Futures 129 (2021) 102733

Contents lists available at ScienceDirect

Futures

journal homepage: www.elsevier.com/locate/futures

Sensemaking and lens-shaping: Identifying citizen contributions to foresight through comparative topic modelling

Aaron B. Rosa[a,*], Niklas Gudowsky[b], Petteri Repo[c]

[a] *Competence Center Foresight, Fraunhofer Institute for Systems and Innovation Research ISI, Karlsruhe, Germany*
[b] *Institute of Technology Assessment, Austrian Academy of Sciences, Vienna, Austria*
[c] *Centre for Consumer Society Research, University of Helsinki, Helsinki, Finland*

ARTICLE INFO

Keywords:
Topic modelling
Comparative foresight
Participation
Natural language processing

ABSTRACT

As foresight activities continue to increase across multiple arenas and types of organizations, the need to develop effective modes of reviewing future-oriented information against long-term goals and policies becomes more pressing. The activities of institutional sensemaking are vital in constructing potential and desired futures, but remain sensitive to organizational culture and ethos, thus raising concerns about whose futures are being constructed. In viewing foresight studies as a critical component in such sensemaking, this research investigates a method of textual analysis that deploys natural language processing algorithms (NLP). In this research, we introduce and apply the methodology of topic modelling for conducting a comparative analysis to explore how citizen-derived foresight differs from other institutional foresight. Finally we present prospects for further employing NLP for strategic foresight and futures studies.

1. Introduction

The European Commission's Strategic Foresight report (European Commission 2020) marks another milestone in the adoption and integration of strategic foresight into governance activities. The continued formalization of foresight as a component of the Commission's activities will increase the amount and scope of future-oriented research efforts and results. This influx of information is likely to outpace the human resources required to account for and accurately regard these different images of the future in time with the demands of policy formulations and decision-makers. Accompanying the institutionalization of strategic foresight, the European Commission has begun approaching the need for an increase in citizen inputs into the missions and goals that will shape EU funding priorities (Mazzucato, 2019). The call for increased citizen engagement may have multiple, interdependent motivations - fostering greater inclusivity, building trust and support for the EU'saspirations, initiating virtuous cycles of reflexive innovation, or increasing futures literacy for example. Whatever the motivations, however, the emphasis on increasing participatory projects and initiatives will certainly lead to an abundance of citizen-based inputs across the spectrum of EU activities. These can be sourced from traditional in-person workshops, digital versions of those workshop formats, configurations of communication platforms and social media, and other types of interactive media (art, games, etc.). We might begin to anticipate the cumulative result being an increase in future-oriented, citizen-sourced statements, ideas, opinions, artifacts, and other types of data. While it will be very interesting to see how citizen-based inputs may change foresight, the sheer volume of the expected inputs create challenges for sensemaking activities within the European Commission or other governing bodies with an eye to the futures.

* Corresponding author.
E-mail addresses: Aaron.Rosa@isi.fraunhofer.de (A.B. Rosa), niklas.gudowsky@oeaw.ac.at (N. Gudowsky), petteri.repo@helsinki.fi (P. Repo).

https://doi.org/10.1016/j.futures.2021.102733
Received 20 November 2019; Received in revised form 22 February 2021; Accepted 19 March 2021
Available online 23 March 2021

Sensemaking is an important endeavor for any organizations that "seek plausibly to understand ambiguous, equivocal or confusing issues or events" (Brown, Colville, & Pye, 2015). Sensemaking is a series of formal and informal activities by which new data is contextualized within the culture, identity, strategy, and operations of said organization in order to inform decisions both in the present and in the near-term futures (Weick, 1995). With an increasing frequency and variety of incoming, citizen-sourced signals about possible futures, coupled with the need for organizations to rapidly 'make sense of' and respond to these information streams, the temptation to utilize automated processes in analysis of these data pools will undoubtedly rise (Boysen, 2020). Indeed, it was this very temptation that animated this research task at its inception. Having conducted a comparative reading analysis of foresight documents from both experts and citizens, and drawn some conclusions regarding differences that emerge between these two sources (Rosa, Gudowsky, & Warnke, 2018), we were curious to discover if the reading process we had engaged could be automated using innovations from the rapidly evolving field of natural language processing (NLP). To be clear, these automated analytical processes do not provide organizational contexts - they rather provide results for further sensemaking to be read through the lens of the organization.

In this article, we explore how outcomes from citizen-based foresight differ from foresight based on expertise. We accomplish this by examining topics that emerge in foresight reports and doing so we also develop and pilot a process for achieving this. As a key methodology, we apply topic modelling (Blei, Ng, & Jordan, 2003), which is a linguistic methodology suitable to the analysis of large volumes of texts. In the upcoming section, we first discuss how computational processing of natural languages can assist in future-oriented sensemaking. Then we proceed to present the methodology of topic modelling and how it is used to compare foresight reports from both expert and citizen sources. Finally, we discuss observed similarities and differences, and reflect on how institutional sensemaking is apt to change when foresight activities become more inclusive and encompassing, and computational processes are introduced to analyze outcomes.

2. Computational text analysis for future-oriented sensemaking

According to the European Forum for Future Looking Activities, sensemaking is in need of improved methods, when compared to the other future-oriented activities of gathering strategic intelligence, selecting priorities, and achieving implementation (EFFLA, 2015). Traditionally, sensemaking plays a central role in the creation of meaning, and consequently in the formation of collective identity that informs organizational action, but this is often an event-focused process done in hindsight (Weick, 1995; Weick, Sutcliffe, & Obstfeld, 2005). While retrospective sensemaking is common in organizations, often serving as the primary conduit between existing hindsight-foresight organizational mechanisms (Nathan, 2004), prospective sensemaking has often been carried out in limited scale by scholars and formal organizations (Sandberg & Tsoukas, 2015) though it remains essential to both public and private policy-makers engaged in future-oriented processes. Sensemaking, particularly with regard to possible futures, is an interpretative-constructive act of collective knowledge creation, building on individual expertise and experience (Könnölä, Salo, Cagnin, Carabias, & Vilkkumaa, 2012; Rohrbeck, Battistella, & Huizingh, 2015), and therefore renders the process somewhere between art and science (Jouvenel, 2017; Loveridge, 2008). According to Mills, Thurlow, and Mills (2010), plausibility plays an important role in the distribution of power within sensemaking activities within organization, particularly in situations where it is tied to elements of organizational culture and identity (gender roles, discriminatory behavior, etc.)(Mills 2010). When policy-making discussions are shaped by the language and culture that dominate an organization, then vectors must be found to infuse prospective policy-making with individuals' input that can be accepted as plausible within those organizations' filtering parameters.

As policy decisions continue the long-turn towards more participatory methods of information gathering (Wagner, Vogt, & Kabst, 2016) and agenda setting (Mazzucato, 2019), the rapid expansion of future-oriented inputs to decision-making continues apace (Metzger et al., 2018; Sotoudeh & Gudowsky, 2018; VOICES, 2014). The value of citizen inputs into such policy processes continues to increase (Wiener, Gattringer, & Strehl, 2018), and we expect influxes of citizen-generated, future-oriented data points to become essential components of policy decisions. However, given the growing number and heterogeneous nature of these data points (sourced from in-person workshops, online forums, surveys, and other media platforms), we can expect sensemaking processes for policy to become increasingly complex and time consuming. Additionally, despite policy-making bodies best efforts, the diversity of cultural, linguistic, and situational conditions that are present in the EU citizenry are difficult to recreate as an organizational culture and ethos. Thus components of sensemaking activities must be developed that can balance institutional pressures and forces with citizen perspectives, without sacrificing the richness of the EU's collective imagination. This means creating a 'public voice' from various sources and types of citizen-provided, future-oriented data. Such a 'voice' can assume the "power to speak" in sensemaking activities (Mills 2010) - offering collective ideas adapted to the language, culture, and identity that define the context of EU governance and future-oriented policy-making. Developments in computational and textual analysis might be useful in creating a type of collective 'voice' from larger data sets, and could make inclusive, prospective sensemaking more achievable and economical to carry out than before.

The utilization of Natural Language Processing algorithms (NLP) for textual analysis has a rich history (Chowdhury, 2003; Lewis & Jones, 1996), with increasing potential uses as the cost of computing capacities has gradually diminished (Coenen, 2011). While the application of computer-aided tools for exploring quantitative data has been at the very center of foresight and forecasting activities since the beginning of the academic field, applying such tools to qualitative data, especially in the sensemaking phase, is underexplored partly because sensemaking in foresight often takes place in dialogic participatory form. Nevertheless, computational text analysis, for instance by natural language processing or text mining, has for some time been seen to bear promises in the field of foresight (Gharehchopogh et al., 2011). Apreda, Bonaccorsi, dell'Orletta, and Fantoni (2016) describe how natural language processing tailored to patent databases can be harnessed to identify emerging technologies. Kayser and Blind (2017) describe the potential for introducing new foresight approaches by using text mining methods on an expanded knowledge base that includes social media, news and patent data. Additionally, they analyze how foresight reports are commonly used as sources for foresight exercises, and that text mining of foresight reports might be applied to generate input into a scenario process (Kayser & Shala, 2016). Here, Boysen (2020) reviews how “Big Data and improved analytics capabilities can expand the knowledge base and act as a corrective to our cognitive biases” in foresight.

The purpose of our research is to develop a mode of (or approach for) comparative analysis that could be applied to policy-oriented foresight reports, thereby reducing the resource requirements for future endeavors and serving as a parallel line of investigation for checking, supporting, and thus improving results from qualitative analysis within forward looking activities. Aligned with these efforts, this research builds on comparative foresight analysis conducted within a large-scale participatory agenda setting activity, namely CIMULACT (Rosa, Gudowsky, & Warnke, 2018, 2018b). Here we focus on developing and applying a computational procedure to augment future comparative processes. Given the results of previous NLP research conducted on citizen-derived foresight – namely the analysis of the 179 citizen visions that emerged from the initial phases of the European-wide CIMULACT project (Repo, Matschoss, & Timonen, 2017) – we anticipate that further development of an application of such processes can contribute to creating efficient and analytically founded modes of comparison.

Here, we see particular potential for utilizing text mining to compare foresight outputs with natural language processing, and specifically topic modelling. In applying topic modelling to a representative corpus of recently published foresight reports, we seek to test the boundaries and limitations of deploying such technologies with respect to comparative foresight research. Just as some NLP technologies have offered crosscutting insights into oft-promoted content within possible and plausible futures scenarios (Kayser & Shala, 2016), we believe NLP can also make institutional biases and undeclared research restraints more visible. Finally, we expect to find modes for improving the sensemaking phase of long-term planning projects, especially those that produce rather large quantities of text, for instance in participatory foresight.

In particular, we hypothesize that NLP can be used as a critical mode of sensemaking that addresses the epistemological gap that might exist between experts, policy-makers, and the citizenry. In doing so, the further development of such techniques could alter the selection of future priorities, and thereby shape the implementation of future-oriented concerns and aspirations as sourced through participatory processes. Here we furthermore hypothesize that such a lens may allow for a streamlined reading of future policy recommendations, if not policies themselves. The utilization of NLP could facilitate the rapid generation of topic models from both citizen-sourced forward-looking activities and expert based research. In turn, this could allow for a more robust framework for policy drafting that ensures equal representation in the field, and recognize both the specific recommendations of experts and the more human-centric concerns generated in citizen-led, participatory efforts.

This article investigates how the contributions from a citizen-derived foresight process differs from those of a large and varied set of other foresight conducted by practitioners and institutions of the field. We examine the differences in terms of topics modelled with natural language processing techniques. To accomplish this, we introduce CIMULACT, a large-scale, participatory, forward-looking activity (FLA) with a focus on comparative foresight work conducted therein. While a full review of this project’s results goes beyond the scope of this contribution, a general overview will set the context and interest for this article and establish the mode of comparison that we believe the proposed NLP process can compliment.

The European CIMULACT project (2015–2018) was designed as a multi-actor, forward looking activity, aimed at contributing to responsible research and innovation agenda setting in the European Union (CIMULACT, 2019). Here, research and innovation agendas served as early entry points for societal needs (Gudowsky & Peissl, 2016). Through a mixed-method process of visioning, workshops, and consultations, the CIMULACT project contributed to the European Union’s research and innovation agenda with recommendations, which were based on citizen-derived conceptions of a desirable and sustainable future (Mission Publiques, 2017).

The project’s methodology was designed to address each of the four categorical necessities of forward looking activities as defined by the EFFLA, with a particular emphasis on developing ’Sensemaking’ capacities and linkages to ’Priority Selection.’ Rosa et al. (2018) describe the comparative process and provide a thorough analysis of the obtained results in light of the project mission. While this extensive work was distributed across the consortium, and produced significant results[1], the labour-intensity of the analysis explains why such comparative foresight has remained a rarity. The efficiency opportunities attributed to computational analysis indeed warrant attention in such foresight studies.

While a human-led textual analysis method of comparing foresight reports is certainly achievable (Rosa et al., 2018b), it requires large resource expenditures and strict qualitative procedures to ensure the reliability of the process of such a large set of textual data. Such an approach can quickly reach its limits of efficacy when increasing the number of reports included in the comparison. The following section describes how research carried out with natural language processing methodology complements and corresponds to previous textual analysis of the same corpus. Topic modelling, in particular, can provide a rigorous and systematic methodology for the comparison of the contents of foresight reports (Repo & Matschoss, 2018).

[1] Findings concluded that, firstly, ‘citizen-based, multi-actor co-created policy advice qualitatively differs considerably from that elicited by expert-based reports, in terms of direction and focus of the proposed R&I agenda (Rosa et al., 2018a). Secondly, that ‘while professional foresight research institutions are more detailed in their examination of specificities concerning future-relevant factors, trends, and emerging issues, distinct qualitative differences emerge within citizen-based foresight initiatives with regard to contextualizing the possible impacts of future research and innovation projects (Gudowsky & Rosa, 2019).’

3. Methodology of topic modelling

Topic modelling, as a subset of natural language processing (NLP), is based on machine learning and textual analysis. Collocations of words are observed in topic modelling, and these collocations contribute to meanings, i.e. topics, which are used to make sense of the corpus under analysis. In essence, topic modelling is a probabilistic and statistical procedure for natural language processing that enables the discovery of latent topics in extensive sets of texts.

For this research, the freely available MALLET machine learning toolkit (McCallum, 2002) was used to apply Latent Dirichlet Allocation (LDA) on the corpus consisting of the 16 foresight reports. LDA is a generative, probabilistic, Bayesian model for the analysis of texts (Blei et al., 2003), which produces results that are relative to the corpus and its modelling. For this research, a hyperparameter optimization every ten iterations was applied to consider the relative weights of the topics in the corpus.

The body texts of the foresight reports formed the corpus of the study. Their texts were prepared for analysis according to standard NLP practices in order to focus on their content through the deletion of stop words, capital letters, punctuations and repetitive sections such as headlines, as well as formatting such as page breaks[2] . Modelling was piloted with 12, 15, 20, 30, and 50 topics to obtain an overall view on how the corpus could be analyzed. Results from modelling with different numbers of topics proved consistent, and indicated that the quality of the data suited the methodology. Topic modelling was then conducted with 20 topics as this number proved to embrace the large variety of topics in the corpus while making them distinct from each other and thereby communicable. This rather large number of examined topics also allows for an elaborate analysis of how the topics are distributed across the 16 foresight reports, thereby providing a comparative take on how differently foresight reports approach the future.

The particular advantage of topic modelling, as compared to forms of qualitative analysis, is that it allows for a technical identification of topics, and can provide indicators on how these topics are distributed across each examined foresight report. This latter feature of topical distribution serves as the base on which contributions from citizen-derived foresight are compared to that of the other types of foresight. Topic modelling indeed fits comparative research designs because it makes it possible to observe topical distributions across varying origins of data (Matschoss, Repo, & Timonen, 2019).

The topic 'title' is labelled by the research team, in accordance to key words and with consideration of other modelling metrics describing word exclusivities, frequencies and probabilities. Generally, the more familiar researchers are with the content of the corpus, the easier interpretation of the topics becomes. However, the tool is also often successfully applied to model topics in very large corpi, consisting for instance of several hundred or thousands of scientific articles. This is a particular strength of the tool, to be applied when processing quantities of text that a research team could not read within a reasonable time frame. Nevertheless, good knowledge of the assessed field is necessary for making sense of the produced topics, which is crucial for choosing parameters of the model, for instance the number of topics to be modelled.

The identified topics were then described in accordance to the key words, and arranged in overarching clusters of topics for readers' convenience. We have taken care to name each of these clusters in a manner that best defines their overall content, and attempted to avoid introducing normative bias. In the upcoming sections, we describe the 20 identified topics and examine the distribution of these topics in the 16 examined foresight reports.

4. Results

The results generated from the method of topic modelling provide us with the appropriate evidence to advance our investigation into the benefit of machine learning textual analysis for sensemaking. This article section comprises (a) identified topics by means of topic modelling, (b) clusters of topics and their description, and (c) the distribution of topics across the investigated foresight reports.

4.1. Topic modelling and clusters in foresight reports

The identification of common topics in the examined foresight reports represents the first step in preparing their topical comparison. The 20 topics identified in the reports by means of topic modelling are presented in Table 1 and described in the upcoming sections. The modelling provides the weight of the topic in terms of the Dirichlet parameter and a list of collocated key words. Four topics with a Dirichlet parameter of a greater magnitude than those of the others emerge in the modelling: Global economic change, Food and energy systems transformation, Public services using data, and Key technologies & market development.

The 20 topics presented in Table 1 are described in overarching topic clusters that connect to these large topics. Our choice to describe topic clusters from the results of the topic modelling methodology was made in an effort to make the results more clearly communicable to a broader audience. Our familiarity with each individual 'expert-based' report, nevertheless, allows us to examine in more detail how the results help create a more neutral comparative lens for sensemaking processes. From the 20 topics resulting from the text analysis, our team derived topic clusters to more succinctly communicate the broader content categories in which the foresight report data could be understood. Accordingly, the topics are described according to the following overarching clusters, which themselves constitute the synthesized contribution of the examined foresight reports: Economic development, Technology and markets, Technological innovation, Governance and public services, Challenge-centered research, and Values and rights. The topic cluster around technological innovation is distinct from the cluster examining the relationship between technology and markets. Further, the

[2] The two CRF reports in the corpus (CRF 2012a; CRF 2012b) were merged into one as the second was considered a follow-up to the first.

Table 1
20 modelled topics in 16 foresight reports.

Dirichlet parameter	Topic title	Key words
1,519	Global economic change	global, world, change, economic, increase, major, growth, time, years, population, economy, emerging, production, climate, future, continue, financial, levels, technological, markets,
1,474	Food and energy systems transformation	development, food, research, system, education, models, order, energy, knowledge, sustainable, local, make, part, processes, social, develop, activities, level, quality, process,
1,425	Public services using data	social, systems, citizens, society, policy, data, life, research, issues, innovation, services, public, health, based, technology, information, governance, societal, individual, security,
1,380	Key technologies and market development	technologies, future, role, areas, public, market, open, key, technology, big, provide, international, sector, potential, support, impact, developments, area, strong, challenges,
0,725	Developing new materials	development, information, materials, technology, high, energy, related, human, devices, based, production, data, time, material, research, scientific, applications, brain, science, environment,
0,663	National intelligence for governance	people, billion, countries, year, political, intelligence, increased, national, government, internet, developing, expected, increasing, number, improve, report, total, state, due, trade,
0,479	Bioeconomy for solving Grand Challenges	energy, innovation, health, change, space, data, climate, environmental, big, key, europe, emerging, bio, food, science, migration, horizon, sustainability, transport, global,
0,448	National economic growth	growth, countries, gdp, trade, productivity, china, economies, europe, energy, gas, higher, investment, rate, capital, labour, rates, population, increase, today, economy,
0,437	Multi-level EU development	european, europe, development, regions, cities, urban, policies, policy, economic, regional, areas, territorial, transport, term, countries, energy, long, cohesion, local, public,
0,325	ICT driven waste management	data, energy, applications, space, european, technologies, products, waste, food, systems, smart, services, production, cloud, cities, ict, security, market, synthetic, industrial,
0,324	Applications for trending technologies	graphene, technology, printing, drones, technologies, vehicles, storage, smart, energy, legislation, electricity, home, impact, terms, level, significant, electronic, access, users, required,
0,308	Challenge driven research	research, innovation, transport, european, climate, energy, challenges, horizon, important, societal, europe, researchers, change, challenge, resource, knowledge, industry, sciences, projects, including,
0,243	Economies of demographic change	economic, population, drivers, horizon, space, trust, technologies, natural, change, europe, technology, migration, ageing, healthcare, actors, degradation, environmental, food, rise, political,
0,192	Energy production and distribution technologies	energy, technologies, technology, cities, printing, urban, percent, gas, shale, states, united, smart, oil, revolution, grid, robots, systems, cost, building, water,
0,164	EU in a globalized world	european, union, states, united, countries, europe, political, china, security, market, social, development, member, inequalities, system, russia, defence, revolution, middle, rise,
0,154	Comparing national digital developments	research, countries, expected, oecd, data, growing, developing, storage, increasingly, manufacturing, systems, iot, number, digital, technologies, satellites, applications, nanomaterials, economies, million,
0,133	Global population and resource scarcity	world, million, water, global, women, africa, billion, food, countries, region, international, america, future, population, china, systems, asia, years, today, crime,
0,064	Societal values and human rights	values, human, sti, rights, weak, people, controversies, care, science, european, europe, change, signal, freedom, signals, related, data, media, living, public,
0,039	Competitive leadership through innovation	market, areas, recent, impacts, technological, product, vanguard, ingress, invention, chains, finland, radical, based, application, growth, finnish, research, model, important, developed,
0,020	Advanced ICT research	data, research, digital, trend, themes, internet, robotics, innovation, european, iot, economic, learning, theme, europe, projects, science, robots, singularity, policy, things,

topic of Societal values and human rights appears as an outlier and is, therefore, described separately.

4.1.1. Economic development

This cluster of topics centres on issues of economic development, mainly of the European Union but also in context of a globalized economy. *Global economic change* is the topic with the most weight in the corpus (Dirichlet parameter of 1.519) and is distributed across the sixteen examined foresight reports. This topic describes global change mechanisms mainly from an economic perspective, with high exclusivity metrics for 'financial' and 'markets'. *EU in a globalized world,* in turn, is mostly a topic unique to ESPAS (2015) and positions strands of development of the European Union within the context of global players such as the USA, Russia and China. *National economic growth, by contrast,* summarizes and compares various traditional parameters of countries' economic growth, such as productivity and labour rates or GDP and is mainly apparent in Gros, Alcidi, and Behrens (2013).

4.1.2. Technology and markets

The three topics in this cluster describe the role of innovation in creating, developing and maintaining markets. *Key technologies and market development* delineates the role of future key technologies in creating international markets. With a high Dirichlet parameter (1.380), this topic is distributed evenly across all foresight reports. *Competitive leadership through innovation* is uniquely represented in Kuusi and Vasamo (2015) and describes the potential market impact of emerging technologies and products. *Comparing national digital developments* focuses on the comparison of data on OECD countries' research and development performance, mainly driven by digitalisation, and is mostly defined by the DASTI report (DASTI, 2016).

4.1.3. Technological innovation

This cluster summarizes four topics that describe streams of cutting-edge technological development. *ICT driven waste management* focuses on the development of data-driven technologies, applications and services for several sectors such as energy, waste management or production. It is almost uniquely represented in the report by the Foresight Network (2014). *Developing new materials* describes the potential role of new materials for developing new technologies, applications, devices and products. It is a rather broadly distributed topic with strong representation in reports by ESPAS (2015); Kuusi and Vasamo (2015); Warnke, Schirrmeister, and Rosa (2017). *Applications for trending technologies* encompasses mainstream trending technologies such as drones and materials such as graphene, and discusses their application in several sectors, but mainly for energy. This topic is mainly represented in van Woensel, Archer, Panades-Estruch, Vrscaj, and Parliament (2015). *Advanced ICT research* is almost exclusively formed by (Burgelman, Chloupková, & Wobbe, 2014), and describes research on advanced matters of digitalisation.

4.1.4. Governance and public services

Two topics in this cluster describe prospects of data-driven governance - meaning the utilization of large scale data collection and analysis as critical inputs for public policy decision making. *Public services using data* is a highly citizen- and society-centred topic distributed across most reports, which is reflected in a high value of the Dirichlet parameter (1.425). It describes how research and innovation in general, and data analysis in particular, impacts social systems, public policy and the life of citizens. *National intelligence for governance* describes the use of large-scale data collection efforts (census data, economic data, environmental data, etc.) as a mode of informing security measures, economic policy, and international environmental cooperation, among other governing decisions. *Multi-level EU development* describes regional development within the EU, focussing on urban issues, transport and cohesion, with a particular emphasis on the utilization of state and regional data collection and analysis. While this topic is mostly defined by Ulied and ESPON (2014) it is also evident in most other reports.

4.1.5. Challenge-centered research

The largest cluster summarizes six topics, which are all characterized by framing research and innovation endeavours along challenge driven issues. *Food and energy systems transformation* is a highly distributed topic with the second highest Dirichlet parameter (1.474). It describes systems transformation towards sustainability at various levels. *Challenge driven research* largely describes the societal challenges as put forward by the European Union, and how these might be addressed by science and technology. It is mostly defined by the report by Duckworth, Lye, Ravetz, and Ringland (2016) and Warnke et al. (2017), while being only partially represented within most of the other reports. *Energy production and distribution technologies* describe technological advances in (traditional, carbon-based) energy resource exploitation. Issues relevant to this topic can mainly be found in the report by Strategic Foresight Initiative (2013). Innovative transport research is mostly found in the reports by the Copenhagen Research Forum (2012) and explores how research and innovation especially in the transport sector can help solve societal challenges such as climate change. *Economies of demographic change* explores drivers for demographic change, technological solutions for ageing societies and their relation to the economy. It is almost exclusively defined by the report by Rousselet (2014). *Global population and resource scarcity* describes global challenges related to water and food systems and exclusively mentions women's roles. This topic can almost uniquely be found in the State of the Future report (Glenn, Gordon, Florescu, Project, & Millennium, 2014), one of the few reports investigated with a scope broader than that of the EU.

Table 2
Comparison of modelled topics with qualitative alignment and degree of coverage (Rosa et al., 2018b).

Modelled topics	Comparative reading: high alignment topics.	Comparative reading: high degree of coverage topics.
Food and energy systems transformation	• Smart energy governance • Making dense and growing urban areas more sustainable and liveable • Responsible use of land	• Smart energy governance • Making dense and growing urban areas more sustainable and liveable • Good quality food for all
Public services using data	• Data for all- share the power of data • Making dense and growing urban areas more sustainable and liveable • Personal and organizational choice management	• Data for all - share the power of data • Making dense and growing urban areas more sustainable and liveable • Dissemination and continuous exploitation of research and innovation in the healthcare system • Personal and organizational choice management
Multi-level EU development	• Transforming technologies for planet and people • Production awareness • Snakes and ladders - scales of issues and actors	• Transforming technologies for planet and people • Production awareness • Community building development
Challenge driven research	• Transforming technologies for planet and people • Evidence-based personalized health care • Meaningful research for community	• Transforming technologies for planet and people • Dissemination and continuous exploitation of research and innovation in the healthcare system • Evidence-based personalized health care • Meaningful research for community
Outliers	• Empowering diversity in community • I am empowered to lead my changes	• Rethinking (the new) job market needs

4.1.6. Values and rights

Societal values and human rights is almost exclusively represented in the report by Remotti et al. (2016). It describes a European perspective on interrelations of society-centred values and human rights with science, technology and innovation. It stands apart from other topics in the modelling, with the clustering of distinctly human perspectives regarding social changes and implying cultural and ethical shifts in the social landscape.

4.2. Distribution of topics across foresight reports

The second step in the topical comparison of the foresight reports involves a more detailed review of how the identified topics are distributed across the examined foresight reports. Table 2 in annex shows this topical distribution in terms of percentages which are provided by the modelling procedure. It can be observed that the citizen induced foresight report by the CIMULACT project (Mission Publiques, 2017) focuses very much on the topic of *Food and energy systems transformation* (47.7 %), significantly on *Public services using data* (27.6 %), to a lesser degree to *Multi-level EU development* (6.2 %), *Key technologies and development* (5.4 %), and *Challenge driven research* (4.0 %). Ten other topics are also (0.2–2.6 %) considered in the project report, indicating that citizen foresight touches a wide range of topics.

The top five topics that saw the highest aggregate distribution across the other foresight reports are: 1) *Global economic change* (12.6 % topical share), 2) *Public services using data* (10.0 %), 3) *Key technologies and market development* (9.4 %), 4) *Food and energy system transformation* (8.3 %), and 5) *Developing new materials* (7.7 %). Within each of these five topics, the CIMULACT report had the highest focus of content within *Food and energy system transformation*, and *Public services using data*, and showed a slightly below average focus of content in the remaining three topics. The foresight reports with the leading focus of content with respect to the remaining 3 topic areas were: the European Strategy and Policy Analysis System (ESPAS) (*Global economic change*), RAND Corporation EU (*Key technologies and market development*), and the OBSERVE project consortium (*Developing new materials*).

Also of note are topics that receive limited focus in the CIMULACT results, but were mentioned in a variety of other foresight reports. While CIMULACT was moderately inclusive of the topic content, three topics standout as absent from the citizen-based foresight project. Firstly, the topic *Economies of demographic change* was addressed in some manner by 11 of the 16 reports, yet is quite neglected in CIMULACT. Secondly, the topic titled *EU in a globalized world* was mentioned by 9 of the 16 reports, but was absent within CIMULACT. Finally, the topic titled *Energy production and distribution technologies* was indicated within 10 of the 16 reports, but not within CIMULACT.

5. Comparative analysis

The topic modelling results provide insights into the relationship between the sources of forward-looking reports and their focus. The CIMULACT contribution to foresight is of particular interest because it is the only report in our corpus that originates from citizens, and therefore serves as the focal point for the analysis of this paper.

5.1. Key differences between citizen- and expert-originated foresight

In direct comparison, the modelling results of CIMULACT were reviewed against those from the expert-based report corpus. It has been observed previously that experts seem to accentuate techno-economical opportunities while citizens appreciate social concerns (Matschoss et al., 2019), and our analysis confirms this. This is an important result because it makes note of how perspectives on the future are constructed and thereby constrained by the observer's framing and capabilities (Ahlqvist & Uotila, 2020). When considering experts foresight results as policy advice, the reported cognitive biases need to be taken in consideration (Winkler & Moser, 2016): e.g. "framing and anchoring as well as the desirability bias, […], the bandwagon effect and belief perseverance. Applying topic modelling to such contributions then reproduces these biases, which may be latent, and makes them comparable. While citizen-based foresight is likely also prone to cognitive bias, these biases may be qualitatively different to such an extent that results can act as a form of benchmark for expert based foresight.

Eight topics emerge as telling examples of the differences in content that citizen and expert foresight produce. Four topics were accentuated with manifold topical shares in citizen originating CIMULACT foresight (*Food and energy systems transformation, Public services using data, Multi-level EU development, and Challenge driven research*), whereas four topics were accentuated in the expertise based reports (*National intelligence for governance, Developing new materials, National economic growth, and Global economic change*). The table in Annex 3 presents the distribution of these topics across the examined reports, with outliers being excluded from the average figures and marked with asterisks (*).

The CIMULACT foresight products accentuate three topics very clearly in the modelling results: *Food and energy systems transformation* (47,7 % share in the topical distribution vs. 8,3% in the other reports), *Public services using data* (27,6 % vs. 10,0 %), Multi-level EU development (6,2 % vs. 1,4 %), and *Challenge driven research* (4,0% vs. 0,9%). It is notable that the first two cover approximately 75 % of the topical distribution of the CIMULACT foresight, with the former close linked to the idea of sustainability transition (Köhler et al., 2019; Raven et al., 2019), and the second to the creation of public good in an information economy (Drahos & Braithwaite, 2017). In effect, this result confirms the need to carry out foresight in which citizens can determine agendas.

Overall, these topics are traceable to concepts of local and regional identity, daily life, and societal functions and cohesion. In short, these topics focus on how people organize, decide, and act in ways that are seen to be beneficial to the group in terms of health and well-being, transport, and regional development of infrastructure and services. While economic growth is also viewed as important, seen in data that we will discuss below, these topics appear to focus on issues that form relationships between peoples and places, but are not strictly mediated by monetary economics. Interesting here is that CIMULACT appears to endorse the use of data, via citizen science and local innovation, for the extension of community benefits, and prefers challenge-driven research, which makes its target more approachable.

By contrast, the expert-based report corpus also situates the use of data as among the priorities, but within a different scope of intended benefits. The relating topic from the expert report corpus, *National intelligence for governance* (3,3% average topical share in the expert reports vs. 0.8 % in CIMULACT), promotes the utilization of data gathering and analysis towards ends that would bolster national security institutions (Glenn et al., 2014), rather than directly benefiting the citizens through service provision. In our view, this is a critical distinction between citizen and expert report texts, as it outlines differences in who is conducting, and benefitting from, research and development efforts: the citizenry or the state?

The citizen originating CIMULACT report positions the economy to play a minor role in the agenda-setting of the future. Indeed, *Global economic change* receives only a 2,6 % topical share (vs. 12,6 % in expert foresight) and *National economic growth* even less (0,2% vs. 2,2 %). We believe that the differences accrue due to competing understandings of concepts of economy. Expert reports focus on economic development that considers growth in GDP, competitiveness, restructuring, and unemployment rates. Again, this confirms the previously observed tendency of experts to accentuate issues relating to economic opportunities and performances. While these matter to citizens, their focus is more on what economic development should offer, for instance in terms of food, energy and public services. These results align well with previous research regarding differences between expert and citizen knowledge bases, and the production of novel research items (Gudowsky & Rosa, 2019). Citizens are more engaged with innovation that leverages local knowledge for the improvement of their community specific needs, and discuss it in these terms.

Meanwhile, experts' macro scale perspective on innovation systems, and their impacts on markets and national economies, reflect different interests. This shows that experts often think or recommend within the somewhat strict boundaries of the prevailing systems' needs. Even when recommending systemic transitions or transformation, experts' perspectives range within barriers that ensure a high connectivity to the status quo, which may make such recommendations more accessible and applicable for policy-making. Furthermore, these results show that a division persists between sensemaking activities within the intellectual community informing the organization, and the aggregated results of diverse contributions presented in the topic-modelling.

Similarly, the topic of *Developing new materials* receives much more interest in the expert reports (7.7 % share vs. 0.7 % share in CIMULACT), highlighting that experts seem more inclined to consider intermediary steps, path dependencies and practical challenges when aspiring futures. Citizen originated foresight, on the contrary, focuses more on the desirability of futures than on the ways to reach these futures (Lee, 2015). These differences do not necessarily need to result in target conflicts between citizen and expert inputs, but rather point towards which views are applicable to which phase of innovation processes. Citizen-based foresight is well suited for agenda setting at a very early stage in the innovation process, providing socially acceptable or desirable targets for orienting innovation responsibly. Expert knowledge is however necessary for establishing the specificity of transition pathways. However, target conflicts may certainly appear when consulting different actor groups for participatory innovation governance, and those need to be made transparent.

Finally, it should be noted that the CIMULACT project report considered 15 of the 20 total topics, indicating that the overall content of CIMULACT was broadly relevant within the context of its peer foresight publications. This is an important finding as a counter argument against those dismissive of participatory activities as non-scientific, or somehow uninformed (Tseng, 2018; Wagner et al., 2016; Yang & Callahan, 2007). These results demonstrate that, on the contrary, the CIMULACT proceedings were broadly informed of, and provided unique perspectives on, a variety of future-oriented inputs - more so than over half of the other reports. This points toward a process of inductive logic at work within diffuse, citizen foresight, if analyzed in aggregate form, implying that collective citizen inputs are both valid and valuable in any context in which they might be used for prospective sensemaking.

As noted above, in two of the most frequently highlighted topics, CIMULACT had the highest focus of content in both. This finding further strengthens the positioning of the CIMULACT results, as not only broadly informed, but also focused on content fields that were seen as primary concerns by peer-level institutional foresight. Linking participatory foresight processes to future-oriented pursuits, for instance technological assessment studies, remains a challenge, perhaps best addressed by mixed methodological approaches (Kaivo-oja, 2017). Whereas topic modelling may not be a suitable approach for the identification of weak signals - given these often emerge from niches that are erased in the aggregate analysis - the method excels at unveiling prevalent topics of contemporary interest. Furthermore, topic modelling as a mode of comparison can help better differentiate along the issues that either group favors more than the other. This type of topic differentiation could present a new angle in sensemaking activities - enabling more clarity of what 'voices' prioritise particular trends and issues and shaping the analytical lens that is used to read foresight results from different sources.

What remains to be demonstrated is whether topic modelling or other computational text analysis methods can be useful in foresight activities like the identification of weak signals, emerging issues, or even wildcards. The parameters that shape topic modelling, for instance, can be changed at the beginning of the textual analysis to create any number of topics, and might be experimented with to produce different results. For example, given a different topic cluster output, and the larger corpus we may anticipate, such a method could produce results that point to gaps in the knowledge base of foresight research, or emerging issues that had avoided detection. Given the definition of wildcards as low-probability, high-impact events (Peterson & Steinmuller, 2009), it again remains unlikely that topic modelling or textual analysis would uncover novel types of wildcards. However, there is a possibility that other forms of textual analysis could help to uncover cascading discontinuity sets - series of small, unrelated events whose cumulative effect can be similar to a wildcard (Barber, 2006). There is also the possibility that given enough citizen-sourced contributions to future-oriented activities, surprises and novelty will emerge from NLP methods of analysis. The question then becomes: will sensemaking activities within governing institutions be able to recognize these results as useful disruptions to ongoing foresight, or might such surprises be lost of organizational culture, or the connection between power and the assessment of plausibility?

5.2. Comparing NLP results to qualitative assessment

The above analysis enables a mode of sensemaking that utilizes the analytical data of the NLP process by allowing us to distinguish between citizen-sourced priorities with regard to desirable futures, and those priorities that emerge from expert or institutional foresight sources. In order to assess the validity of these results, we further compare them with results from a previous study that accomplished the same comparison through the use of qualitative methodology which produced two metrics: qualitative alignment and degree of coverage. Qualitative alignment concerned a semantic analysis of statements while degree of coverage examined terminological references (Rosa et al., 2018b).

In the initial CIMULACT comparative study, the topics that demonstrated the highest alignment between citizens and experts were (in descending order): 1) Data for all - share the power of data, 2) Transforming technologies for planet and people, 3) Making dense and growing urban areas more sustainable and liveable, 4) Evidence-based personalized healthcare, 5) Smart-energy governance, 6) Empowering diversity in community, 7) Production awareness, 8) Snakes and ladders - connecting scales of issue and actors, 9) I am empowered to lead my changes, 10) Meaningful research for community, 11) Personal and organizational choice management, and 12) Responsible use of land. The metrics on degree of coverage contributed to four additional topics: 13) Dissemination and continuous exploitation of research and innovation in the healthcare system, 14) Good quality food for all, 15) Rethinking (the new) job market needs, and 16) Community building development.

Based on this information, we can link the qualitative research topic comparison to the topic modelling results to better gauge the sensemaking capacity that NLP processes enable. In effect, we compare the qualitative results against the four key citizen-emphasized topics identified in the modelling: 1) Food and energy systems transformation, 2) Public services using data, 3) Multi-level EU development, and 4) Challenge driven research. As viewed in Table 2, the topics of highest importance to CIMULACT as identified by the topic modelling process fit nearly all of the CIMULACT research topics identified as best aligned or covered with expert reports via a comparative reading methodology. These findings demonstrate that the topic modelling process is able to identify areas of interest that are demonstrably similar to qualitative, comparative readings conducted by human teams.

For example, topic modelling confirms results of the qualitative analyses that showed (a) energy discussed as an integrative topic for urban development is a distinct concern compared to the individual technologies that might underlie such systemic change (2018b, Rosa et al., 2018a) participatory energy governance is a rather unique topic derived from citizen-based foresight (Gudowsky & Rosa, 2019). Furthermore, the modelling results pair food and energy systems under a single banner, which is supported by the high number of CIMULACT topics dealing with either food, or energy, and often an arena (urbanization, land use planning, etc.) in which both came into play. Similarly, the topic of *Public services using data* and *challenge driven research* is clearly identified across multiple aligned CIMULACT research topics highlighted in the initial comparative reading. Further, counterparts to the modelled topics of *Multi-level EU development* and *Challenge driven research* can be readily found, and the outliers are few and relate to value-laden calls for change, which topic modelling does not respond to very well.

Both topic modelling and the applied qualitative methodology provide analytical metrics. These metrics are not directly comparable because modelling targeted the identification of a meaningful number of exclusive topics in the full corpus of reports while the applied qualitative methodology targeted the alignment and coverage of CIMULACT report topics across the other reports. This is also the reason for the lesser number of examined topics in the modelling. Nevertheless, metrics provided by both approaches indicate magnitudes of differences, and work well to their aims.

We also find the topic modelling results notable for what they identify as topics that CIMULACT does not focus on. Key examples of topics that are shared across expert reports but are prevalent only to a limited degree in CIMULACT include: 1) National intelligence for governance, 2) Global economic change, 3) National economic growth, and 4) Developing new materials. In the qualitative analysis, we had observed that there was an overarching bias towards both econometrics and techno-optimism within the reviewed expert-based foresight reports (Rosa et al., 2018b). In examining the topic modelling results, these initial observations can now better be buttressed with the results of the 20 topics modelled in the NLP analysis. We would read these four topics as particularly aligned with economy centered concerns (2 and 3), and precise technological agendas (4), with national intelligence for governance being a more of a macro-level societal security concern (1). Additionally, given that only six of the modelled topics did not include some derivation of ’economic’ or ’technology,’ a stronger case might be made regarding the unique contributions of citizen-based foresight activities, and the types of institutional focuses that reflected within expert-based foresight reports more generally.

Given that the human teams spent nearly 8 months in total analyzing the CIMULACT topics in comparison with the expert-based reports that compose our corpus, whereas the topic modelling algorithm and topic analysis took a matter of weeks, there seems to be a strong case that topic modelling methods can certainly offer a new economy of resources during sensemaking phases. However, when topic modeling is applied without sufficient knowledge about the corpus or the field that is assessed, this bears several risks. It could for instance narrow the spaces for imagination and creativity within the sensemaking phase considerably. Further, knowledge of the field helps to create meaningful research designs. Topic modelling could indeed be seen as a powerful tool for summarizing content at a meta-level, and here the tool has specific analytical strength. Nevertheless, while this may provide valuable insights, it may not necessarily lead to a change of perspective. It may for instance lead to reproducing cognitive biases underlying the source documents. Topic modelling could be effectively used to identify non-alignments, which would require additional qualitative effort and additional data. While the CIMULACT results (Rosa et al., 2018) and subsequent publications have shown the multiple use cases for a human-based comparative reading methodology, the resource savings that this research points to are also notable, providing improved opportunities to conduct more extensive participatory agenda setting processes.

6. Conclusions

In this article, we have investigated if there is a value added in utilizing NLP, namely topic modelling, for generating a lens of analysis for the comparison of forward-looking activity results within a context of sensemaking for policy decisions. We asked if and how NLP analysis might confront resource challenges involved in comparative foresight analysis, and if the results of NLP analysis enable new modes of comparative research within foresight and futures studies. Overall, our analysis indicates that the results of the topic modelling succeed in shaping a useful analytical lens for comparative foresight projects. Of course, even if many of the promises attributed to natural language processing (NLP) can be validated, Grimmer and Stewart (2013) emphasize that such methods cannot be substitutes for careful thought and close reading, especially in political contexts, which require extensive and problem-specific contextual knowledge. In this research, for instance, topic modelling benefitted from our researchers' prior knowledge of the corpus' content - a prerequisite for any attempt to accommodate topic model results in context - enabling more efficient and accurate analysis. Sense making defines the heart of the intersection between policy and foresight, and citizen voices can be an informative component at this interface. This research shows the types of unique contributions that participatory methods collect, and demonstrates how topic modelling can enable sensemaking that is better informed by citizen engagement in an efficient and effective manner.

In considering the resource intensiveness of a participatory project - larger funding budgets, time requirements, and more complex management - there is an expectation for such projects to provide deliverable results in line with the investment made. We have demonstrated that the processes deployed during this research points the way towards a greater efficiency in the necessary analytical work that accompanies comparative literature review. The total time invested in the preparation of texts for NLP, the NLP processes themselves, and the analysis of the NLP results was calculated to be less than a single person month. This is in comparison to the nearly eight person months dedicated to the comparative textual analysis work that was conducted within the CIMULACT project. Thus, when time or budget concerns reinforce the restriction of 'sensemaking' activities that are essential to capturing latent topics in multi-actor processes, we are confident that NLP processes can be used to perform quick checks as well as thorough analyses. However, this research does not address the very real risk that exists when automated processes are invoked because of their mechanical efficiency - namely, data-based biases that result in the reduced consideration of nuanced uses of language that might be highly imaginative or creative within a culturally-sensitive context. While NLP offers efficiency, further research should be conducted to assess the qualitative costs associated with this approach, and suggest methods that might complement topic modelling to capture the creative texture and situated richness that participatory contributions provide to broaden the scope of organizational imagination.

We think that the nature of citizen-led, multi-actor, participatory processes creates a threshold for sensemaking activities that many national and international governing bodies are wary to cross. Firstly, if CIMULACT is any indication, then these processes result in qualitatively different research agendas, compared with the results of 'expert-based' research. In this article, we have constructed and piloted a research design that can be used to explore the divergence of expert- and citizen-based research agendas. The context-sensitive nature of citizen-led results creates a pool of variables that may be unfamiliar, or unapproachable, to policy-makers, thus compounding the difficulties such projects face from the onset (Repo & Matschoss, 2018). Despite the important connections that participatory processes can build between higher-level governing policy and the aspirations of the citizenry, it is understandable how such processes have not become standard practice due to the required efforts. We suggest continued testing of different modes of NLP analysis to lower the threshold for, and encourage greater adoption of, citizen-led participatory processes. Additionally, we view the long tail of this research as pointing towards modes of creating greater transparency and accountability within governance institutions, with respect to policy-creation that reflects the long-term aspirations of the collective citizenry.

The European Forum on Forward Looking Activities (EFFLA) has become a critical dialogue for defining the art and praxis of foresight and futures studies, and their integration within governing bodies and organizational policy. As approaches, methods and organizational backgrounds are highly diverse, one primary recommendation issued from this forum was the definition of 'necessary elements' that should be addressed in upcoming FLA: strategic Intelligence, sensemaking, selecting priorities, and implementation (EFFLA, 2015). They furthermore indicate that of these steps, especially the sensemaking phase is somewhat weakly represented and in need of attention and improved methods. Our conclusions point to a mode of textual analysis that can identify unique thematic and topical areas of special interest based on a large corpus of future-oriented reporting. Such identification can be utilized to mark perspectives that are both over- and under-represented, allowing policy makers, and others, to better 'make-sense' of the informational inputs they rely upon. Future research should be focused on testing this model of textual analysis against a new corpus, and using NLP analysis in the examination of future-oriented texts that are not delivered in report form, but might be gathered from online wikis, dynamic argumentative Delphi survey, or citizen-based forums and observatories. Given new sources of future-oriented information, we think this technique might continue to provide useful insights and can foster new modes of synthesis for policy creation.

Acknowledgements

The CIMULACT project was funded by the Horizon 2020 Framework Program of the European Union H2020-EU.5.c. Grant agreement ID: 665948.

Annex 1 Foresight reports composing the corpus

Report Title	Report Shortname for this research (Table 1)	Year Pub.	Time Horizon	Methods	Institution	Location	Authoring Inst.
Using foresight to support the next strategic programming period of Horizon 2020 (2016-2018)	EFFLA	2014	H2020	Drivers of Change, 'Disruptors'	European Commission	EU (Brussels)	Vincent Rousselet & Associates Ltd
100 opportunities for Finland and the world: Radical Technology Inquirer (RTI) for anticipation/evaluation of technological breakthroughs.	Finland 100	2014	Not Explicit	Technological Forecast,	Committee for the Future	Finland	VTT, What Futures Ltd., Solveto Ltd.
An OECD Horizon Scan of Megatrends and Technology Trends in the Context of Future Research Policy	OECD - Megatrends	2016	2025	Trend Analysis, Wildcards, Technological Forecast,	Danish Agency for Science, Technology and Innovation	Denmark	OECD
Recommendations for an optimized implementation of Horizon 2020	CRF	2012	H2020	Trend Analysis	Copenhagen Research Forum	Denmark	Copenhagen Research Forum
Visions for Horizon 2020. Copenhagen Research Forum II	CRF	2012	H2020	Scenario Development, Technological Forecast,	Copenhagen Research Forum	Denmark	Copenhagen Research Forum
Foresight Services to Support Strategic Programming within Horizon 2020	RAND	2014	H2020	Scenario Development, Technological Forecast, Trend analysis	European Commission	EU (Brussels)	RAND Europe
Envisioning 2030: US Strategy for the Coming Technology Revolution	Envisioning 2030	2013	2030	Trend analysis, Technological Forecasting	The Atlantic Council of the United States	USA	The Atlantic Council of the United States
Ten Technologies which could change our lives: Potential impacts and policy implications	Ten Technologies	2015	Not Explicit (21st C)	Technological Forecasting	European Parliament	EU (Brussels)	European Par-liamentary Research Service
Preparing the Commission for future opportunities- Foresight network fiches	Fisches 2030	2015	2030	Technological Forecasting, Trend Analysis, Workshops	European Commission	EU (Brussels)	EC Foresight Network, 21 DGs participated.
Global Trends to 2030: Can the EU meet the challenges ahead?	ESPAS 2030	2015	2030	Trend Analysis	European Strategy and Policy Analysis System	EU (Brussels)	European Strategy and Policy Analysis System
The Global Economy in 2030: Trends and Strategies for Europe	Global Economy 2030	2013	2030	Trend Analysis	European Strat-egy and Policy Analysis System	EU (Brussels)	Centre for European Policy Studies (CEPS)
Making Europe Open and Polycentric. Vision and Scenarios for the European Territory towards 2050	ESPON 2050	2014	2050	Scenario Development, Trend Analysis, Mapping	ESPON monitoring committee	EU (Luxembourg)	MCRIT, IGEAT, Tersyn, HAS RCERS, S&W, IOM, IGSO, RIKS, SGH, ISIS, U of Thessaly, ERSILIA Foundation
Strategic Foresight: Towards the 3rd Strategic Programme of Horizon 2020	Strategic Foresight	2015	H2020	Drivers of Change, Scenario Development (2030),	European Commission	EU (Brussels)	SAMI Consulting
European Value Changes Signals, Drivers, and Impact on EU Research and Innovation Policies	European Value Changes	2016	Not Explicit (H2020)	Trend Analysis, Weak Signals	European Commission, DG R&I	EU (Germany)	IISA, Fraunhofer ISI
OBSERVE: Observing Emergence.	OBSERVE	2017		Horizon Scanning, Sensemaking, Online Consultation.	European Commission	EU (Germany)	Fraunhofer ISI
State of the Future, 2013-2014	State of the Future	2014		Trend analysis, Horizon Scanning	The Millenium Project	USA	The Millenium Project

Annex 2 Distribution of topics across foresight reports (%)

CLUSTER	Economic development			Values and rights	Technological innovation				Technology and markets			Governance and public services			Challenge centered research					
REPORT\TOPIC	Global economic change	National economic growth	EU in a globalized world	Societal values and human rights	Developing new materials	Applications for trending technologies	ICT driven waste management	Advanced ICT research	Key technologies and market development	Competitive leadership through innovation	Comparing national digital developments	Public services using data	Multi-level EU development	National intelligence for governance	Food and energy systems transformation	Bioeconomy for solving Grand Challenges	Challenge driven research	Energy production and distribution technologies	Economies of demographic change	Global population and resource scarcity
CIMULACT	2,6	0,2	0,0	0,5	0,7	0,8	0,4	0,0	5,4	0,0	0,2	27,6	6,2	0,8	47,7	2,4	4,0	0,0	0,0	0,6
ESPAS-2030	26,2	7,0	31,3	0,0	0,3	0,0	0,2	0,0	8,7	0,0	0,4	10,0	2,9	4,3	5,0	0,9	0,4	0,0	1,5	0,7
Global Economy 2030	23,6	53,1	0,9	0,0	1,1	0,4	0,0	0,0	8,7	0,0	0,0	1,6	1,7	4,1	4,1	0,1	0,2	0,1	0,3	0,0
ESPON 2050	14,1	4,8	0,0	0,0	0,0	0,0	0,2	0,0	4,7	0,0	0,0	4,3	62,2	0,3	8,4	0,4	0,1	0,5	0,0	0,0
Strategic Foresight	12,2	0,0	0,2	0,0	3,2	0,5	2,0	0,0	11,4	0,0	0,0	15,2	1,4	3,6	10,4	37,4	0,7	0,0	1,7	0,0
European Value Change	6,5	1,0	0,1	48,9	1,5	0,1	0,0	0,0	6,6	0,0	0,0	25,8	0,5	1,5	6,3	0,3	0,9	0,0	0,0	0,0
OBSERVE	2,1	0,0	0,0	0,0	31,4	3,1	0,9	0,0	5,4	0,2	0,3	11,7	0,0	0,1	6,6	29,4	6,1	0,3	0,3	1,8
State of the Future	13,7	1,9	1,0	0,7	5,1	0,2	0,6	0,0	3,5	0,0	0,1	5,4	1,4	25,5	7,1	0,3	0,4	1,4	0,6	31,3
EFFLA	18,3	2,8	1,5	0,2	2,4	0,0	0,4	0,0	8,0	0,0	0,2	12,2	2,0	4,9	7,2	5,0	0,0	1,0	32,7	1,3
Finland 100	2,5	0,7	0,0	0,0	29,8	4,9	0,2	0,0	6,9	35,3	0,0	3,0	0,1	5,9	9,0	0,0	0,0	1,4	0,0	0,0
OECD - Megatrends	19,5	6,5	0,5	0,0	9,3	1,4	2,1	0,0	10,5	0,0	24,2	7,9	1,1	4,6	7,5	2,2	0,0	0,5	1,2	0,8
CRF	6,5	0,9	0,0	0,0	2,8	0,0	2,0	0,0	12,9	0,0	0,0	10,7	3,0	1,3	18,5	0,5	40,7	0,0	0,2	0,0
RAND	9,8	2,0	0,3	0,0	2,7	0,9	0,7	31,8	21,7	0,0	0,3	10,7	1,6	3,3	10,9	0,5	1,1	0,3	1,5	0,0
Envisioning 2030	16,0	1,6	0,0	0,1	10,5	2,5	1,9	0,0	9,2	0,0	0,9	4,3	0,1	7,6	3,4	0,8	0,0	41,2	0,0	0,0
Ten Technologies	7,7	0,0	0,0	0,0	7,2	44,0	2,6	0,0	11,5	0,0	0,1	11,7	1,3	3,0	9,3	1,1	0,4	0,0	0,1	0,0
Fisches 2030	9,9	1,0	0,1	0,0	8,9	3,2	28,1	0,1	11,0	0,0	1,2	15,8	2,5	2,1	10,4	2,1	2,8	0,2	0,5	0,0

Annex 3 Topics which are accentuated in the CIMULACT report and in other foresight reports (topical share, %)

	Accentuated in CIMULACT				**Accentuated in other foresight reports**			
REPORT \ TOPIC	Food and energy systems transformation	Public services using data	Multi-level EU development	Challenge driven research	National intelligence for Governance	Global economic change	National economic growth	Developing new materials
CIMULACT	47,7	27,6	6.2	4,0	0,8	2,6	0,2	0,7
Average of other foresight reports	8,3	10,0	1.4*	0,9**	3,3***	12,6	2,2****	7,7
ESPAS-2030	*5,0*	*10,0*	*2,9*	*0,4*	*4,3*	*26,2*	*7,0*	*0,3*
Global Economy 2030	*4,1*	*1,6*	*1,7*	*0,2*	*4,1*	*23,6*	*53,1*	*1,1*
ESPON 2050	*8,4*	*4,3*	*62,2*	*0,1*	*0,3*	*14,1*	*4,8*	*0,0*
Strategic Foresight	*10,4*	*15,2*	*1,4*	*0,7*	*3,6*	*12,2*	*0,0*	*3,2*
European Value Change	*6,3*	*25,8*	*0,5*	*0,9*	*1,5*	*6,5*	*1,0*	*1,5*
OBSERVE	*6,6*	*11,7*	*0,0*	*6,1*	*0,1*	*2,1*	*0,0*	*31,4*
State of the Future	*7,1*	*5,4*	*1,4*	*0,4*	*25,5*	*13,7*	*1,9*	*5,1*
EFFLA	*7,2*	*12,2*	*2,0*	*0,0*	*4,9*	*18,3*	*2,8*	*2,4*
Finland 100	*9,0*	*3,0*	*0,1*	*0,0*	*5,9*	*2,5*	*0,7*	*29,8*
OECD - Megatrends	*7,5*	*7,9*	*1,1*	*0,0*	*4,6*	*19,5*	*6,5*	*9,3*
CRF	*18,5*	*10,7*	*3,0*	*40,7*	*1,3*	*6,5*	*0,9*	*2,8*
RAND	*10,9*	*10,7*	*1,6*	*1,1*	*3,3*	*9,8*	*2,0*	*2,7*
Envisioning 2030	*3,4*	*4,3*	*0,1*	*0,0*	*7,6*	*16,0*	*1,6*	*10,5*
Ten Technologies	*9,3*	*11,7*	*1,3*	*0,4*	*3,0*	*7,7*	*0,0*	*7,2*
Fiches 2030	*10,4*	*15,8*	*2,5*	*2,8*	*2,1*	*9,9*	*1,0*	*8,9*

*Excludes the EPSON 2050 report with a 62.2 % topical share
** Excludes the CRF report with a 40.7 % topical share.
*** Excludes the State of the Future report with a 25.5 % topical share
**** Excludes the Global Economy 2030 report with a 53.1 % topical share.

References

Ahlqvist, T., & Uotila, T. (2020). Contextualising weak signals: Towards a relational theory of futures knowledge. *Futures, 119*, 102543.
Apreda, R., Bonaccorsi, A., dell'Orletta, F., & Fantoni, G. (2016). Functional technology foresight. A novel methodology to identify emerging technologies. *European Journal of Futures Research, 4*(1), 13.

Barber, M. P. (2006). Wildcards – Signals from a future near you. *Journal of future Studies, 11*(1). Tamkang University.
Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3*(Jan), 993–1022.
Boysen, A. (2020). Mine the gap: Augmenting foresight methodologies with data analytics. *World Futures Review, 12*(2), 239–248. https://doi.org/10.1177/1946756720905639.
Brown, A. D., Colville, I., & Pye, A. (2015). Making sense of sensemaking in organization studies. *Organization Studies, 36*(2), 265–277. https://doi.org/10.1177/0170840614559259.
Burgelman, J.-C., Chloupková, J., & Wobbe, W. (2014). Foresight in support of European research and innovation policies. The European Commission is preparing the funding of grand societal challenges. *European Journal of Futures Research, 2*(1), 55. https://doi.org/10.1007/s40309-014-0055-4.
Chowdhury, G. G. (2003). Natural language processing. *Annual Review of Information Science and Technology, 37*(1), 51–89.
CIMULACT. (2019). Cimulact project homepage. *CIMULACT project*. Brussels, checked on 10/10/2019.
Coenen, F. (2011). Data mining. Past, present and futuren. *The Knowledge Engineering Review, 26*(1), 25–29.
Copenhagen Research Forum. (2012). *Visions for Horizon 2020*. CRF.
DASTI. (2016). An OECD horizon scan of megatrends and technology trends in the context of future research policy: Danish agency for science. *Technology and innovation*.
Drahos, P., & Braithwaite, J. (2017). Information feudalism. *Who owns the knowledge economy*. Routledge.
Duckworth, M., Lye, D., Ravetz, J., & Ringland, G. (2016). Commission, European; Innovation, directorate-general for research and; Consulting, SAMI. *Strategic foresight towards the third strategic programming of Horizon 2020*. Luxembourg: Publications Office.
EFFLA. (2015). *How to design a European foresight process that contributes to a European challenge driven R&I strategy process - Policy Brief No. 2*. European Commission.
ESPAS. (2015).): Global trends to 2030. Can the EU meet the challenges ahead?. *European strategy and policy analysis system*.
Foresight Network. (2014). *Preparing the Commission for future opportunities. Foresight network fiches 2030*. European Commission.
Analysis and evaluation of unstructured data. In Gharehchopogh, Soleimanian, F., Khalifelu, & Abbasi, Z. (Eds.), *Text mining versus natural language processing*, (2011). IEEE.
Glenn, J. C., Gordon, T. J., Florescu, E., Project, U., & Millennium, N. (2014). *2013-2014 State of the future*. Washington, D.C: The Millennium Project.
Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis, 21*(3), 267–297.
Gros, D., Alcidi, C., & Behrens, A. (2013). The global economy in 2030. *Trends and strategies for Europe*. Brussels: Centre for European Policy Studies [u.a.].
Gudowsky, N., & Peissl, W. (2016). Human centred science and technology—Transdisciplinary foresight and co-creation as tools for active needs-based innovation governance. *European Journal of Futures Research, 4*(1), 8.
Gudowsky, N., & Rosa, A. (2019). Bridging epistemologies—Identifying uniqueness of lay and expert knowledge for agenda setting. *Futures, 109*, 24–38.
Jouvenel, B. (2017). *The art of conjecture*. Routledge.
Kaivo-oja, J. (2017). Towards better participatory processes in technology foresight: How to link participatory foresight research to the methodological machinery of qualitative research and phenomenology? *Futures, 86*, 94–106.
Kayser, V., & Blind, K. (2017). Extending the knowledge base of foresight. The contribution of text mining. *Technological Forecasting and Social Change, 116*, 208–215.
Kayser, V., & Shala, E. (2016). Generating futures from text—Scenario development using text mining. *Anticipating future innovation pathways through large data analysis* (pp. 229–245). Springer.
Köhler, J., Geels, F. W., Kern, F., Markard, J., Onsongo, E., Wieczorek, A., et al. (2019). An agenda for sustainability transitions research. State of the art and future directions. *Environmental innovation and societal transitions*.
Könnölä, T., Salo, A., Cagnin, C., Carabias, V., & Vilkkumaa, E. (2012). Facing the future. Scanning, synthesizing and sensemaking in horizon scanning. *Science & Public Policy, 39*(2), 222–231.
Kuusi, O., & Vasamo, A.-L. (2015). 100 opportunities for Finland and the world. *Radical technology inquirer (RTI) for anticipation/evaluation of technological breakthroughts* (english edition). Helsinki: Committee for the Future, Parliament of Finland.
Lee, W.-I. (2015). Technology foresight and its implementation strategy. *International Journal of Innovation Management and Technology, 6*(6), 383.
Lewis, D. D., & Jones, K. S. (1996). Natural language processing for information retrieval. *Communications of the ACM, 39*(1), 92–101.
Loveridge, D. (2008). *Foresight. The art and science of anticipating the future*. Routledge.
Matschoss, K., Repo, P., & Timonen, Päivi (2019). Embedding European citizen visions in sustainability transition. Comparative analysis across 30 European countries. *Futures, 112*, p. 102437.
Mazzucato, M. (2019). *Governing missions in the European Union*. Brussels: European Comission. Available online at https://ec.europa.eu/info/publications/governing-missions-governing-missions-european-union_en, checked on 9/17/2019.
McCallum, A. K. (2002). *Mallet. A machine learning for language toolkit*. http://mallet.cs.umass.edu.
Metzger, M. J., Murray-Rust, D., Houtkamp, J., Jensen, A., La Riviere, I., Paterson, J. S., et al. (2018). How do Europeans want to live in 2040? Citizen visions and their consequences for European land use. *Regional Environmental Change, 18*(3), 789–802. https://doi.org/10.1007/s10113-016-1091-3.
Mills, J. H., Thurlow, A., & Mills, A. J. (2010). Making sense of sensemaking: The critical sensemaking approach. Qualitative research in organizations and management. *International Journal*.
Mission Publiques. (2017). Deliverable 2.2 – Social needs based research programme scenarios including 10 to 15 simulated calls for H2020. With assistance of fraunhofer ISI, strategic design scenarios. *Politecnico di Milano*. Brussels: CIMULACT.
Nathan, M. L. (2004). How past becomes prologue: A sensemaking interpretation of the hindsight-foresight relationship given the circumstances of crisis. *Futures, 36*(2), 181–199.
Peterson, J. L., & Steinmuller, K. (2009). "Wild cards" in futures research methodology CD-Rom(Ver. 3). In Glenn & Gordon (Ed.), *Millennium project*. http://www.millennium-project.org/wp-content/uploads/2020/02/10-Wild-Cards.pdf.
Raven, R., Sengers, F., Spaeth, P., Xie, L., Cheshmehzangi, A., & de Jong, M. (2019). Urban experimentation and institutional arrangements. *European Planning Studies, 27*(2), 258–281.
Remotti, L. A., Damvakeraki, T., Nioras, A., Britzolakis, S., Erdmann, L., Cuhls, K., et al. (2016). *European value changes. Signals, drivers, and impact on EU research and innovation policies: Final report*. Luxembourg: Publications Office.
Repo, P., & Matschoss, K. (2018). Citizen visions for European futures—Methodological considerations and implications. *European Journal of Futures Research, 6*(1), 20.
Repo, P., Matschoss, K., & Timonen, Päivi (2017). Sustainable futures. *Comparing Methodologies for Analyzing Citizen Visions in Europe. In Sociology Study, 7*(5), 246–262.
Rohrbeck, R., Battistella, C., & Huizingh, E. (2015). Corporate foresight. An emerging field with a rich tradition. *Technological Forecasting and Social Change, 101*, 1–9.
Rosa, A., Gudowsky, N., & Warnke, P. (2018). But do they deliver? Participatory agenda setting on the test bed. *European Journal of Futures Research, 6*(1), 14.
Rosa, A. B., Gudowsky, N., & Warnke, P. (2018b). *Deliverable 5.2—Report on comparison of research topics from CIMULACT with those from expert oriented foresight studies. In CIMULACT—Citizen and multi-actor consultation on horizon*, 2020.
Rousselet, V. (2014). *Using Foresight to support the next strategic programming period of Horizon 2020*, 2016-2018. European Commission.
Sandberg, Jörgen, & Tsoukas, H. (2015). Making sense of the sensemaking perspective. Its constituents, limitations, and opportunities for further development. *Journal of Organizational Behavior, 36*(S1), S6–S32.
Sotoudeh, M., & Gudowsky, N. (2018). CIVISTI – a forward-looking method based on citizens' visions. *fped, 5*(2), 73. https://doi.org/10.14746/fped.2016.5.2.22.
Strategic Foresight Initiative. (2013). Envisioning 2030. *US strategy for the coming technology revolution*. Atlantic Council.
Tseng, Y. (2018). Public administrators' attitudes toward citizen participation. *Case evidence from the water resources agency in Taiwan*.
Ulied, A., & ESPON. (2014). Making Europe open and polycentric. *Vision and scenarios for the European territory towards 2050*. Luxembourg: ESPON.

van Woensel, L., Archer, G., Panades-Estruch, L., Vrscaj, D., & Parliament, E. (2015). Services, directorate-General for parliamentary research. *Ten technologies which could change our lives. Potential impacts and policy implications: In depth analysis*. Brussels: European Commission.

VOICES. (2014). *VOICES (Views, opinions and ideas of citizens in Europe on science)*. Available online at http://www.voicesforinnovation.eu.

Wagner, S. A., Vogt, S., & Kabst, Rüdiger (2016). The future of public participation. Empirical analysis from the viewpoint of policy-makers. *Technological Forecasting and Social Change, 106*, 65–73.

Warnke, P., Schirrmeister, E., & Rosa, A. (2017). *FET proactive topic generation workshop background document*. Brussels: OBSERVE.

Weick, K. E. (1995). *Sensemaking in organizations* (Vol. 3). Sage.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science, 16*(4), 409–421.

Wiener, M., Gattringer, R., & Strehl, F. (2018). Collaborative open foresight - A new approach for inspiring discontinuous and sustainability-oriented innovations. *Technological Forecasting and Social Change*, 119370.

Winkler, J., & Moser, R. (2016). Biases in future-oriented Delphi studies: A cognitive perspective. *Technological Forecasting and Social Change* (Volume 105,, 63–76. https://doi.org/10.1016/j.techfore.2016.01.021.

Yang, K., & Callahan, K. (2007). Citizen involvement efforts and bureaucratic responsiveness. Participatory values, stakeholder pressures, and administrative practicality. *Public Administration Review, 67*(2), 249–264.