STUDIA UBB EPHEMERIDES, LXIII, 1, 2018 (p. 59 – 78)
DOI:10.24193/subbeph.2018.1.03

AN INSTITUTIONAL VOCABULARY FOR SUSTAINABILITY. THE COMPUTATIONAL APPROACH

ADELA FOFIU[1*] AND RALUCA IGREȚ[2]

ABSTRACT. Sustainability is a concept that incorporates a multitude of perceptions, attitudes and practices, from opinion leaders to decision makers, generating a lack of common framework and vision. Through our research, we question this lack at its origin: language. This paper explores institutional discourses in the European space, with the aim of collecting the vocabulary used in communicating sustainability principles. Automated tools are applied to explore raw text data of public discourses practiced in three European dimensions: European institutions (political dimension), UN consultative bodies (civil dimension) and European news outlets (media dimension). We use a mixed methods and mixed techniques research design, which combines text mining through Topic Modelling (TM) with a glossary-based approach using key words in context (KIWC) analysis. We analyze our corpus to discover keywords, topics proportion, social mood trends, for each type and source of discourse.

Keywords: sustainability lexicon, European discourses, topic modeling, discourse analysis, social moods, NLP

[1] *Senior Researcher, Schloer Consulting Group, Bucharest, Romania.*
[2] *Co-author, PhD candidate in Administration and Public Policies at Babeş–Bolyai University, Cluj-Napoca, Romania*
* *Corresponding author: adela.fofiu@scgee.com*

Introduction

Language is the most dynamic tool that creates and recreates the cultural reality and the frameworks in which this reality is perceived and experienced. Attaining a sustainable future starts at the level of speech, including the vocabulary used to convey sustainability related concepts. Research projects on sustainable lexicons in different social environments show that the vocabulary is insufficiently developed for building an upright and integrated vision of what sustainability is for at least three societal pillars: government, business and civil society. Multiple meanings and attitudes in terms of communicating sustainability principles have led to slow and rigidly adaptive strategies in creating mass awareness about the need to shift towards a mentality of sustainability in all aspects of human activity.

The main purpose of this research is exploring institutional discourses in the European space to collect the vocabulary used in communicating sustainability principles and help build an institutional vocabulary for creating a common vision and framework. Framing theories are used to support the assumption that the way sustainability is communicated today is insufficient to promote cultural change. This paper uses semi-automated content analysis methods in order to extract the main topics that sustainability is framed with. We next proceed to discuss the value of discourse in building sustainable futures. We then discuss discursive frames, prior to a brief introduction in what topic modeling means for qualitativists and humanists. We then present the design and the results of the semi-automated process of topic extraction.

The importance of discourse in transitioning towards sustainability

Globally, the Anthropocene - the idea that the planet has entered a new geological era where the human footprint on ecosystems is irreversible, is gaining more and more popularity[3]. In many practice communities this concept stimulates a certain perception of the urgent

[3] *http://anthropocene.info/anthropocene-timeline.php*

need to develop and implement a common sustainable vision, including the UN's sustainable development goals[4].

The project GLAMURS Green Lifestyles, Alternative Models, and Upscaling Regional Sustainability[5] explored the complex dynamics of economic, political, cultural, social and technological factors that influence sustainable lifestyles and the transition to a green economy. In the closing conference of 2016, Adina Dumitru, PI of the project, discussed one of the key findings from a 3-year research across 7 regions of Europe: sustainable lifestyles do not reach a level of social contagion to determine and to stimulate the transition to sustainable economic models because they receive little recognition from public institutions, in particular from administrations and governments, and from mass media. (Dumitru 2016)

In this regard, the Lexicon of Sustainability[6] initiated in 2009 in the US is a crowdsourcing effort to create a lexicon to help communities express, implement and enrich the imagination of a sustainable world. Gravitating around the "local" concept, the ongoing project gathers the vocabulary of sustainability from practitioners and activists, in a bottom-up approach. For now, this lexicon has not affected US public policies, even though it is popular in civil society, and is becoming more and more known across the academic community. This is consistent with the understanding of Adina Dumitru/GLAMURS about how lexical contagion and practice can vanish in the absence of a positive institutional sanction. This is the context in which the exploration of institutional language in the European space is timely.

Frames in discourse analysis

In an interview for The Guardian (Clark, 2013), George Lakoff explains how the Enlightenment era assumed a Cartesian approach to reason as being conscious, logical and rational, a mainstream

[4] *http://www.un.org/sustainabledevelopment/sustainable-development-goals/*
[5] *http://glamurs.eu/*
[6] *http://lexiconofsustinability.com/*

mentality still widely adopted. Nevertheless, we have since learned that reason is 98% unconscious (Lakoff, 2013 in Franks and Turner, 2013). How people think and talk derives from personal frames which are unconscious processes, sensitive to subjective experience and agendas. Therefore, framing cannot be avoided and requires attention when used by opinion makers. There is little cultural progress to publicly admitting that specific linguistic structures are used in public discourses to enable certain political ideologies. For Lakoff, ideologies are nothing else than systems of frames using certain words (ideological language) which activate the ideological system in the eyes and ears of the public to support or dismiss certain ideas and policies.

What sustainability means is dependent not only on linguistics, but also on how communicators frame sustainability principles. Sustainability is a topic which gained exponential presence in public discourses since the early 2000s, when global media started reporting environmental concerns at an unprecedented rate. The global political tensions around environmental issues were highlighted in 2007, with nations still not endorsing the UN sponsored Kyoto Protocol (United Nations 1997) and the writings on the subject accelerated.

Coverage of environmental issues is no less susceptible to framing. Lakoff (2010) points out that environmental framing is everywhere in the news: depending from which framing a public person speaks (conservative or progressive), they implicitly promote a particular agenda. Conservatives are pushing for the term 'climate change' instead of 'global warming' for being less frightening, 'climate' having a nicer connotation, while 'change' leaves out any human cause of the change (Lakoff, 2010, p.71). On the other hand, progressives use frames such as 'environmental protection', 'well-being' or 'personal responsibility.' The progressive moral system rejects market fundamentalism, sees government as necessary for improving environmental conditions and understands phenomena like global warming as systemic, not direct causation. The conservative moral system sees man above nature, promotes a Let-the-Market-Decide ideology and thinks more in terms of direct rather than systemic

causation (Lakoff, 2010, p.76). Although the public uses versions of language from both conservative and progressive value-systems, the current state of the global environmental crisis shows that conservatives have prevailed in using their language and activating conservative frames.

The use of word play (metaphors, narratives, domain-specific lexicon) and their implicit framings are crucial in communicating sustainability, particularly in how it guides cognition processes and attitudes within the public (Semino et al, 2016). Using positive or negative words with respect to a certain topic e.g. 'death' described as an 'ending' or a 'beginning', a 'tragedy' or a 'natural process of life', suggest different attitudes and perceptions of death. These framings can reinforce or influence different views of understanding death and furthermore affect how receivers relate to it.

The use of specific words, metaphors and framings can be considered in three perspectives (Semino et al, 2016): cognitive (e.g. Lakoff and Johnson, 1980), discourse based - who uses them, why, in what contexts, and with what possible effects and consequences (Cameron et al, 2010), and practice based - how it can help or hinder communication in particular institutional settings (e.g. Reisfield and Wilson, 2004). The three perspectives over discourse can be linked to the three societal dimensions able to influence mentalities through their corresponding framings: government (authority, political - public dimension) influencing moral frames; news outlets (media dimension) influencing attitudes and opinions, and third-sector (civil society, NGOs – civil dimension) influencing actions.

Either being mediated in the news, social media or websites, all three discourse dimensions primarily use specific framings in debating specific topics. The lack of consensus on sustainability may well be due to different frames and languages used by the conservative and progressive value-systems. An agreement that the environment is interdependent with other areas of society such as economics, energy, food, health, trade, and security can be achieved by looking at the frames which promote a global view and can make way for a bipartisan

discourse influencing media coverage and institutionalized frames (Lakoff, 2010). In finding a simple, non-technical language which communicates the facts, repetition is key for pushing a certain frame.

Language needs to be taken seriously and critically. Methodologically, applying a global view to detect language and frame patterns in communicating sustainability entails a complex approach. Critical discourse analysis (CDA) infers a variety of approaches towards the social analysis of discourse (Fairclough, 2012). CDA starts from the premise that content analysis and abstract analysis of context is insufficient and proposes a three-dimensional model for studying communication processes: the text, the discursive practice that produces the text and the socio-historical context of discursive practice. This approach implies working transdisciplinary with discourse, in dialogue with other disciplines and theories addressing contemporary processes of social change (Fairclough, 2012. p. 452). Analyzing texts includes interdiscursive analysis of how genres, discourses and styles are articulated together, while taking into account that these have a hybrid character, as Lakoff's framework includes an awareness of how conservative and progressive language frames intertwine.

Consistent with Lakoff's view of the emotional, less rational man in making decisions and speaking about the world is the socionomic theory of social moods. Robert Prechter, who initiated the field of socionomics, defines social moods as "a shared mental state among humans that arises from social interaction. Social mood predisposes individuals in the group toward emotions, beliefs and actions. It fluctuates constantly in a fractal pattern. It is unconscious, unremembered and endogenously regulated." (Prechter, NA) The innovative approach of socionomics and its understanding of social moods is that the way a group or a population feels about their future is what shapes events (Casti, 2010). Conventional wisdom has it that events, especially collective events, determine how people feel. Here, we adopt the socionomic view of collective states and explore how people express what they feel about sustainability, understanding that these feelings are the source of framing, which determines our cultural capacity to imagine futures.

As follows, in this research we propose an integrated multi-level framework for discourse analysis applied to a corpus of public discourses used to communicate sustainability by European institutions, European NGOs and European news outlets. We use social moods and Lakoff's frames to explore institutional discourses practiced in the European space and showcase how they can be used to detect language biases, disambiguate concepts and their frames of reference.

Methodological considerations and the research procedure

For the purpose of developing an institutional vocabulary for communicating sustainability principles, text mining (Feldman and Sanger, 2007) and a glossary-based approach are used on a complex corpus of public discourses built from three sources:

- Policy agents: The European Administrative Office: the official website of the European Union - www.europa.eu;

- Civil society: The Advisory body to the UN: Global Ecovillage Network official website - www.ecovillage.org, and the official website of Global Ecovillage Network Europe - www.gen-europe.org;

- Mass media: The Guardian website www.theguardian.com, Politico website www.politico.eu, BBC website www.bbc.com, The Independent website www.independent.co.uk.

Having a thematic aim - sustainability, the selection of articles in the sample will be dictated by keywords consisting of three discursive categories: economic sustainability, environmental sustainability and social sustainability (Triple Bottom Line, Slaper & Hall, 2011). The publication period considered for the online materials is January 2016 - April 2017. To observe the variation of discourse on our chosen topic, a sampling pace of one article per week was used if the sources allowed for this frequency. Where this was not feasible, the selection followed the exhaustion of the keywords over a course of each month. The purpose of using this sampling pace is (i) to observe the variation of various frequencies by each data source, (ii) to enable the possibility of creating time series illustrating the variation of moods and sentiment and (iii) to facilitate frame interpretation according to the socio-historical context.

The steps of our research process in extracting and exploring the data consisted of selecting the corpus (by keywords, resulting in 623 articles from all sources), preprocessing the files to prepare them for import in the appropriate software, text mining using topic modeling for extracting main topics per source, identifying frames and social moods per source using a glossary-based approach, and finally data analysis and interpretation.

Text mining

The method of choice for performing natural language processing on our corpus is topic modeling. This method has been initially developed for descriptive and exploratory purposes, as it looks for patterns in how words are used (Graham, Weingart and Milligan, 2017). Texts are considered as a collection of topics, while topics are considered as a collection of words, and each word has a certain probability of belonging to a topic (De Angelis, 2015).

Topic modeling comes with consistent advantages for researchers who are interested in inductive and forensic approaches to text as data, given that it is an unsupervised method of text analysis. In a conventional approach to content analysis and text analysis, specifically, creating a manual coding system is paramount. This requires that the text is known and understood by the researcher prior to performing data collection and analysis. This is why oftentimes a pilot-sample of text is manually analyzed and used to extract a coding system, later to be used in the annotation and analysis of the larger sample. This is time consuming and raises questions regarding inter-coder agreement when several people work on the same data. Unsupervised methods of text mining and text analysis, on another hand, allow for the discovery and emergence of coding categories not previously thought of (De Angelis, 2015). They also allow for the creation of a common understanding of the text among coders and researchers, significantly reducing the time invested in reaching inter-coder agreement.

Structural topic modeling is a specific type of unsupervised text mining that includes structures derived from document meta-data in the analysis of corpora. In other words, the researcher is capable, through structural topic models, to observe how document time and date, author's gender, length of text, etc., structures the detected topics. Roberts, Stewart, Tingley and Airoldi (2013) used STM in the analysis of differing media coverage of China's rise. The analysis on news stories that included the word "China" from 1997 to 2006 from five major newswire services allowed the topic prevalence in the model to vary by year and news source. This strategy indicated in the generated time series how important historical events are associated to differences between the topics these newspapers chose to discuss in relation to these events (Roberts et al, 2013).

The first step in our research design is to perform topic modeling without exploring for structures. In this sense, we decided to work with MALLET[7] (MAchine Learning for LanguagE Toolkit), the friendliest NLP tool for the less mathematically inclined, that requires the use of command line. MALLET is an excellent and simple tool to run text through topic modeling algorithms, with a few strokes of the keyboard. We have used the topic model prepared by Graham et al. (2017), which extracts 20 topics of each input sample.

We have also used Voyant Tools[8] in order to generate some topic visualizations, starting from key words in context (KWIC). This tool allowed us to understand how the words that compose topics interact with each other, in concept maps that highlight linguistic frames in the chosen discourses.

A glossary-based approach

Next, we have built a glossary of positive and negative social moods, using WordNet[9]. This is a tool developed by Princeton

[7] *http://mallet.cs.umass.edu/*
[8] *http://voyant-tools.org/*
[9] *https://wordnet.princeton.edu/*

University, which comprises a large lexical database of English: "Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations." (Princeton University, 2010) We searched for each social mood noun from Prechter's list in WordNet and manually collected each first order synonym. We have stored the newly generated glossary in Yoshikoder[10], a cross-platform multilingual content analysis program, ready for applied text analysis on our samples using keywords in context and concordance analysis.

Results and data analysis

Our topic modeling approach extracted 20 topics for each sample, by source. As Table 1 below shows, there is a slight divergence in how the media, policy makers and the civil society construct their discourses on sustainability. The cognitive framing of government discourses revolves around the impact of business and industry on the environment and the public through growth, products and services, at both country and international levels.

Table 1. Topic models with the highest weight, as computed by MALLET

Media/Guardian	18	social future time environment waste areas trade increase recent provide report action large levels current real improve rights systems costs
Media/BBC	9	sustainability people world global food years climate environmental local change community life university time place land system living home society

[10] http://yoshikoder.sourceforge.net/index.html

Media/Independent	16	sustainable business water environmental development companies businesses hemp environment company it's emissions green industry pollution plastic goals care
Media/Politico	16	home brexit politics supermarkets nhs worst read theresa crisis they're department shirt trust student win marriage class threats statistics logo
Government/EU	0	social sustainable public time food growth environment make business environmental including international impact services system industry access data country products
Civil Society/GENE	0	environmental sustainability sustainable climate company it's report emissions green home power farmers carbon access data communities land goals air human
Civil Society/GEN	3	social food work business public waste report environmental make sustainability growth years data government part industry development impact change services

The discourse-based framing of mass media equally regards the impact of businesses, trade, companies on the environment, but with a closer look to specific manifestations of this impact (waste, costs, pollution, plastic, emissions). Media also focuses on the role of people and communities (local, community, people, university) and on the value of natural resources (water, hemp, food). The practice-based framing of civil society changes the perspective and highlights the role of government and industry in sustainable growth with a focus on goals, carbon emissions, waste and on the impact on farmers, land and air.

Concept mapping with Voyant Tools on the three subsamples - media, government, civil society discourses - illustrates the topic models that MALLET has extracted. The media discourse is the most complex of the three, being the only one that frames sustainability as a complex issue, at the interplay of economy, society and environment (Fig. 1).

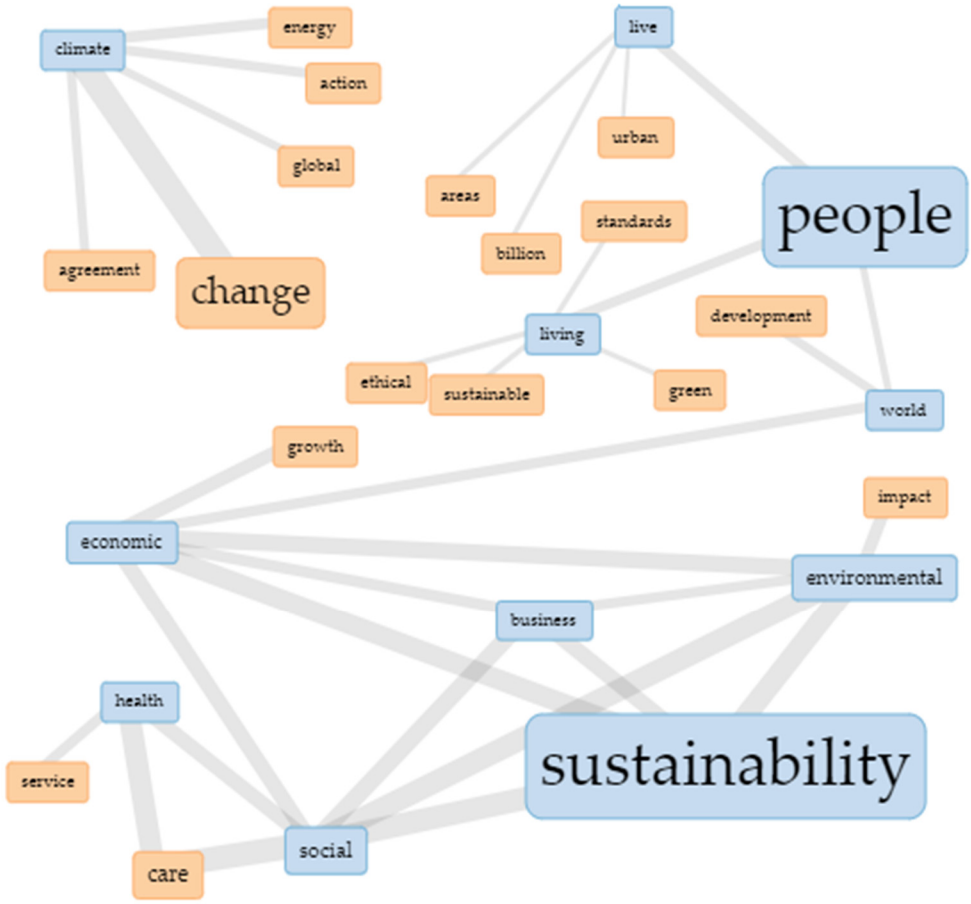

Fig. 1. Concept mapping the mass media discourse on sustainability

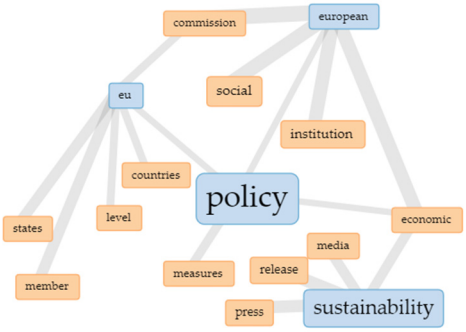

Fig. 2. Concept mapping the government discourse on sustainability

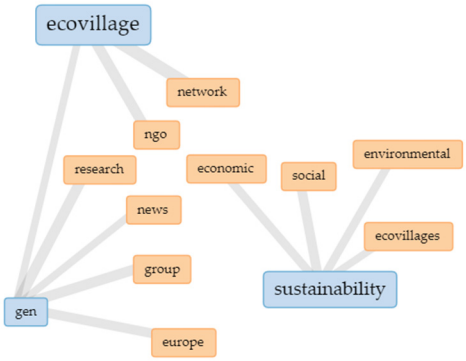

Fig. 3. Concept mapping the civil society discourse on sustainability

The civil society discusses sustainability as environmental, social and economic, but does not connect the three dimensions (Fig. 2). At its turn, policy makers discuss sustainability as an economic issue in the European sphere, with no reference to the other two aspects highlighted by the media and by the civil society (Fig. 3).

An overview of social moods and emotional frames in the discourses, using our glossary-based approach, shows that all three societal actors discuss sustainability issues in a predominantly positive note. It is interesting to note, though, that the tendency towards negative social moods is the highest in the media discourse (Fig. 4), with almost 10% stronger than the government or civil society discourses (Fig. 5 and 6). This begs the question of what particular negative mood predominates each type of discourse. Thus, we have chosen three negative moods and three positive moods to look at. Figure 7 shows that fear predominates the discourses of all three societal actors and it is the highest in the media discourse. This is associated with supportiveness in all the discourses, which is by far the highest in the discourse of the civil society (Fig. 8). But how are these social moods associated with our key word, sustainability?

We have returned to Yoshikoder to address this question, by generating a concordance analysis of KIWC sustainability, then running the social moods glossary on the resulting concordance list. Figures 9 and 10 illustrate that the overall negative tendency of fear and the overall positive tendency of supportiveness still hold valid when looked at in the context of KIWC sustainability.

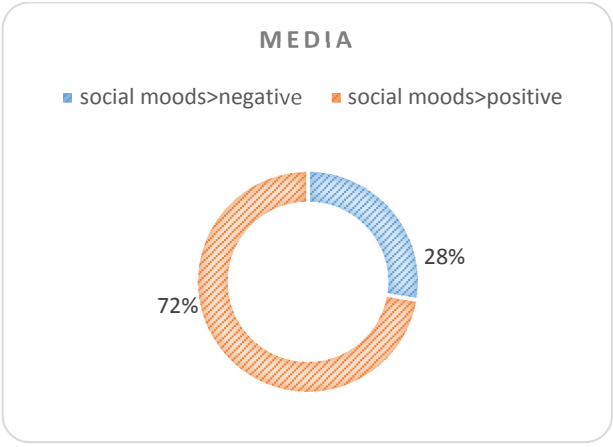

Fig. 4. Negative social moods vs positive social moods in media discourses

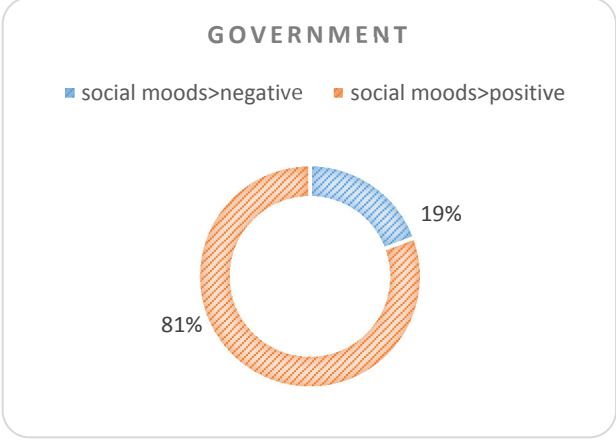

Fig. 5. Negative social moods vs positive social moods in government discourses

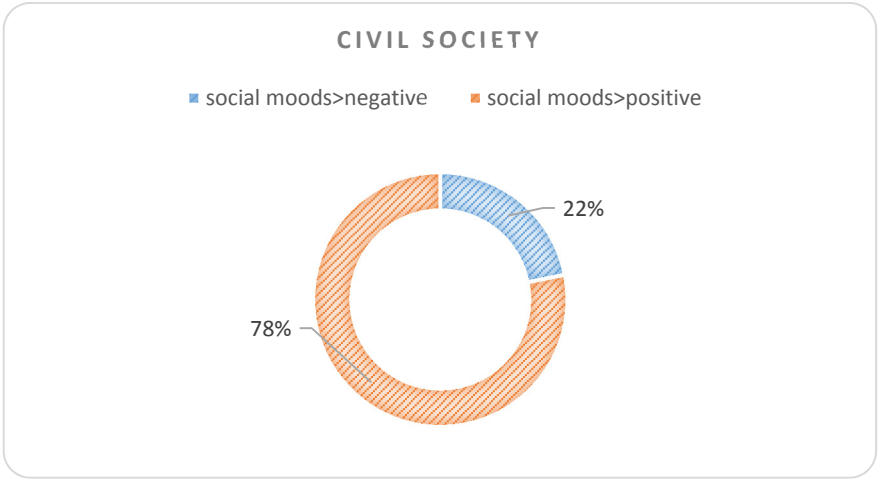

Fig. 6. Negative social moods vs positive social moods in civil society discourses

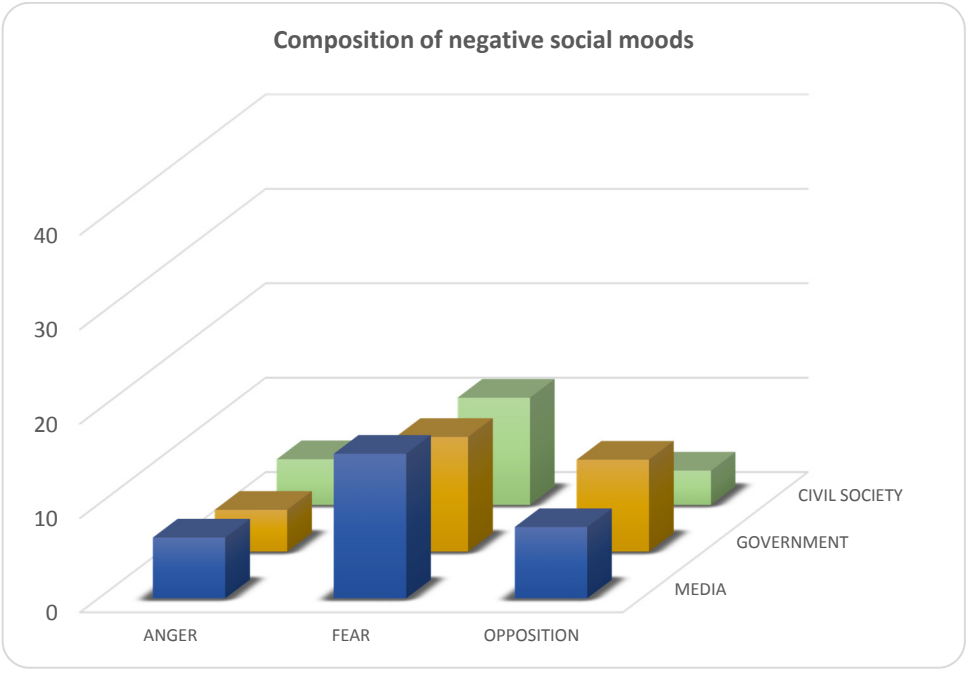

Fig. 7. Composition of negative social moods in all three discourse samples

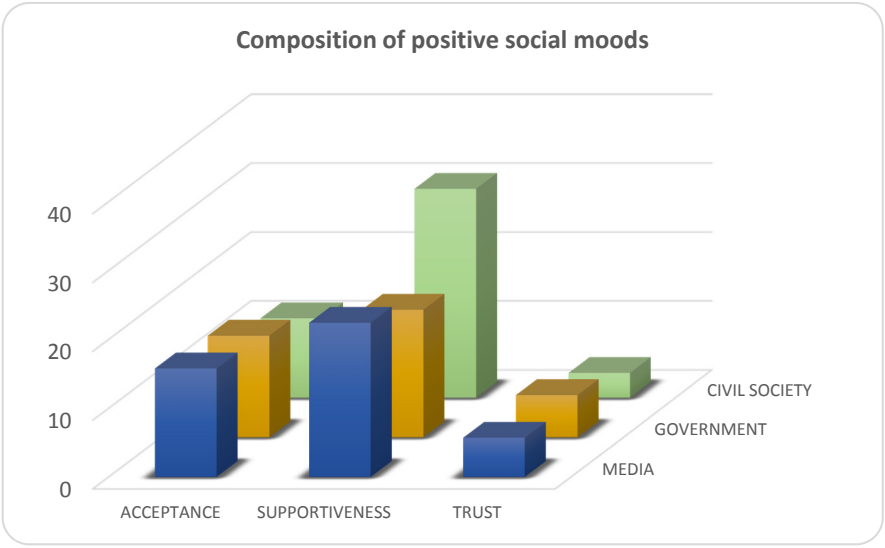

Fig. 8. Composition of positive social moods in all three discourse samples

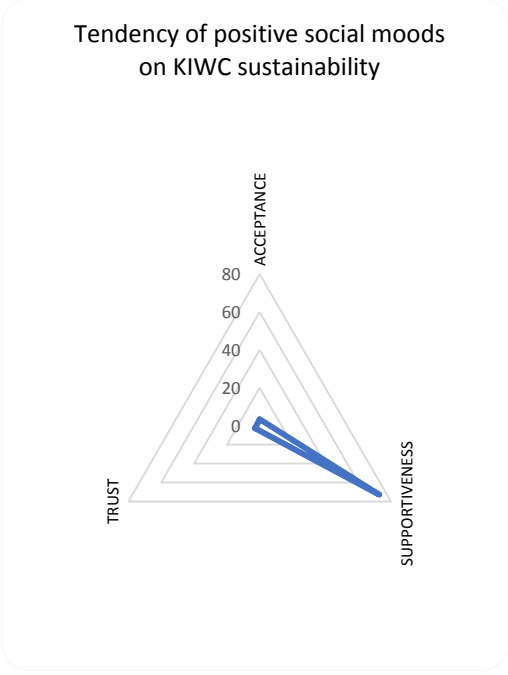

Fig. 9. Tendency of positive social moods on KIWC sustainability

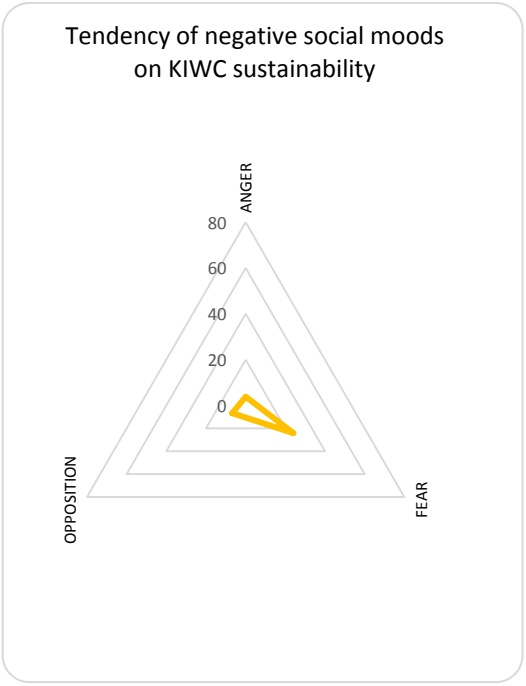

Fig. 10. Tendency of negative social moods on KIWC sustainability

Discussion and conclusions

The main objective of our mixed methods and mixed techniques research design was to explore how sustainability is communicated by three types of European institutional actors: mass media, policy makers/ government and civil society. On our overall sample of 623 articles published between January 2016 and April 2017, we have applied text mining techniques, a glossary-based approach and key words in context analysis. We have thus identified that mass media constructs the most complex discourse addressing issues related to sustainability, by discussing this concept as a phenomenon at the intersection of economy, society and environment, mediated by the important role of businesses. The other two actors understand sustainability in more simple terms,

by not connecting the various dimensions of the phenomenon and by constructing a disconnected discourse. In this sense, the discursive frames that mass media uses on the issue of sustainability are notable and indicate, in comparison, that media is a much more active and complex actor when it comes to stressing the importance of sustainability. Even though our research is exploratory, this result is somewhat surprising, as common sense would have indicated that the civil society is more prone to highlighting the sensitive issues of our age. It is consistent, then, to observe that media discourses are dominated by supportiveness and fear, a strong combination that indicates that the value of sustainable future is constructed as sensitive and of importance. Of the three actors, civil society manifests the highest supportiveness for the value of sustainability as practice, while media manifests the highest fear related to this issue. The policy-oriented discourses prove to be the least active and critical, given that sustainability is represented as solely economic and dependent on country reports and international agreements. This is not surprising, at a second glance, since it confirms the observations of researchers at GLAMURS: social change for sustainable futures is not sanctioned by governments, leaving the most advocacy and practice on the responsibility of communities and media.

The methodological and procedural cocktail that we have employed in processing and analyzing our sample and sub-samples, as demonstration of how powerful content analysis can be, is an encouraging exercise to continue importing knowledge from computational sciences to the field of social sciences and communication studies. Text mining can greatly benefit the development of critical thinking towards the many types of discourses that the public space accommodates, increasing awareness of how media products work and of the role communicators from all fields – media, government, civil society – play in the representation and construction of our futures.

References

CASTI, John, (2010), Mood Matters. From Rising Skirt Lengths to the Collapse of World Powers, Copernicus Books, Springer Science+Business Media, New York.

CLARK, A. M., (2013) George Lakoff to green marketers: use the F-word, The Guardian, available online at *https://www.theguardian.com/sustainable-business/george-lakoff-green-marketing*, retrieved on 27 March 2018.

DE ANGELIS, Andrea, (2015), Structural Topic Models to study political texts: an application to the Five Star Movement's blog, available online at *https://me.eui.eu/andrea-de-angelis/blog/structural-topic-models-to-study-political-text-an-application-to-the-five-star-movements-blog/*, retrieved on 10 April 2018.

DUMITRU, Adina, (2016), GLAMURS overview. Framing discussion: Research, policy and practice, within the conference presentation Making time for sustainability: promising avenues for transitions to sustainable lifestyles and a green economy, 30 noiembrie-1 decembrie 2016, Bruxelles, Belgia.

FAIRCLOUGH, N. (2012), Critical Discourse Analysis, International Advances in Engineering and technology (IAET), Vol. 7 July 2012, International Scientific Researchers, accessed on March 28, 2018 *http://scholarism.net/FullText/2012071.pdf*

FAIRCLOUGH, Norman and Isabella FAIRCLOUGH, (2012), Political Discourse Analysis: A Methods for Advanced Students, London: Routledge.

FELDMAN, R., & SANGER, J. (2007). The text mining handbook: advanced approaches in analyzing unstructured data. Cambridge University Press.

GRAHAM, Shawn, SCOTT Weingart, and Ian MILLIGAN, (2017), Getting Started with Topic Modeling and MALLET, available online at *https://programminghistorian.org/lessons/topic-modeling-and-mallet*, retrieved 15 December 2017.

JABER, David, (2009), The Language of Sustainability: Why Words Matter, available online at *https://www.greenbiz.com/blog/2009/11/18/language-sustainability-why-words-matter*, retrieved 20 April 2017.

KNOWLES, Bran; Lynne BLAIR; Stuart Walker, (2014), Toward a Sustainability Lexicon and Pattern Language?, Workshop position paper submitted to CHI'14, workshop entitled: What have we learned? A SIGCHI HCI & Sustainability community work-shop, available online at *http://eprints.lancs.ac.uk/69376/*, retrieved 20 April 2017.

LAKOFF, George, (2008), Idea Framing, Metaphors, and Your Brain, YouTube video by For a.tv, published on July 16, 2008, accessed on March 27, 2018, *https://www.youtube.com/watch?v=S_CWBjyIERY*.

LAKOFF, George, (2010), Why it Matters How We Frame the Environment, Environmental Communication, 4:1, 70-81, accessed on March 27, 2018 *https://www.tandfonline.com/doi/pdf/10.1080/17524030903529749*

PRECHTER, Robert, (NA), Social Mood, available online at *https://www.socionomics.net/2011/04/social-mood/* accessed on March 27, 2018.

PRECHTER, Robert, (2011), Socionomic Causality and the Standard View, lecture at the World Futurist Society in July, 2010, available online at *https://www.youtube.com/watch?v=xpZAyIiJlyE*, accessed on March 27, 2018.

ROBERTS, Margaret E., Brandon M. STEWART, Dustin TINGLEY, and Edoardo M. AIROLDI, (2013), The Structural Topic Model and Applied Social Science, available online at *https://scholar.princeton.edu/files/bstewart/files/stmnips2013.pdf*, retrieved on 10 April 2018.

SLAPER, Timothy; Tanya HALL, (2011), The Triple Bottom Line: What Is It and How Does It Work?, Indiana Business Review, vol. 86, issue 1.

SEMINO, E., DEMJEN, Z., and DEMMEN, J., (2016), An Integrated Approach to Metaphor and Framing in Cognition, Discourse, and Practice, with an Application to Metaphors for Cancer, Applied Linguistics 2016: 1-22, Oxford University Press.