

Integrating and Differentiating Meanings in Tweeting about the 5th

IPCC report

Kim Holmberg^{1,*}, Iina Hellsten²

¹Research Unit for the Sociology of Education, University of Turku, Finland

²Department of Organization Sciences, VU University Amsterdam, The Netherlands

**kim.j.holmberg@utu.fi*

Abstract

The publication of the 5th IPCC (Intergovernmental Panel on Climate Change) Working Group 1 report in September 2013 was highly debated in Twitter. In this paper we focused on Tweets that mention “IPCC”, and in particular the content and sentiment of the tweets sent by tweeters that were identified as unconvinced or as convinced towards the scientific basis of global warming. Our results indicate that the content and sentiment of those convinced reflect mainly information sharing activities instead of expressing opinions or participation in the debate. Climate change science is, however, challenged by some unconvinced tweeters who tend to use more negative words in their tweets. Our theoretical contribution is on the processes of meaning making around the IPCC report in relation to different groups of tweeters. We identify how certain words may be given different meanings by different groups, and how certain words have a differentiating function between the groups and integrating function within the groups. Our results increase our knowledge about the content of climate change debate in social media and on Twitter in particular and contribute to research interested in how words function as differentiating and integrating meanings between and within social groups.

Introduction

The IPCC reports assess scientific information around human-induced climate change and set grounds for mitigation and adaptation policies. The publication of the IPCC's Working Group 1 report (IPCC, 2013) represented an important event in a climate change debate that has become increasingly polarized, especially in some countries like the United States (Painter, 2014). This polarization has emerged between people convinced by the findings of climate change science, and those who are not convinced by them or by some of them (e.g., Hoffman, 2001; Corner *et al.*, 2012). Tweeting about the fifth IPCC report provides a rich dataset to investigate whether this polarization also manifests itself in tweets sent around the publication of the IPCC reports.

We analyze Tweets that mention "IPCC" over a period of two weeks around the publication of the IPCC's WG1 Summary for Policymakers report, on 27 September 2013. We used this data to develop further the idea of integration and differentiation of meanings (Leydesdorff, 1994), as represented by the words used by different groups of people. Our theoretical approach is interested in how words can function both as means of integration and as means of differentiation in meanings across different social groups or communities. On a more practical level the goal of this research was to analyze the content of the tweets sent by two different groups (unconvinced and convinced) and to map the sets of shared and unique words and phrases across the two groups. Through a content analysis of the words and phrases that can be considered as either integrating or differentiating meanings within and between the groups, we can analyze the different meanings that are attached to Twitter conversations connected to climate change

science in general and the IPCC in particular. Our results will increase understanding of the content of tweeting about climate change and the IPCC.

Differentiation and integration of meanings

The dynamics of meaning-making and meaning-giving has been studied by Leydesdorff (2011) at different social levels. At the group level, meanings tend to integrate into shared sets of common sense, such as shared jokes and metaphors that may strengthen the in-group identity. At the system level, self-organization tends to lead to differentiation of meanings into specialist jargons, for instance across scientific disciplines. The dynamics of integration and separation of meanings at different levels of social groups and their meanings calls for further research (Hellsten & Groenewegen, 2013).

The process of differentiation and integration of meanings can be observed by studying how words are used in a variety of texts (Leydesdorff & Hellsten, 2005). Some words, such as metaphors and key phrases, provide a common ground for integration of different discourses in society. Terms such as “Frankenfoods” or “climategate” are powerful tools of integrating meanings around issues such as genetically modified foods and climate change because they open up specific perspectives on the issues (Hellsten, 2003; Hellsten & Vasiliadou, 2015) and the related debates. Other words, in turn, work for differentiation of meanings by functioning on a more general level, and thereby opening up a set of possible meanings that can be further re-defined in the communications. For example, words like “stem cells” can contribute to boundary construction as the meaning is defined differently in different context and by different

communities. “Stem cells” mean different issues to medical professionals than to lay publics, for instance (Leydesdorff & Hellsten, 2006). This idea of words as integrating meanings into simple perspectives, and words as providing differentiation across the social groups, can be analyzed by investigating the various types of words used by different communities in their communications about the same topic.

Semantic networks

Semantic network research has developed generally in two ways that approach word networks differently. One branch of research builds upon content analysis and uses manual coding of the phrases or words to be included in the analysis (e.g., Danowski, 2012; Carley & Kaufer, 1993; Diesner, 2013), while the second approach tends to produce co-word maps without human coding, either focusing on the underlying, implicit semantics in co-word networks (e.g., Leydesdorff, 1994; Leydesdorff & Hellsten, 2005; Vlieger & Leydesdorff, 2011) or extracting noun phrases from the analyzed texts (e.g., BrodinDanell, 2014; Heersmink *et al.*, 2011).

In the early approach by Carley and Kaufer (1993), concepts were defined as the nodes in the network and density was considered as the frequency of co-occurrences of the concepts in the network. Later on, this approach has evolved into established and systematic research on structures of concept networks using dedicated software packages, like AutoMap and ORA, that are able to help automating the analysis of large sets of texts (e.g., Pfeffer & Carley, 2012; Diesner & Carley, 2005; Diesner 2013; Martin, Pfeffer & Carley 2013). This approach extends upon content analysis and requires human coding of the concepts. In a similar way, Danowski (2012; also

Danowski & Cepela, 2010) has developed an include list to semi-automate which words will be taken into account in the analysis.

The second strand of semantic co-word networks, Leydesdorff (1997; Leydesdorff & Hellsten, 2005; 2006) has developed an automated semantic co-word maps approach to map the implicit frames in text documents without human coding. This so called vector-space model of mapping words is based on word-document matrices (Turney & Pantel, 2010). In the semantic maps approach (Leydesdorff & Hellsten, 2005) context is defined as a specific topic under discussion. It takes into account not only dyads of co-occurring words, but also single words, triads etc. In addition, the method is able to take into account the relations of co-occurring words and the positions of such words in a set of documents. In the resulting semantic maps, meaning is considered as positioning the words used in the communication within network relations (Leydesdorff & Hellsten, 2005). This approach is suitable for identifying implicit semantic relations in large sets of text documents.

In this paper, VOSviewer software was used for the semantic network analysis of the words used in the set of tweets. The approach included in the VOSviewer software (Van Eck & Waltman, 2010) builds on Van Eck *et al.* (2010) and instead of finding the co-occurrences of words it focuses on the co-occurrences of noun phrases in the texts. VOSviewer uses part-of-speech tagging to identify for instance the verbs, nouns, and adjectives in the texts and then extracts the noun phrases, i.e. “word sequences that consist exclusively of nouns and adjectives and that end with a noun”. Then the most relevant noun phrases are chosen based on the distribution of their co-occurrences “compared with the overall distribution of co-occurrences over noun phrases” (Van Eck

& Waltman, 2011). By choosing the most relevant noun phrases, frequent and therefore more general noun phrases are eliminated from the analysis, while the focus is turned on the more specific, descriptive noun phrases.

Research aims and questions

In this paper, we will focus on a set of shared noun phrases that seem to connect different groups of people engaged in the climate change communication on Twitter and a set of words or noun phrases unique to the different groups and that are thought to separate the groups from each other. We will focus on the most frequently used noun phrases by two different groups of tweeters that were active in tweeting about the publication of the IPCC's Working Group 1 report in September 2013. This case study and medium is interesting for our methodological purposes, namely to experiment with methods to analyze semantic networks from the perspective of integration and differentiation of meanings and the related sentiments. The IPCC report provided a clear cut case study where it was also possible to manually categorize the tweeters according to their stance towards climate change science. This enabled us to connect the tweeters and the meanings communicated via Twitter around one well-defined event; the publication of the IPCC's WG1 report.

Our overall research goal is to analyze the integrating and differentiating meanings, as mapped from the content of the tweets, both within the two groups of tweeters and between them. In addition, we also want to analyze the sentiment of the tweets and evaluate whether the measured sentiments reflect the two groups' stance in the climate

change debate. The research goals can be manifested in the following research questions:

1. What kind of topics and meanings a) integrate and b) differentiate the two groups of tweeters (those convinced and those unconvinced by findings emerging from climate change science)?
2. To what extent does the sentiment of the tweets reflect the two groups' stance in the climate change debate?

Data and Methods

The data consists of English language tweets containing the acronym "IPCC" that were sent between September 17 and October 8, 2013, by tweeters whose stance in the climate change debate had been identified by Pearce et al. (2014). Pearce et al. (2014) identified 239 Twitter users that were frequently mentioned in tweets about the IPCC or frequently authored tweets about the IPCC. These usernames were then coded by two researchers according to the users' stance in the climate change debate as convinced / supportive, neutral or unconvinced / unsupportive. Inter-coder agreement was calculated using the standard Cohen's Kappa (Neuendorf, 2002), which gave a moderate agreement of 0.582. We used the 239 coded Twitter users and their tweets as a starting point for our analysis. In the coding conducted by Pearce et al. (2014) more than half of the 239 Twitter users were coded as convinced by findings emerging from climate change science, roughly a quarter as unconvinced and the rest as neutrals, i.e. not taking a stance in the climate change debate. Most of the tweeters coded as neutral consisted of news media and other sources focusing on merely disseminating

information about the newly released IPCC report. A total of 4,932 tweets were sent by 126 convinced users, 1,052 by 44 neutral users and 4,360 by 54 skeptic users. Fifteen of the 239 Twitter users did not send any tweets, but were frequently mentioned in the tweets. For this research we only used the tweets sent by those tweeters that were coded as convinced or unconvinced.

The noun phrases were extracted from the tweets with VOSviewer (Van Eck & Waltman, 2010). A threshold of a minimum of 20 occurrences was chosen in order to focus the analysis on the most frequently used noun phrases. VOSviewer's default settings for relevance scores were used in order to focus the investigation on the most relevant noun phrases. Based on the co-occurrences of the noun phrases in the tweets semantic networks were created for visual analysis in Gephi (Bastian *et al.* 2009). First a base map was created which contained the most frequently used noun phrases by either group, then the map was filtered according to whether the noun phrases were used by unconvinced, convinced or by both groups. The position of the nodes remained intact for all the maps, but the occurrences of the nodes, the size of nodes and the edges between the nodes changed with each filtered map. This allowed us to compare and identify the similarities and differences in the network structures and in the use of noun phrases within and across the groups.

We also wanted to investigate whether the sentiment of the tweets reflected the different groups' stance in the climate change debate, assuming that those unconvinced about the climate change science would use more negative words in their tweets about the IPCC. The sentiments of the tweets were measured with SentiStrength (<http://sentistrength.wlv.ac.uk/>), which has been developed specifically to detect the

sentiment strength of short informal texts (Thelwall et al., 2011), such as tweets. For each tweet SentiStrength calculates positive and negative sentiment strengths (between 1 and 5, and between -1 and -5), and hence a single tweet can simultaneously have a positive sentiment value and a negative sentiment value. From these an average sentiment score for both groups were calculated.

Results

In this section, we will first discuss the main similarities and differences in the noun phrases used by the two groups of tweeters. Figure 1 shows the semantic networks from the tweets of the two groups combined, while in figures 2-4 the maps have been filtered so that they only show the words used by both groups (Figure 2), words only used by those coded as convinced about the climate change science (Figure 3) and words only used by those coded as unconvinced (Figure 4). The layout of the maps is the same for each map, i.e. the position of the nodes remains intact, but the size of the nodes changes to reflect the proportional number of occurrences accordingly and the edges represent the co-occurrences of the words and the frequencies with which the words have occurred together in tweets.

Figure 1 shows that overall the conversation focused clearly on the actual report (e.g., ipcc climate report, ipcc ar5 report) and the event surrounding its release (e.g., Stockholm, meeting), however, certain subtopics related to the reasons (e.g., carbon, dominant cause) and the implications (e.g., warmest, hottest day) are also visible. By focusing the analysis on the shared and unique words by the groups of tweeters we are able to focus on the attached meanings by the groups separately.

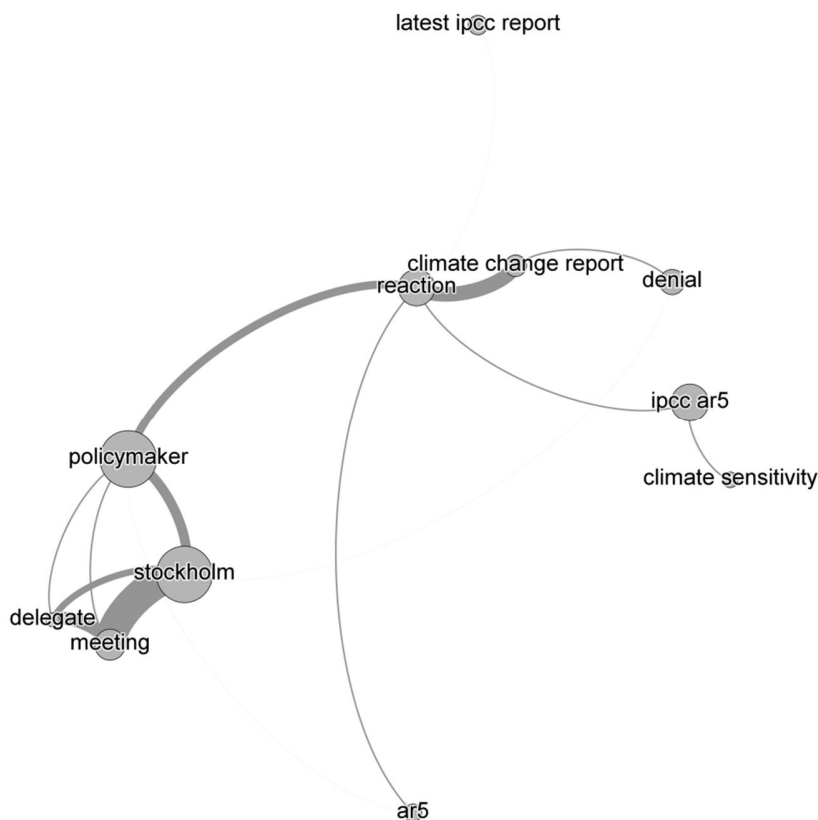


Figure 2. The most frequent noun phrases used by both groups of tweeters

Perhaps even more interestingly, both groups mention “reaction” and “denial”. “Denial” was frequently used by both groups, but in different context with both groups accusing the other of being in denial. There were no single popular tweet causing the word “denial” to appear frequently in the data, as the word was linked to different contexts and different sources. The convinced group of tweeters for instance tweeted and retweeted about “The five stages of climate denial” in tweets like:

ICYM: The 5 stages of climate denial are on display ahead of the IPCC report |

Dana Nuccitelli <http://t.co/xw2UBpmsfB>

while the skeptics tweeted about the denial of IPCC in tweets such as:

RT @[...]: Active denial from the IPCC? Rewriting their past failures to better suit observed temps? <http://t.co/zwHyy5nwin> Surely...

While both groups mentioned “reaction” and shared URLs to blog entries and news articles covering the reactions to the newly released report, they linked to resources supporting their own view about the climate change. These shared words are rather general and mainly related to the publication of the report and reactions thereof, instead of expressing opinions about climate change science, however, they seem to connect and integrate the tweeters into a common context, namely the IPCC report.

Many of the words and noun phrases in the tweets by those convinced about climate change science (Figure 3) connect directly to the report (e.g., ipcc climate report, climate report, working group) or to the content of the report (e.g., finding, risk, influence), while other words are connected to the media attention and press coverage of the event (e.g., live blog, livestream, press release, press conference, journalist, guardian and bbc news). These seem to mainly reflect the information sharing efforts of the group and do not indicate any efforts made to comment on the topic or to express one’s own opinions. However, it is interesting that these two news sources are dominant. The Guardian has, overall, a ‘convinced’ stance, while the BBC tries to be more neutral.

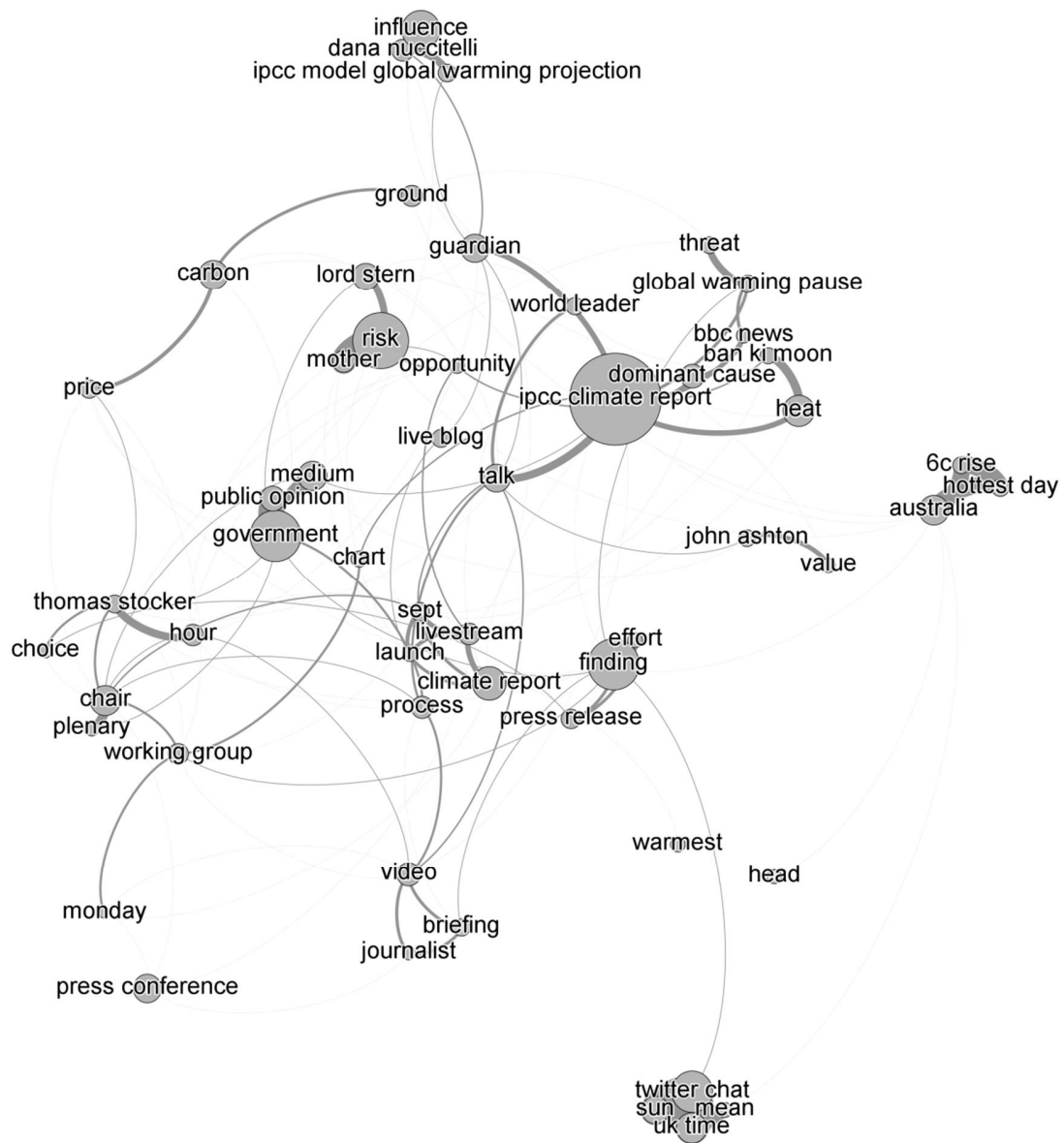


Figure 3. The most frequent noun phrases used only by the convinced group

The names that appear in Figure 3 are public figures, such as the secretary general of the UN and academics engaged in climate science or related areas of research and that were connected to the IPCC WG1 report in different ways. Figure 3 is also influenced by

some frequently tweeted and retweeted tweets that can be seen as tightly connected clusters towards the edges of the graph. For instance tweets like

@[...] we're having a twitter chat about what #IPCC #AR5 means for #health this Sun 5pm GMT/6pm UK time - join us if you can! #IPCChealth

and slightly different variations of

IPCC report: Australia can expect 6C rise on hottest days <http://t.co/t9f5icOTmp> #climate #Australia

can be seen as tight clusters in the map.

Although words referring to the actual report appear in the semantic map based on the tweets from the unconvinced tweeters (Figure 4) as well, these words are often connected to negative or criticizing words, such as “fiction”, “shame”, “failure”, “fault”, and “models wrong”. These can be thought to reflect the tweeters’ disbelief and skepticism towards the IPCC and the WG1 report.

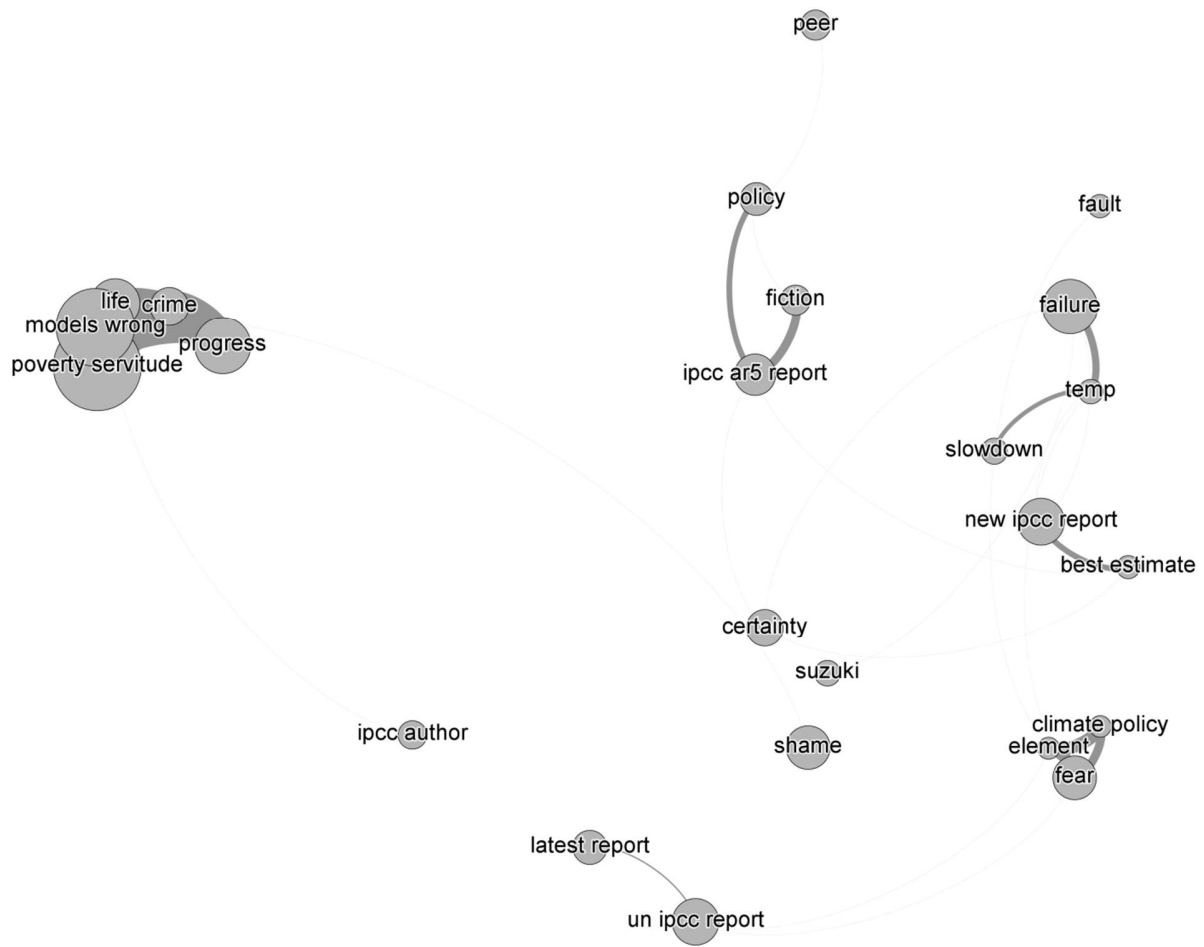


Figure 4. The most frequent noun phrases used only by the skeptic group

The semantic map of the skeptics group also shows some frequently tweeted and retweeted content. The tightly connected cluster (upper left part of Figure 4) with the words “life”, “crime”, “models wrong”, “poverty servitude” and “progress”, is the result of frequent tweeting of variations of the following tweet:

@[...] IPCC Models Wrong - Humanity Sentenced to Poverty/Servitude 4 Sin of progress <http://t.co/3wtN9CE7nY> <http://t.co/rfVRcNnlU5>

The sentiment analysis of the tweets from both groups revealed that those with a unconvinced stance towards anthropogenic climate change had the most negative sentiment in their tweets (-0.618). Those convinced of anthropogenic climate change still scored a negative value (-0.221) but lower than the skeptics, possibly reflecting their more neutral attitude and participation in the debate for mainly information sharing purposes.

Table 3. Sentiment of the tweets from each group of tweeters

	Mean positive sentiment	Mean negative sentiment	Total sentiment score
Skeptic	1.185	-1.798	-0.614
Convinced	1.228	-1.448	-0.221

In summary, the content of the tweets reflects different discourses within the climate change debate, as represented by the tweets on IPCC's WG1 report, and the sentiment of the tweets seem to be connected to the stance the two groups have in the climate change debate.

Discussion and conclusions

The shared set of words and phrases that were used by both groups are general words related to the event or the actual report. At the system level, these words function as a shared ground for the debate, but at the group level their meanings are further defined by the set of unique words used by that group, i.e., different context for "denial" and

“reaction”. The set of unique words, in turn, provides insights into opposing views on what the IPCC report may mean or how the report is being communicated, with skeptics for instance talking about “shame”, “models wrong” and “failure” while those convinced are tweeting about the report in itself and the implications it may have (e.g., influence, risk). Further, while the set of shared words seems to provide integration of meanings across the groups, these shared words may mean different things to the opposing groups, as further defined by the use of unique words that provide more general interpretations to what the IPCC report is about. This set of unique words provides integration of meanings within the groups while functioning as boundaries across the communities.

Our results contribute to research interested in how words function as differentiating meanings across social groups and how they may provide integration between the groups (Leydesdorff & Hellsten, 2005; Hellsten & Groenewegen, 2013). In our case study, general words indicating the topic of discussion seemed to connect the different communities, while words indicating a stance towards the general topic functioned on one hand as differentiating the groups from each other and on the other hand as integrating within the groups. Further research is needed to develop methodological approaches to positioning different types of words within bounded territories (Cardon *et al.*, 2011) as a more detailed approach to integrating or differentiating meanings across communities.

We can conclude the answer to our first research question by stating that the shared words used by both groups give general focus and provide context for the communications, while within the different communities they give different

interpretations, meanings and foci. In addition, the unique words used by the communities may have a differentiating function between the communities, while having an integrating function within the communities. To answer the second research question the sentiment strength of the tweets by the groups were classified and the results indicate that the sentiment of the tweets do, at least to a certain level, reflect the groups' stance in the climate change debate, with the skeptic tweets scoring the lowest negative score and the convinced scoring close to a neutral score. The fact that the convinced scored a negative sentiment score (although clearly closer to neutral than the skeptics) may reflect a level of pessimism in the tweets or it may be a reflection of the chosen vocabulary in their demand for action.

In addition to the above mentioned findings, the visualization techniques used in this study proved to be successful in mapping similarities and differences between the semantic networks of the researched groups of tweeters. By focusing the maps on the different groups we could also map and visualize network dynamics, shared interests and dividing topics.

About the authors

Dr. Kim Holmberg is a research associate at the Research Unit for the Sociology of Education at the University of Turku, Finland, where he works on questions related to bibliometrics, altmetrics, open science and social media. He is also an Honorary Research Fellow at the Statistical Cybermetrics Research Group at the University of Wolverhampton, UK, and he holds the Title of Docent in Informetrics at Åbo Akademi University, Finland. His academic background includes periods as a postdoc researcher

at the University of Wolverhampton, UK, and at the VU University Amsterdam, The Netherlands. He has worked on topics such as hyperlink networks, climate change communications in social media, disciplinary differences in online scholarly communication, and educational use of virtual environments.

E-mail: kim.j.holmberg@utu.fi

Dr. Iina Hellsten is Associate Professor at the VU University Amsterdam, Department of Organization Sciences and affiliated to the Network Institute. Her research has focused on the politics of metaphors, the dynamics of social avalanches in communication networks, and the development of new methods for the analysis of Web based texts, including social and semantic network analysis. Her research has been published in communication sciences (e.g. *Journal of Computer-Mediated Communication*, *New Media & Society*) and in information sciences (e.g. *JASIST*, *Internet Research*). Her current research interests focus on the dynamics of social avalanches, see <http://www.fsw.vu.nl/nl/onderzoek/onderzoeksprogrammas/organisatiewetenschappen/tracking-networks-of-communications/index.asp>

E-mail: i.r.hellsten@vu.nl

References

Bastian M., Heymann S., Jacomy M. (2009). Gephi: an open source software for exploring and manipulating networks. In the *Proceedings of the International AAAI Conference on Weblogs and Social Media*. Available at <http://gephi.org/publications/gephi-bastian-feb09.pdf> (accessed 11 February 2014).

- Brodin Danell, J.-A. (2014). Reception of integrative and complementary medicine (ICM) in scientific journals: A citation and co-word analysis. *Scientometrics*, vol. 98, no. 2, pp. 807-821.
- Cardon, D., Fouetillou, G. & Roth, C. (2011). Two paths of glory: Structural positions and trajectories of websites within their topical territories, In the *Proceedings of Fifth International AAAI Conference on Weblogs and Social Media*. Available at <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/download/2765/3300> (accessed 20 March 20 2014).
- Corner, A., Whitmarsh, L. & Xenias, D. (2012). Uncertainty, skepticism and attitude towards climate change: Biased assimilation and attitude polarization. *Climatic Change*, vol. 114, pp. 732-735.
- Danowski, J. A. (2012). Analyzing change over time in organizations' publics with a semantic network include list: An illustration with Facebook. In the *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2012*, pp. 954-959.
- Danowski, J. A. & Cepela, N. (2010). Automatic mapping of social network of actors from text corpora: Time series analysis. In N. Memon *et al.* (eds.) *Data mining for social network data*, *Annals of Information Systems*, vol. 12, pp. 31-45.
- Diesner, J. (2013). From texts to networks: Detecting and managing the impact of methodological choices for extracting network data from text data, *KunstIntell*, vol. 27, pp 75-78.

Diesner, J., & Carley, K. M. (2005). Revealing Social Structure from Texts: Meta-Matrix Text Analysis as a novel method for Network Text Analysis. In V. K. Narayanan & D. J. Armstrong (Eds.), *Causal Mapping for Information Systems and Technology Research*, pp. 81-108. Harrisburg, PA: Idea Group Publishing.

Heersmink, R., Van den Hoven, J., Van Eck, N.J., & Van den Berg, J. (2011). Bibliometric mapping of computer and information ethics. *Ethics and Information Technology*, vol. 13, no. 3, pp. 241-249.

Hellsten, I. (2003). Focus on Metaphors: The case of 'Frankenfood' on the web. *Journal of Computer-Mediated Communication*, vol. 8, no. 4.

Hellsten, I. & Groenewegen, P. (2013). Combining actors and meanings: Three strategies for network theory, presentation at the Sunbelt conference 2013 in Hamburg, Germany.

Hellsten, I. & Vasilieadou, E. (2015). The creation of the climategate hype in newspapers and blogs: Mixed methods approach. *Internet Research*, vol. 25, no. 4, pp. 589-609. doi:10.1108/IntR-05-2014-0130.

Hoffman, A.J. (2011). Talking past each other? Cultural framing of skeptical and convinced logics in the climate change debate. *Organization and Environment*, vol. 24, pp. 3–33. doi:10.1177/1086026611404336.

Intergovernmental Panel on Climate Change (2013). Summary for Policymakers. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, et al., editors. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the*

Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

Cambridge, United Kingdom and New York, NY, USA: Cambridge University

Press. Available at <https://www.ipcc.ch/report/ar5/wg1/> (accessed 31 March 2014).

Leydesdorff, L. (1994). The evolution of communication systems. *International Journal of Systems Research and Information Sciences*, vol. 6, pp. 219-230.

Leydesdorff, L. (1997). Why Words and Co-Words Cannot Map the Development of the Sciences. *Journal of the American Society for Information Science*, vol. 48, no. 5, pp. 418-427.

Leydesdorff, L. (2011). "Meaning" as a sociological concept: A review of the modeling, mapping, and simulation of the communication of knowledge and meaning. *Social Science Information*, vol. 50, no. 3-4, pp. 391-413.

Leydesdorff, L. & Hellsten, I. (2005). Metaphors and Diaphors in Science Communication: Mapping the Case of Stem-Cell Research. *Science Communication*, vol. 27, no. 1, pp. 64-99.

Leydesdorff, L. & Hellsten, I. (2006). Measuring the meanings of words in contexts: automated analysis of 'Monarch butterflies', 'Frankenfoods' and 'stem cells'. *Scientometrics*, vol. 67, no. 2, pp. 231-258.

Martin, M.K., Pfeffer, J. & Carley, K. (2013). Network text analysis of conceptual overlap in interviews, newspaper articles and keywords, *Social Network Analysis and Mining*, vol. 3, no. 4, pp. 1165-1177.

Painter, J. (2014). *Climate Change in the Media. Reporting Risk and Uncertainty.*

University of Oxford, Reuters Institute for the Study of Journalism.

Pfeffer, J. & Carley, K. M. (2012). Rapid modeling and analyzing networks extracted from pre-structured news articles. *Computational Mathematic Organization Theory*, vol. 18, pp. 280-299.

Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, vol. 62, no. 2, pp. 406-418.

Turney, P.D. & Pantel, P. (2010). From Frequency to Meaning: Vector Space Models of Semantics. *Journal of Artificial Intelligence Research*, vol. 37, pp. 141-188.

Van Eck, N.J., & Waltman, L. (2011). Text mining and visualization using VOSviewer. *ISSI Newsletter*, vol. 7, no. 3, pp. 50-54.

Van Eck, N.J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, vol. 84, no. 2, pp. 523-538.

Van Eck, N.J., Waltman, L., Noyons, E.C.M., & Buter, R.K. (2010). Automatic term identification for bibliometric mapping. *Scientometrics*, vol. 82, no. 3, pp. 581–596.

Vlieger, E. & Leydesdorff, L. (2011). Content Analysis and the Measurement of Meaning: The Visualization of Frames of Collections of Messages, *The Public Journal of Semiotics*, vol. 3, no. 1, pp. 28-50.