Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

TOWARDS AN AUTOMATIC CONSTRUCTION OF SIMULATION SCENARIOS: A SYSTEMATIC REVIEW

Christopher W.H. Davis

Microsoft
1414 NW Northrup
Portland, OR 97209, USA

Antonie J. Jetter

Dept. of Engineering & Technology Management
Portland State University
Portland, OR 97201, USA

Philippe J. Giabbanelli

Dept. of Computer Science & Software Engineering
Miami University
Oxford, OH 45056, USA

ABSTRACT

A predictive simulation is built on a conceptual model (e.g., to identify relevant constructs and relationships) and serves to estimate the potential effects of 'what-if' scenarios. Developing the conceptual model and plausible scenarios has long been a time-consuming activity, often involving the manual processes of identifying and engaging with experts, then performing desk research, and finally crafting a compelling narrative about the potential futures captured as scenarios. Automation could speed-up these activities, particularly through text mining. We performed the first review on automation for simulation scenario building. Starting with 420 articles published between 1995 and 2022, we reduced them to 11 relevant works. We examined them through four research questions concerning data collection, extraction of individual elements, connecting elements of insight and (degree of automation of) scenario generation. Our review identifies opportunities to guide this growing research area by emphasizing consistency and transparency in the choice of datasets or methods.

1 INTRODUCTION

Model commissioners use predictive simulations to identify the consequences of 'potential futures', known as *what-if scenarios*. These scenarios are often crafted by practitioners as stories (Bowman, MacKay, Masrani, and McKiernan 2013). Originally developed by industry practitioners to manage future uncertainty, the field of *scenario planning* has not only grown in popularity over time, but has remained focused on practice, so that over half of the publications on scenarios are about methodology (Tiberius, Siglow, and Sendra-García 2020). There are several schools of thought when it comes to scenario planning and their methods span a continuum from purely qualitative techniques that rely on intuition and experience to formal, quantitative techniques that use, among others, statistical analysis (Amer, Daim, and Jetter 2013). At the beginning of any scenario study, a *topic* or domain is chosen and research is performed to identify and analyze influencing *factors* that shape the future. The next analytical step focuses on a subset of factors that are understood to have strong influence on future developments and for which the future states are uncertain. For example, a scenario planner estimates that the factor might have a future value ranging from "low" to "medium high", but cannot confidently forecast a specific (or 'crisp') value. The goal of scenario planning is to identify plausible combinations of outcome states of scenario drivers, based on the assumption that

979-8-3503-6966-3/23/$31.00 ©2023 IEEE

the drivers either impact each other directly or because they are impacted by the same underlying trends. For instance, a public health simulation model for suicide prevention may include parameters such as minimum wage, after school programs, and lethality of suicide attempts. A scenario team may agree that a combination of "increase in minimum wage by $2 per hour, no additional after schools programs, lower lethality by 15% due to secure gun storage" could occur together and decide to run these values in the simulation. The same team may reject combination of factor that include a decrease in minimum wage or an increase in lethality, as they cannot think of any way in which such situations could occur. The goal of scenario planning is to identify a small number (typically 3-8) of plausible, alternative scenarios that are quite different from each other and, together, cover much of the "cone of uncertainty" about the future (Amer, Daim, and Jetter 2013). These alternative futures are subsequently described as "*narratives*" to make them accessible to planning teams (Lindgren and Bandhold 2009).

Compelling and relevant scenarios are instrumental to support the applicability of a simulation model, yet the process is very time consuming. Automating the development of scenarios has thus been recognized as a valuable possibility (Raford 2015). For example, computational methods applied to Big Data (and particularly online data) have been suggested as a way to rapidly assemble evidence that can potentially speed up the process of scenario development (Raford 2015; Pillutla and Giabbanelli 2019). Some have explored *Foresight Support Systems* to aid in the creation of scenarios or other types of foresight methodologies (Keller et al. 2015). In particular, *text mining* has been linked to scenario development because the benefits of text mining closely align with the stages of scenario development (Kayser and Blind 2017). In scenario planning, teams typically start with interviews before moving into a desk-research phase where interviews are analyzed and combined with other data sources. These steps could potentially benefit from automation in the form of text mining from big data, which would allow to get insights from large volumes of text by finding topics and their interrelationships. Similarly, the use of Natural Language Processing (NLP) can derive higher-level meanings from the text, thus serving to identify the importance of certain topics within a corpus, and assemble such topics to create the skeleton of scenarios. Said otherwise, the use of text mining and NLP can augment or even dramatically speed up efforts for the scenario development process.

The potential for automation in scenario building has only been recently explored. In a 2022 review on data analytics in the broader domain of technological forecasting, scenario planning was only mentioned in the context of expert opinions (e.g., obtained through interviews), without any automation (Feng et al. 2022). An earlier synthesis from 2021 found only two studies (from 2016 and 2018) where automation was used to 'generate scenario concepts' (Lee 2021). The emergence of automated scenario development was in its infancy and closely mirrored broader practices in computational methods for technological forecasting (Lee 2021). Specifically, studies performed text preprocessing (e.g., tokenization, stop word removal), topic extraction (keyword extraction via tf-idf and topic modeling via LDA methods detailed in Section 2.2), and clustering (Kim et al. 2019; Jang et al. 2021).

Scenario planning is crucial when it comes to using a simulation model for prediction and decision-making. Computational methods such as text mining can speed-up the process of creating useful scenarios. While several approaches have been explored, there are still open challenges. A systematic review could therefore serve as a starting point for future research in those areas and accelerate the progress in level of automation. The present review contributes to this need by examining how the literature has used computational methods to speed up the process of scenario building. Specifically, we conducted a systematic review of text mining in scenario development to identify articles up to February 2022, going as far back as 1995. Although previous reviews have already covered much of this period (Feng et al. 2022; Lee 2021), they only considered the top 10 (Lee 2021) or top 50 journals in the field of technology and innovation management (Feng et al. 2022). However, the emerging practice of automating scenarios is of interest for Modeling & Simulation in general, rather than solely for innovation management. This may explain the paucity of relevant studies found by targeting journals within a single field. In contrast, our study uses multiple databases and analyzes relevant articles with respect to four *Research Questions* (RQs), which correspond to the successive steps of the scenario development process:

RQ1 What was the approach to *data collection*? In particular, we examine whether it is automated and which sources are used (e.g., Web scraping, social media). If data sources readily used in a simulation project can also serve to define scenarios, then efforts would be lessened for the simulation team. For instance, a building simulation could already be extracting occupancy schedules from social media (Lu, Feng, Pang, Yang, and O'Neill 2021) and thus use the same source to craft scenarios that define potential changes (e.g., transitioning from personal offices to open spaces).

RQ2 Once data was collected, how were elements *extracted* from the corpus? This extraction may be manual (for a small corpus) or automatized, for instance through topic modeling.

RQ3 After individual elements have been obtained, how are they combined to *derive insight*? They may be structured into a conceptual model, which is a starting point for a simulation and thus naturally aligns with the development of a simulation. Alternatively, insight can be achieved by classification models, or by a manual analysis.

RQ4 Finally, how are *scenarios generated*? This may employ a dedicated, specialized interface through which participants guide the generation process. Such an interface may be integrated with a simulation platform. Alternatively, a more fluid approach may be taken, centered on focus groups.

2 BACKGROUND

2.1 Approaches to Scenario Planning

Qualitative and quantitative scenario techniques (and the many mixed approaches in between) differ in how knowledge about factors, their future states, and plausible combinations of factor states is integrated (Voinov et al. 2018). Purely qualitative techniques trust planning experts to integrate their knowledge into plausible futures by mapping connections between factors (Chakraborty and McMillan 2015) using drawings or other visualizations (Buckman et al. 2019). Fully quantitative techniques might estimate the range of uncertainty for each factor through statistical techniques, thus informing models in terms of the structural uncertainty regarding relations between factors (Bishop, Hines, and Collins 2007) or through the parametric uncertainty surrounding each factor (Vilkkumaa et al. 2018). Some of the most common quantitative criteria include *plausibility*, where there is a conceivable path to a certain potential scenario from the present state (Van der Heijden 2005; Alcamo and Henrichs 2008; Bradfield, Derbyshire, and Wright 2016; Spaniol and Rowland 2019); *internal consistency*, to ensure scenarios do not contradict themselves (Bradfield, Derbyshire, and Wright 2016; Spaniol and Rowland 2019); and *relevance* of the scenarios to the team (Van der Heijden 2005; Alcamo and Henrichs 2008; Durance and Godet 2010). There are several other measures which include differentiation between a set of selected scenarios, the creativity of the scenarios (Alcamo and Henrichs 2008), whether they reflect uncertainty and help people create new perspectives (Van der Heijden 2005), and how transparent they are (Durance and Godet 2010).

While a *model* can be validated, a *scenario* is not usually validated through the same comparison of point estimates (Chermack et al. 2001). Rather, the validation focuses on the *process* of scenario building by ensuring that each step is conducted rigorously, that all relevant drivers are identified, that their future states are carefully investigated, that factor combinations are plausible, and that the resulting scenarios are meaningful and relevant for decisions they are supposed to support (Alcamo and Henrichs 2008). The many different process steps make it necessary to involve several different roles in the scenario planning process. *Domain experts* (a.k.a. subject-matter experts) are people with specialized knowledge (based on research and practice) in a particular domain such as technology trends, changes to the marketplace, or competitors. They can be internal or external to an organization. *Stakeholders* represent various roles within an organization that have a vested interest in the generated scenarios. Stakeholders can include managers, team leaders, executives, strategic marketing, or product developers. Individuals who requested the simulation model (i.e., model commissioners) can be domain experts and/or stakeholders, as these categories are not exclusive. Figure 1 represents their involvement in the different stages of the scenario process.

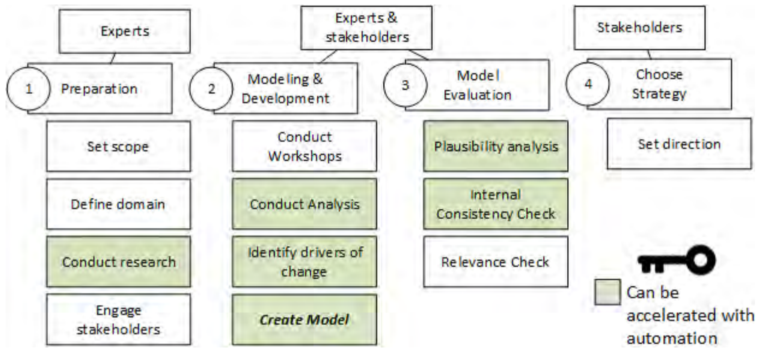

Figure 1: General steps and involvement of participants in scenario development. Our review covers up to the creation of the model, thus we focus on the first two steps (preparation and development). Scenarios are used in the final step by stakeholders, but the use of scenarios or public communications about simulation results are beyond the scope of this review.

2.2 Applied Text Mining

The process of extracting insights from data is characterized as "mining" hence the term data mining is used for any type of data. Web mining is used when the data exists in the World Wide Web, and *text mining* (or text data mining) is used for written resources consisting of unstructured text (Macanovic 2022). To make data accessible for mining, it needs to be collected and made available in a format that can be mined. Techniques that support or automate this process are characterized as "*scraping*". For example, web scraping, web harvesting, or web data extraction are all used for techniques that extract data from websites. These websites can be a focused community that discusses the specific domain of interest (Jang, Park, and Seol 2021), patents databases (Kim et al. 2019), publication databases (Lee 2021), or online communities where experts share scenarios (e.g., Gartner, Futurism, Kurzweil Accelerating Intelligence, World Future Society, Future TimeLine, MIT Technology Review),

People use different vocabularies and speech patterns in different contexts, so that the language used in a specific online community differs from, for example, the language used in scientific journals. Spoken language also differs greatly from written texts. Accordingly, there are many approaches to make sense of language by understanding the context in which it is used. The associated field of research and practice, *Natural Language Processing* (NLP), broadly refers to the study and development of computer systems that can interpret speech and text as humans naturally speak or write. Once data sources are identified and text is retrieved, it often undergoes a *pre-processing* step. For instance, words that are not deemed informative can be removed (e.g., 'a', 'an', 'the'), and words can be mapped onto a root form and synonyms can be combined to facilitate the identification of frequent words (e.g., lemmatization). These objectives may be realized through generic lexical databases for English (e.g., WordNet) or via specialized databases for the application domain of the scenario (e.g., the TechWord database for technological information by Jang, Jeong, and Yoon 2021). Note that the tasks associated with pre-processing (e.g., word removal) are not necessarily performed *as soon as* the data is obtained: as shown in Figure 2, certain pipelines only start applying these tasks to filter potential candidates obtained by an intermediate algorithm.

NLP can remove bottlenecks of the scenario creation process by finding insights in large amounts of text data. One method used to achieve this goal is *Topic Modelling* (TM), which uses a statistical approach for discovering the abstract 'topics' that occur in a collection of documents. That is, Topic Modeling is a form of unsupervised Machine Learning (ML) that takes in unstructured data and identifies themes based on the words in the corpus. This is made possible using distributional semantics, where co-occurrence, similarity, and frequency of words are used to determine their meaning. One popular implementation is called *Latent Semantic Analysis* (LSA), which produces a set of concepts related to the documents and

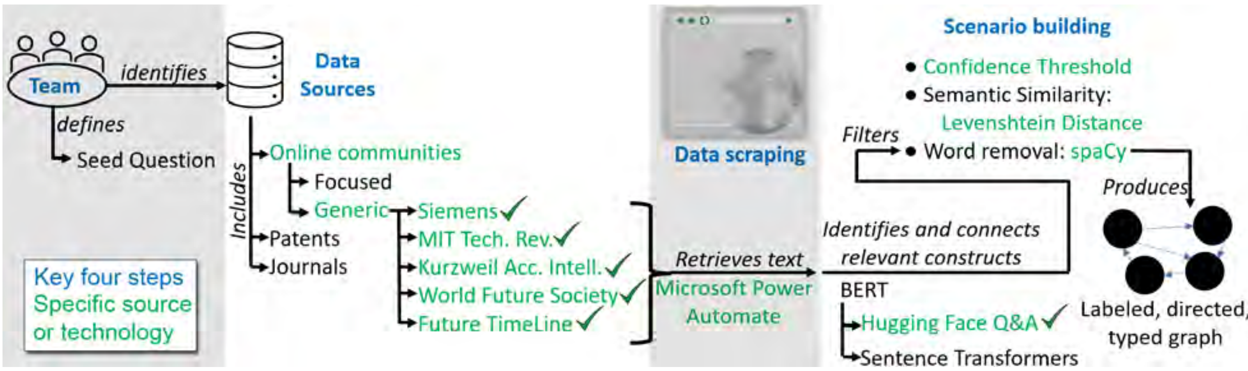

Figure 2: Process employed in (Davis, Jetter, and Giabbanelli 2022) to build scenarios from text. Some elements of the process require the attention of experts (e.g., specifying the three filters, identifying sources) while others are fully automated. A scenario is returned as a graph.

terms. Other common techniques are the *Latent Dirichlet Allocation* (LDA) or word2vec, which convert words into vectors then apply clustering algorithms to establish groups based on uses within the text.

As exemplified in Figure 2, a study can involve a variety of data sources and articulate several technologies to generate candidate scenarios from text and filter them. In this example, a question-answering (Q&A) BERT machine learning module serves to extract and connect concepts that relate to a team's initial question. This follows a similar process to how a facilitator would interact with a domain expert to elicitate a model by asking successive questions to tease out concepts and relationships. Some of the tasks in this sample process require human intervention, for instance to identify relevant sources or set model parameters (e.g., confidence threshold) that prune candidate scenarios. Although a *mix* of human- and automated-components is typical for emerging technologies that build scenarios, our review contributes to precisely identifying where automation is available and which areas require greater research efforts.

3 METHODS

The terms "text mining", "web mining", and "data mining" are frequently interchanged in the literature because typically data in the form of text is mined or scraped from the internet, or web. Consequently, we combined them as search terms. Note that search terms are necessary limited and do not seek to account for all possible synonyms. For example, 'scenario planning' is often used in a given paper together with 'scenario development' or 'scenario building', hence adding search terms would not yield additional articles. We used the following search terms:

("text mining" OR "data mining" OR "web mining" OR "nlp") AND "scenario development"

Previous studies have shown that search engines can cover different parts of the literature, thus a best practice for a comprehensive review is to employ several engines. Consequently, we performed our search in Science Direct (used by the publisher Elsevier), Web of Science Core Collection, Google Scholar, IEEE Xplore, ACM, and Emerald Insight (for its coverage of engineering topics) databases. We looked for works published from 1995 up to the date of data collection, February 2022. We then spent the subsequent year filtering and analyzing the content using the PRISMA process.

The PRISMA process for reviews consists of five consecutive steps (Moher et al. 2009). Our application of these steps is provided on a permanent repository at https://doi.org/10.5281/zenodo.7799300 for full disclosure. First, we tracked articles found from each database. Then, we performed a Title and Abstract Review (TAR) to remove papers that met the search keywords, but did not refer to text mining or NLP in the context of scenarios. Indeed, our study focuses on the synergy between these computational methods of textual analysis and the development of scenarios, hence papers that did not jointly create scenarios with text mining or NLP are excluded. Articles that passed this criterion were further scanned to confirm the presence of text mining or NLP among the methods used in the study. Finally, articles were read in detail to confirm that text mining or NLP were used together with scenario development, rather than co-appearing. Articles that passed all criteria were then examined with respect to our four research questions (RQs).

4 RESULTS

4.1 Articles Found at Each Step of the Search Process

After checking the title and search previews, we examined a total of 420 abstracts from 6 sources: Google Scholar (262), Science Direct (111), Emerald Insight (20), ACM (18), Web of Science Core Collection (5), and IEEE Xplore (4). After removing duplicates and reviewing abstracts to find articles that mentioned text mining in the context of scenarios, we cut the list down to 161 articles. We conducted a high-level review on these articles to find those that mentioned text mining and scenarios as methods, which cut the list down to 41 articles. We then reviewed these documents in detail to understand the type of data and methodologies they used and how they combined automated approaches with scenario planning. Some of the papers were review papers that mentioned text mining or the potential of text mining for scenarios but did not construct a scenario or detail a proposed method. Some of the papers attempted to speed up the scenario process using automation but did not use text-based approaches. Other techniques included using statistical measures from data collected through simulation, storing manually created scenarios in databases for reuse (Kishita et al. 2020), and distributing scenarios created by executives across the rest of an organization (Ramirez et al. 2017). After excluding these papers, *eleven studies were left* that used text-based automation to aid in the creation of scenarios (Figure 3).

The remaining eleven studies were published from 2016 to 2021, with a relatively constant output for each year: two studies in 2016, one in 2017, two in 2018, one in 2019, three in 2020, and two in 2021. Although journals were the primary publication venue (8 out of 11 studies), studies were each published in a different journal at the exception of Technological Forecasting & Social Change, which accounted for two studies. The relatively flat output year-over-year suggests that we are in the early days of automating scenario generation, as these practices have not yet been widely embraced by other research groups. The dispersed nature of the publications (ranging from *Expert Systems with Applications* to the *AAAI* conference or *Futures*) confirms the suggestion that this area is in its infancy, as there is no typical forum to disseminate results. This reinforces the need for a timely review to nurture progress in this area.

4.2 RQ1: Data Collection

For data collection, all the studies except for Retek 2021 used web scraping, the process of using automation to navigate to various web pages, copy text from them, and move it to a data base for further processing. Sources range from sites that have aggregated opinions on the future (e.g. MIT Technology), to twitter feeds that look for specific terms, to journals, patents, and industry reports, to manual subjective inputs. These sources differ greatly in content, length, and structure (Table 1). The four articles that used websites all originated from Seoul National University, which may explain some of their shared characteristics in topics (vehicles), number of sources (all used five websites), and sample size (a few thousand articles). Despite these similarities, studies still varied noticeably in terms of source selection: only one source was shared across studies (MIT Technology Review) and a total of nine different sources were involved.

The four studies relying on news feeds did not fully disclose their sources or the volume of data. Only two studies give an approximate indication for volume: Gokhberg et al. (2020) mentioned "about" 15 million articles, while Park et al. (2016) reported "approximately" 48,000 articles. Only one study provided a list (albeit incomplete) of sources. Indications about sources include "the best of the global news portal" (Gokhberg et al. 2020), using the Watson Discovery service news collection (Feblowitz et al. 2021), using "all the major technology innovation and venture capital news websites" (Kuzminov et al. 2018), or tapping into "29 domestic press agencies" (Park et al. 2016). The lack of information on news

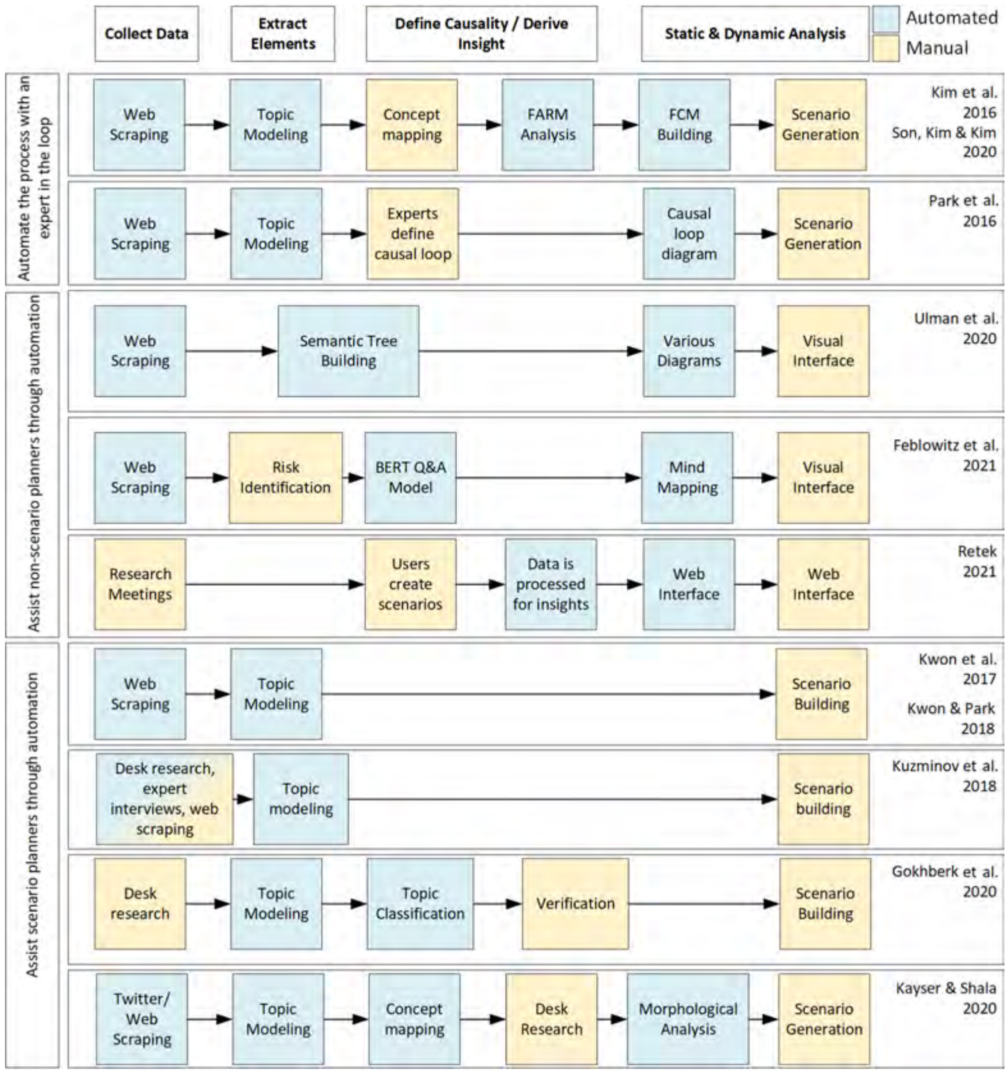

Figure 3: Steps taken in each study and whether automation was involved. Data can be downloaded in table form at https://doi.org/10.5281/zenodo.8067456, including the list of all 11 sources and their URLs.

feeds suggests that such a source should be subject to greater scrutiny and disclosure in order to ensure the *replicability* of past studies (which is part of a broader challenge in the simulation community).

4.3 RQ2: Extracting Individual Elements

In all cases, the data collection produces unstructured text data, consisting of nouns, adjectives, verbs, names, dates, Twitter hashtags, etc. which need to be processed. All studies apart from Feblowitz et al. (2021) use common methods of pre-processing such as stop-word removal which removes the most common words in the English language to avoid focusing on words such as 'and' or 'the', lemmatization which groups inflected forms of the same word, and word vectorization which turns the words to numbered vectors that can be used with Machine Learning. All the studies then use automation to identify constructs of interest to the scenario study, such as scenario driving forces (e.g., economy, technology, social trends), which is done through Topic Modeling. Kayser and Shala (2020) automate concept mapping using commercial software supplemented with LDA, where most other studies that specify their technique use LSA as the primary concept identification automation with the exception of Feblowitz et al. (2021). They also supplement

Table 1: Sources and methods used in each study. A detailed list of websites used as data sources is provided at https://doi.org/10.5281/zenodo.8067430.

		Automation	
Study	Data sources used	Topic modeling	Other
(Kim et al. 2016)	Generic Technology Websites	Yes (LSA)	FARM
(Kwon, Kim, and Park 2017)	Generic Technology Websites	Yes (LSA)	
(Kwon and Park 2018)	Generic Technology Websites	Yes (LSA)	
(Kayser and Shala 2020)	Twitter	Yes (LDA)	
(Gokhberg et al. 2020)	Patents, journals, news feeds, forecast reports	Yes (word2vec)	
(Feblowitz et al. 2021)	News feeds from commercial APIs	No	BERT
(Retek 2021)	User-entered data	Yes	
(Kuzminov et al. 2018)	Research papers, patents, analytical reports, news feeds.	Yes	
(Park et al. 2016)	News feeds	Yes (LSA)	
(Ulman et al. 2020)	EU Documents, scientific journals	Yes	
(Son, Kim, and Kim 2020)	Generic Technology Websites	Yes (LSA)	

topic modelling with morphological analysis (e.g., to remove inflectional endings and convert words to a root form) to explore all potential options to manually create scenarios (Kayser and Shala 2020).

4.4 RQ3: Connecting Elements for Insight

Six studies used a type of graph to connect the elements that were extracted (Kim et al. 2016; Son et al. 2020; Park et al. 2016; Ulman et al. 2020; Feblowitz et al. 2021; Kayser and Shala 2020). These graphs serve to encode both the *structure* of interactions between elements (i.e., which constructs are connected) and the *nature* of these interactions (i.e., how one construct interacts with another). In reality, these interactions can be complex as they be nonlinear (e.g., doubling global warming may cause more than double the impact on polar ice loss), may vary over time (e.g., winter drought is seasonal), involve time delays (e.g., the effects of global warming), or depend on other conditions (Figure 4). Just like models, scenarios are a simplification of reality to assist in specific decision-making tasks. Consequently, each study chooses to focus on only a few potential aspects of the interactions.

The simplest graphical model is a *conceptual model* (Kayser and Shala 2020), which lists associations between constructs (Figure 4-a). Note that the authors used a *variation* of a conceptual model produced by the software `Leximancer`, which adds 'bubbles' to represent automatically inferred themes. Causal models contain more information as they specify the direction of causation (Figure 4-b). The authors of this study also chose to use a variation by having different types of edges and various categories of nodes (Park et al. 2016). The only studies that strictly adhered to a standard graphical representation used a Fuzzy Cognitive Map (FCM), which is a simulation model in which directed edges have values in the range [-1, 1] to indicate the strength of causation (Kim et al. 2016; Son et al. 2020).

Overall, these results suggest that (i) graphical models are the most popular choice when elements are connected, but (ii) these graphical models are very purpose-built hence there is no dominating approach or shared standards beyond the use of nodes and edges.

4.5 RQ4: Scenario Generation

We discern three approaches to generate and provide interactions with scenarios, depending on the target user and their expected involvement. First, several studies use automation to *help a broad audience* to create scenarios (Feblowitz et al. 2021; Ulman et al. 2020; Retek 2021). This category situates scenarios

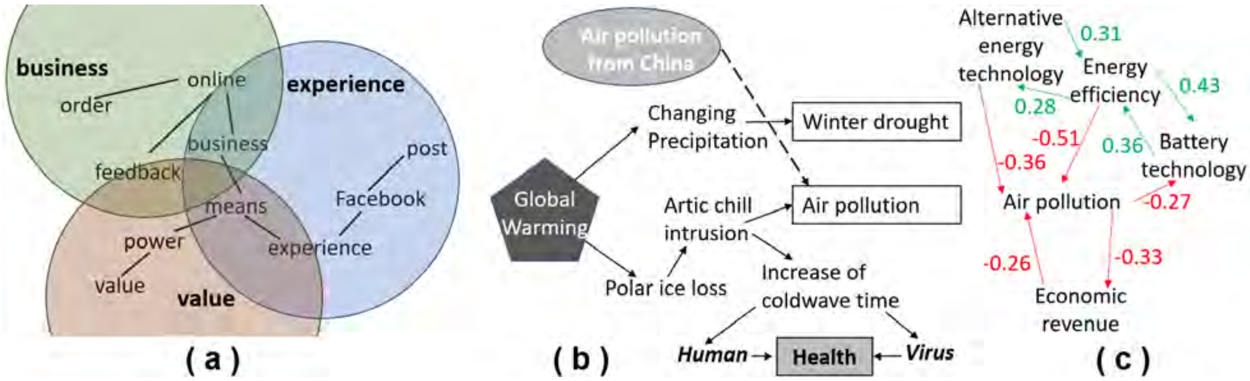

Figure 4: Graphical models used across studies include conceptual models (labeled, undirected graphs) in Kayser and Shala (2020) (a), causal models (labeled, directed, typed graphs) in Park et al. (2016) (b), and Fuzzy Cognitive Maps (labeled, directed, weighted graphs with simulation abilities) in Kim et al. (2016), Son et al. (2020) (c). The models shown here are *excerpts from the original models in each paper*.

within research on citizen science, such that individuals whose lives are affected by a scenario can provide input. Second, there are studies focused on *assisting scenario planners* in creating scenarios (Gokhberg et al. 2020; Kuzminov et al. 2018; Kayser and Shala 2020; Kwon et al. 2017). Instead of diverse community members, this category is geared towards the experts who would previously have engaged in desk research to manually create scenarios and can now automate familiar tasks. The third category emphasizes automation by framing scenario design as an application of data science. Instead of helping experts to speed up some of their tasks, the algorithm is now helped by an *expert in the loop* to determine parameter values (Kim et al. 2016; Kwon and Park 2018; Park et al. 2016; Son et al. 2020). Each group uses text mining for different purposes.

Studies that are intended to help a broad audience create scenarios characterized by a specialized platform that requires user input (Retek 2021) or facilitates the scenario planning process through prompts (Ulman et al. 2020; Feblowitz et al. 2021). These studies use text mining to assist the users of the system and present the results of text mining without the ability to modify the process, keeping the machine learning component inside of a black box. Ulman et al. (2020) designed a system to bring policy makers and rural stakeholders to a common understanding of potential scenarios to create better collaborative policy development. Feblowitz et al. (2021) created a platform where users can define risks they are concerned about to get a better understanding of driving forces associated with them. Retek (2021) outlined an interactive environment that facilitates the scenario process, thus using automation to remove the desk research and need for outside experts. Although these platforms share the goal of bringing scenarios to a broader audience, they achieve it differently. Feblowitz et al. (2021) used commercial APIs to access scraped data, Ulman, Šimek, Masner, Kogut, et al. (2020) used scientific journals and published reports, and Retek (2021) asked users to input data into the system which was then further analysed. We also note a diversity of visualizations in the respective platforms, as Feblowitz et al. (2021) uses mind maps, Ulman et al. (2020) uses a variety of diagrams, and Retek (2021) uses wordclouds among other charts.

The second and most diverse branch of research involves studies intended to **assist scenario planning experts**. These studies encompass automation that complements desk research tasks (Gokhberg et al. 2020; Kuzminov et al. 2018; Kayser and Shala 2020) or outputs data and insights that can be used to create scenarios by experts or in focus groups (Kwon, Kim, and Park 2017). Because text mining is used to assist experts that have access to a white box, the text mining stops short of providing recommendations allowing the experts to make judgements based on findings. All these studies use topic modelling, either via LSA or LDA. Of the five studies in this group, two started the process with desk research to identify the most useful sources of data across news feeds, journals, patents, and forecast reports before processing them (Gokhberg et al. 2020; Kuzminov et al. 2018); Kuzminov, Loginova, and Khabirova (2018) ised text mining during the desk research process. Data sources depend on the study and include web scraping data referenced on twitter (Kayser and Shala 2020), websites that demonstrate thought leadership (Kwon and Park 2018; Kwon et al. 2017), or multiple sources (Gokhberg et al. 2020; Kuzminov et al. 2018).

The third branch of research uses a **domain expert in the loop** in the process of creating quantitative scenarios. These studies turn the scenario process into a data science exercise, where automation is used to find and aggregate data and to create the quantitative model used for scenarios but a person in the middle determines key parameters to the model that affect the outcome. The person in this case would need to be a domain expert as well as a data expert. In these studies experts may interpret the results of web scraping by categorizing them manually (Kim et al. 2016; Son et al. 2020) or define connections between topics found in previous steps (Park et al. 2016). These studies output quantitative models in the form of FCM (Kim et al. 2016; Son et al. 2020) or a causal loop diagram (Park et al. 2016), as discussed in section 4.4.

5 DISCUSSION

Scenario planning has been shown to help people remove blinders, reduce bias, and make better strategic decisions. In two recent reviews, scenario planning was seen as either exclusively manual (e.g., via the traditional approach of expert interviews) or only automatized in two studies (Feng et al. 2022; Lee 2021). Our review on automation for scenario generation found 11 articles, thus identifying a nascent community of practice. We examined this corpus through four research questions (RQs). Overall, all studies leveraged web scrapers to automatically retrieve data, but data sources were highly heterogeneous across the studies (RQ1). Studies relying on news feeds did not systematically disclose their sources or volume, while studies using websites only had one website in common (despite originating from the same research group). Although each study had a clear rationale for choosing their sources, an open question remains: could the choice of data sources be made more consistent based on the type of scenario? A starting point to support replicability and guide choices would be to transparently disclose the criteria to select sources, the list of sources identified, and the volume of data obtained.

Methods for data extraction were more consistent (RQ2), as they use typical pre-processing operations for dealing with text and often involve topic modeling. This part of the process appears strongly aligned with existing practices in the broader domain of modeling technological trends based on textual data. There may be unexplored opportunities in creating unique methods of data extraction in the context of scenario generation, particularly in preparation for the ensuing step of connecting elements for insight (RQ3). This part often involved a graphical model, which (in its simplest form) tracked associations between concepts and could be as sophisticated as a simulation model able to derive quantitative causal effects. Since all methods operated on the same type of dataset (i.e., a textual corpus), it is possible that consolidation of methods will occur in this space such that the method able to derive the most detailed insight will become a de facto standard in future studies. At present, such a method would be Fuzzy Cognitive Mapping.

When it comes to the final step of scenario generation (RQ4), although causal detection is making progress in the machine learning community, defining causality is still a manual task at present. Only one of eleven studies generated a quantitative model that was used to generate scenarios, while all other studies relied on individuals building models once the data had been gathered and processed. Whether automation is used to assist expert scenario planners, the broader community that wants to use scenarios, or data scientists with expertise in a domain, we see that manual interpretation holds a key role to distil results and make the final interpretation from models and text clusters to narrative forms that are typically used in scenario planning. This observation suggests that text mining holds potential for speeding up the scenario process, but creating scenarios ultimately still involves interpretation and manual effort. After all, it is the narrative form of scenarios that speak to us as humans; stories capture our imagination, help us better absorb information, and evoke means to frame potential futures and how to manage them. Given that the main bottleneck in the scenario planning process is the need for experts and lengthy research, computational intelligence may not *completely* automate scenarios but rather contributes to speeding up time consuming aspects such that humans can focus their time where it is most needed.

6 CONCLUSION

Automating the creation of scenarios is an emerging space that can aid expert scenario planners, bring the benefits of scenarios to lay people, or be incorporated into modeling and simulation processes. The current dominant method for the automation of scenario planning is topic modeling of various kinds followed by the generation of a graphical model, which occasionally aligns with the conceptual models used as preliminary artifacts to a simulation model. These practices suggest room for diversification with other Machine Learning techniques to create and integrate scenarios across application domains. These practices also show that simulation models that already align with these practices (e.g., shared data sources or types of conceptual models) may be prime candidates to incorporate scenario automation techniques with minimal disruption. Speeding up the scenario process carries implications for future research. For instance, stakeholder engagement across stages of a simulation process is beneficial to support the translation from research findings to implementation. On the one hand, contenders of automation intended for repetitive low-level tasks have argued that it may come at the expense of stakeholder engagement. While we suggest that humans are still needed to craft compelling stories, this may change by integrating scenario automation with other automation techniques such as ChatGPT. On the other hand, generating scenarios faster implies the ability to create many more scenarios for different groups who do not typically use them. This could lead to scenarios that are refreshed more frequently or even kept up to date in real time. As scenarios are typically used for strategic thinking, studying how they can be integrated into business processes and the repercussions of these integrations constitutes another worthy future investigation.

REFERENCES

Alcamo, J., and T. Henrichs. 2008. *Towards Guidelines for Environmental Scenario Analysis*, Volume 2, 13–35. Elsevier.

Amer, M., T. U. Daim, and A. Jetter. 2013. “A review of scenario planning”. *Futures* 46:23–40.

Bishop, P., A. Hines, and T. Collins. 2007. “The Current State of Scenario Development: an Overview of Techniques”. *foresight*.

Bowman, G., R. B. MacKay, S. Masrani, and P. McKiernan. 2013. “Storytelling and the Scenario Process: Understanding Success and Failure”. *Technological Forecasting and Social Change* 80(4):735–748.

Bradfield, R., J. Derbyshire, and G. Wright. 2016. “The Critical Role of History in Scenario Thinking: Augmenting Causal Analysis Within the Intuitive Logics Sscenario Development Methodology”. *Futures* 77:56–66.

Buckman, S., M. A. de Alarcon, and J. Maigret. 2019. “Tracing Shoreline Flooding: Using Visualization Approaches to Inform Resilience Planning for Small Great Lakes Communities”. *Applied Geography* 113:102097.

Chakraborty, A., and A. McMillan. 2015. “Scenario Planning for Urban Planners: Toward a Practitioner’s Guide”. *Journal of the American Planning Association* 81(1):18–29.

Chermack, T. J., S. A. Lynham, and W. E. Ruona. 2001. “A Review of Scenario Planning Literature”. *Futures Research Quarterly* 17(2):7–32.

Davis, C. W., A. J. Jetter, and P. J. Giabbanelli. 2022. “Automatically Generating Scenarios from a Text Corpus: A Case Study on Electric Vehicles”. *Sustainability* 14(13):7938.

Durance, P., and M. Godet. 2010. “Scenario Building: Uses and Abuses”. *Technol Forecast Soc Change* 77(9):1488–1492.

Feblowitz, M., O. Hassanzadeh, M. Katz, S. Sohrabi, K. Srinivas, and O. Udrea. 2021. “IBM Scenario Planning Advisor: A Neuro-Symbolic ERM Solution”. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Volume 35, 16032–16034.

Feng, L., Q. Wang, J. Wang, and K.-Y. Lin. 2022. “A Review of Technological Forecasting from the Perspective of Complex Systems”. *Entropy* 24(6).

Gokhberg, L., I. Kuzminov, E. Khabirova, and T. Thurner. 2020. “Advanced Text-Mining for Trend Analysis of Russia’s Extractive Industries”. *Futures* 115:102476.

Jang, H., Y. Jeong, and B. Yoon. 2021. “TechWord: Development of a Technology Lexical Database for Structuring Textual Technology Information Based on Natural Language Processing”. *Expert Systems and Applications* 164:114042.

Jang, W., Y. Park, and H. Seol. 2021. “Identifying Emerging Technologies Using Expert Opinions on the Future: A Topic Modeling and Fuzzy Clustering Approach”. *Scientometrics* 126:6505–6532.

Kayser, V., and K. Blind. 2017. “Extending the Knowledge Base of Foresight: The Contribution of Text Mining”. *Technological Forecasting and Social Change* 116:208–215.

Kayser, V., and E. Shala. 2020. “Scenario Development Using Web Mining for Outlining Technology Futures”. *Technological Forecasting and Social Change* 156:120086.

Keller, J., C. Markmann, and A. Heiko. 2015. “Foresight Support Systems to Facilitate Regional Innovations: A Conceptualization Case for a German Logistics Cluster”. *Technological Forecasting and Social Change* 97:15–28.

Kim, J., M. Han, Y. Lee, and Y. Park. 2016. "Futuristic Data-Driven Scenario Building: Incorporating Text Mining and Fuzzy Association Rule Mining Into Fuzzy Cognitive Map". *Expert Systems with Applications* 57:311–323.

Kim, K. H., Y. J. Han, S. Lee, S. W. Cho, and C. Lee. 2019. "Text Mining for Patent Analysis to Forecast Emerging Technologies in Wireless Power Transfer". *Sustainability* 11(22).

Kishita, Y., Y. Mizuno, S. Fukushige, and Y. Umeda. 2020. "Scenario Structuring Methodology for Computer-Aided Scenario Design: An Application to Envisioning Sustainable Futures". *Technological Forecasting and Social Change* 160:120207.

Kuzminov, I., I. V. Loginova, and E. Khabirova. 2018. "Stress Scenario Development: Global Challenges for the Russian Agricultural Sector". *Higher School of Economics Research Paper No. WP BRP* 88.

Kwon, H., J. Kim, and Y. Park. 2017. "Applying LSA Text Mining Technique in Envisioning Social Impacts of Emerging Technologies: The Case of Drone Technology". *Technovation* 60:15–28.

Kwon, H., and Y. Park. 2018. "Proactive Development of Emerging Technology in a Socially Responsible Manner: Data-Driven Problem Solving Process Using Latent Semantic Analysis". *Journal of Engineering and Technology Management* 50:45–60.

Lee, C. 2021. "A Review of Data Analytics in Technological Forecasting". *Technological Forecasting and Social Change* 166:120646.

Lindgren, M., and H. Bandhold. 2009. *Scenario Planning in Practice*, 49–117. London: Palgrave Macmillan UK.

Lu, X., F. Feng, Z. Pang, T. Yang, and Z. O'Neill. 2021. "Extracting Typical Occupancy Schedules From Social Media (TOSSM) and Its Integration With Building Energy Modeling". In *Building Simulation*, Volume 14, 25–41. Springer.

Macanovic, A. 2022. "Text Mining for Social Science – The State and the Future of Computational Text Analysis in Sociology". *Social Science Research* 108:102784.

Moher, D., A. Liberati, J. Tetzlaff, D. G. Altman, and P. Group*. 2009. "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA S tatement". *Annals of Internal Medicine* 151(4):264–269.

Park, S., D. Kim, J. Kim, J. Chung, and J. Lee. 2016. "Future Disaster Scenario Using Big Data: A Case Study of Extreme Cold Wave". *International Journal of Design & Nature and Ecodynamics* 11(3):362–369.

Pillutla, V. S., and P. J. Giabbanelli. 2019. "Iterative Generation of Insight From Text Collections Through Mutually Reinforcing Visualizations and Fuzzy Cognitive Maps". *Applied Soft Computing* 76:459–472.

Raford, N. 2015. "Online Foresight Platforms: Evidence for Their Impact on Scenario Planning & Strategic Foresight". *Technological Forecasting and Social Change* 97:65–76.

Ramirez, R., S. Churchhouse, A. Palermo, and J. Hoffmann. 2017. "Using Scenario Planning to Reshape Strategy". *MIT Sloan Management Review* 58(4).

Retek, M. 2021. "Scenario building in an interactive environment and online communication". *Technological Forecasting and Social Change* 162:120395.

Son, C., J. Kim, and Y. Kim. 2020. "Developing Scenario-Based Technology Roadmap in the Big Data Era: an Utilisation of Fuzzy Cognitive Map and Text Mining Techniques". *Technology Analysis & Strategic Management* 32(3):272–291.

Spaniol, M. J., and N. J. Rowland. 2019. "Defining Scenario". *Futures & Foresight Science* 1(1):e3.

Tiberius, V., C. Siglow, and J. Sendra-García. 2020. "Scenarios in Business and Management: The Current Stock and Research Opportunities". *Journal of Business Research* 121:235–242.

Ulman, M., P. Šimek, J. Masner, P. Kogut et al. 2020. "Towards Future Oriented Collaborative Policy Development for Rural Areas and People". *AGRIS on-line Papers in Economics and Informatics* 12(665-2020-1229):111–124.

Van der Heijden, K. 2005. *Scenarios: the Art of Strategic conversation*. John Wiley & Sons.

Vilkkumaa, E., J. Liesiö, A. Salo, and L. Ilmola-Sheppard. 2018. "Scenario-Based Portfolio Model for Building Robust and Proactive Strategies". *European Journal of Operational Research* 266(1):205–220.

Voinov, A., K. Jenni, S. Gray, N. Kolagani, P. D. Glynn, P. Bommel et al. 2018. "Tools and Methods in Participatory Modeling: Selecting the Right Tool for the Job". *Environmental Modelling & Software* 109:232–255.

AUTHOR BIOGRAPHIES

CHRISTOPHER C. W. DAVIS received a Ph.D. in Engineering & Technology Management from Portland State University. His doctoral thesis was defended in 2022 and focused on scenario acceleration through automated modelling. He is now a Senior Principal Engineering Manager at Microsoft. His email address is davis.christopher.w@gmail.com.

ANTONIE J. JETTER is the Associate Dean for Research at the Maseeh College, and a Professor of Engineering and Technology Management at Portland State University. She works the intersection of management, system modeling, and engineering. She holds a Ph.D. in Technology and Innovation Management from RWTH Technical University. Her email address is ajetter@pdx.edu.

PHILIPPE J. GIABBANELLI received his Ph.D. from Simon Fraser University, and currently is Associate Professor in the Computer Science & Software Engineering Department at Miami University. His research often employs mixed methods, for instance to automate simulation tasks by using machine learning. His email address is giabbapj@miamioh.edu.