 625–660, Feb. 2010.
[62] A. A. Lopes, R. Pinho, F. V. Paulovich, and R. Minghim, "Visual text mining using association rules," *Comput. Graph.*, vol. 31, no. 3, pp. 316–326, 2007.
[63] M. V. Mäntylä, D. Graziotin, and M. Kuutila, "The evolution of sentiment analysis—A review of research topics, venues, and top cited papers," *Comput. Sci. Rev.*, vol. 27, pp. 16–32, Feb. 2018.
[64] S. Lee, J. Song, and Y. Kim, "An empirical comparison of four text mining methods," *J. Comput. Inf. Syst.*, vol. 51, no. 1, pp. 1–10, 2010.
[65] L. Y. Y. Lu, C.-H. Hsieh, and J. S. Liu, "Development trajectory and research themes of foresight," *Technol. Forecasting Social Change*, vol. 112, pp. 347–356, Nov. 2016.
[66] A. Kontostathis, L. M. Galitsky, W. M. Pottenger, S. Roy, and D. J. Phelps, "A survey of emerging trend detection in textual data mining," in *Survey of Text Mining*. New York, NY, USA: Springer, 2004, pp. 185–224.
[67] J. F. Hair, Jr., "Knowledge creation in marketing: The role of predictive analytics," *Eur. Bus. Rev.*, vol. 19, no. 4, pp. 303–315, Jul. 2007.
[68] T. H. Szymanski and S. Behdin, "Traffic provisioning in a future Internet," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2012, pp. 5872–5878.
[69] P. O. Seglen, "Why the impact factor of journals should not be used for evaluating research," *Brit. Med. J.*, vol. 314, no. 7079, pp. 498–502, 1997.
[70] M. H. MacRoberts and B. R. MacRoberts, "Problems of citation analysis: A study of uncited and seldom-cited influences," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 61, no. 1, pp. 1–12, 2010.
[71] M. Yassine and H. Hajj, "A framework for emotion mining from text in online social networks," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Dec. 2010, pp. 1136–1142.
[72] J. E. Hirsch, "An index to quantify an individual's scientific research output," *Proc. Nat. Acad. Sci. USA*, vol. 102, no. 46, pp. 16569–16572, Nov. 2005.
[73] R. E. LeBlanc and B. Quintiliano, "Recycling C.R.A.P.: Reframing a popular research mnemonic for library instruction," *Pennsylvania Libraries, Res. Pract.*, vol. 3, no. 2, pp. 115–121, 2015.
[74] J. D. West, M. C. Jensen, R. J. Dandrea, G. J. Gordon, and C. T. Bergstrom, "Author-level Eigenfactor metrics: Evaluating the influence of authors, institutions, and countries within the social science research network community," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 64, no. 4, pp. 787–801, 2013.
[75] R. Rekik, I. Kallel, J. Casillas, and A. M. Alimi, "Assessing web sites quality: A systematic literature review by text and association rules mining," *Int. J. Inf. Manage.*, vol. 38, no. 1, pp. 201–216, 2018.
[76] K. Kucher and A. Kerren, "Text visualization techniques: Taxonomy, visual survey, and community insights," in *Proc. IEEE Pacific Vis. Symp.*, Apr. 2015, pp. 117–121.
[77] M. Bastian, S. Heymann, and M. Jacomy, "Gephi: An open source software for exploring and manipulating networks," in *Proc. Int. AAAI Conf. Web Social Media*, 2009, pp. 361–362.
[78] U. Bubeck and H. K. Büning, "Learning Boolean specifications," *Artif. Intell.*, vol. 229, pp. 246–257, Dec. 2015.

[79] D. Krimpmann, "IT/IS organisation design in the digital age—A literature review," *Int. J. Social, Behav., Educ., Econ., Bus. Ind. Eng.*, vol. 9, no. 4, pp. 1208–1218, 2015.
[80] E. W. T. Ngai, "Customer relationship management research (1992–2002): An academic literature review and classification," *Marketing Intell. Planning*, vol. 23, no. 6, pp. 582–605, 2005.
[81] O. Wahlberg, C. Strandberg, H. Sundberg, and K. W. Sandberg, "Trends, topics and under-researched areas in CRM research: A literature review," *Int. J. Public Inf. Syst.*, vol. 2009, no. 3, pp. 191–208, 2009.
[82] E. W. T. Ngai, L. Xiu, and D. C. K. Chau, "Application of data mining techniques in customer relationship management: A literature review and classification," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 2592–2602, 2009.
[83] Z. Soltani and N. J. Navimipour, "Customer relationship management mechanisms: A systematic review of the state of the art literature and recommendations for future research," *Comput. Hum. Behav.*, vol. 61, pp. 667–688, Aug. 2016.
[84] C. Dirican, "The impacts of robotics, artificial intelligence on business and economics," *Procedia-Social Behav. Sci.*, vol. 195, pp. 564–573, Jul. 2015.
[85] S. Diffley, P. McCole, and E. Carvajal-Trujillo, "Examining social customer relationship management among Irish hotels," *Int. J. Contemp. Hospitality Manage.*, vol. 30, no. 2, pp. 1072–1091, 2018.
[86] S. E. Griffis, S. Rao, T. J. Goldsby, and T. T. Niranjan, "The customer consequences of returns in online retailing: An empirical analysis," *J. Oper. Manage.*, vol. 30, no. 4, pp. 282–294, May 2012.
[87] M. Klumpp, H. Abidi, S. Bioly, R. Buchkremer, S. Ebener, and G. Sandhaus, "Logistics dynamics and demographic change," in *Dynamics in Logistics*, M. Freitag, H. Kotzab, and J. Pannek, Eds. Cham, Switzerland: Springer, 2017, pp. 347–362. doi: 10.1007/978-3-319-45117-6_31.
[88] M. D. Uncles, G. R. Dowling, and K. Hammond, "Customer loyalty and customer loyalty programs," *J. Consum. Marketing*, vol. 20, no. 4, pp. 294–316, 2003.
[89] K. Wegrzyn-Wolska, L. Bougueroua, and G. Dziczkowski, "Social media analysis for e-health and medical purposes," in *Proc. Int. Conf. Comput. Aspects Social Netw. (CASoN)*, Oct. 2011, vol. 11, no. 1, pp. 278–283.
[90] S. Tkatek, O. Abdoun, and J. Abouchabaka, "An expert system for a constrained mobility management of human resources," in *Proc. Int. Colloq. Logistics Supply Chain Manage. (LOGISTIQUA)*, Apr. 2017, pp. 53–58.
[91] J. A. Zachman, "A framework for information systems architecture," *IBM Syst. J.*, vol. 26, no. 3, pp. 276–292, 1987.
[92] J. Ortigosa-Hernández, J. D. Rodríguez, L. Alzate, M. Lucania, I. Inza, and J. A. Lozano, "Approaching Sentiment Analysis by using semi-supervised learning of multi-dimensional classifiers," *Neurocomputing*, vol. 92, pp. 98–115, Sep. 2012.
[93] J. Dougherty, R. Kohavi, and M. Sahami, "Supervised and unsupervised discretization of continuous features," in *Proc. 12th Int. Conf. Mach. Learn.*, 1995, pp. 194–202.
[94] D.-Y. Chiu and Y.-C. Pan, "Topic knowledge map and knowledge structure constructions with genetic algorithm, information retrieval, and multi-dimension scaling method," *Knowl.-Based Syst.*, vol. 67, pp. 412–428, Sep. 2014.
[95] A. Ibáñez, R. Armañanzas, C. Bielza, and P. Larrañaga, "Genetic algorithms and Gaussian Bayesian networks to uncover the predictive core set of bibliometric indices," *J. Assoc. Inf. Sci. Technol.*, vol. 67, no. 7, pp. 1703–1721, Jul. 2016.
[96] J. Janulevičius, L. Marozas, A. Čenys, N. Goranin, and S. Ramanauskaitė, "Enterprise architecture modeling based on cloud computing security ontology as a reference model," in *Proc. Open Conf. Elect., Electron. Inf. Sci. (eStream)*, Apr. 2017, pp. 1–6.
[97] G. Weichhart, A. Molina, D. Chen, L. E. Whitman, and F. Vernadat, "Challenges and current developments for sensing, smart and sustainable enterprise systems," *Comput. Ind.*, vol. 79, pp. 34–46, Jun. 2016.
[98] A. J. Yepes, J. G. Mork, B. Wilkowski, D. D. Fushman, and A. R. Aronson, "MEDLINE MeSH indexing: Lessons learned from machine learning and future directions," in *Proc. 2nd ACM SIGHIT Int. Health Inform. Symp. (IHI)*, 2012, pp. 737–742.
[99] J. D. Stanaway *et al.*, "The global burden of dengue: An analysis from the global burden of disease study 2013," *Lancet Infectious Diseases*, vol. 16, no. 6, pp. 712–723, Jun. 2016.
[100] Y. Yamamoto and T. Takagi, "Biomedical knowledge navigation by literature clustering," *J. Biomed. Inf.*, vol. 40, no. 2, pp. 114–130, 2007.
[101] M. Beigbeder, M. Géry, and C. Largeron, "Using proximity and tag weights for focused retrieval in structured documents," *Knowl. Inf. Syst.*, vol. 44, no. 1, pp. 51–76, Jul. 2015.
[102] I. Atal, J.-D. Zeitoun, A. Névéol, P. Ravaud, R. Porcher, and L. Trinquart, "Automatic classification of registered clinical trials towards the global burden of diseases taxonomy of diseases and injuries," *BMC Bioinf.*, vol. 17, no. 1, p. 392, Dec. 2016.
[103] J. McCarthy, M. L. Minsky, N. Rochester, and C. E. Shannon, "A proposal for the dartmouth summer research project on artificial intelligence, August 31, 1955," *AI Mag.*, vol. 27, no. 4, p. 12, 2006.
[104] W. H. Ware, "Introduction to session on learning machines," in *Proc. Western Joint Comput. Conf. (AFIPS)*, Mar. 1955, pp. 85–86.
[105] A. Gilchrist, "Library and information science abstracts: A brief progress report," *Aslib Proc.*, vol. 21, no. 8, pp. 325–327, Aug. 1969.
[106] H. P. Luhn, "A business intelligence system," *IBM J. Res. Develop.*, vol. 2, no. 4, pp. 314–319, Oct. 1958.
[107] B. H. Wixom and H. J. Watson, "An empirical investigation of the factors affecting data warehousing success," *MIS Quart.*, vol. 25, no. 1, pp. 17–41, Mar. 2001.
[108] G. Salton, A. Singhal, M. Mitra, and C. Buckley, "Automatic text structuring and summarization," *Inf. Process. Manage.*, vol. 33, no. 2, pp. 193–207, 1997.