



Data-driven analysis and prediction of norm acceptance

Ralf Krestel^{1,2} · Annegret Kuhn² · Wilhelm Hasselbring²

Accepted: 14 June 2022 / Published online: 20 July 2022
© The Author(s) 2022

Abstract

That norms matter for politics is a widely shared observation. Existing political science research on norm diffusion, norm localization, and contestations is, however, constrained due to methodological manageability of empirical data. To face this research challenge, we propose an interdisciplinary research collaboration between political and computer science. Using the show case of energy politics, we want to conduct unsupervised and semi-supervised content analysis and fusion with the help of automated text mining methods to analyze the influence of different types of so-called norm entrepreneurs on the public acceptance and, respectively, contestations of different energy policies.

Norm acceptance and its conflicts with interests

Our goal is to analyze major political narratives, as well as their discursive contestations with regard to different political topics of crucial societal relevance in order to explain, and in a long-term research perspective also make predictions about the societal (non-)acceptance of political discourses and norm diffusion. A (first) major focus of our empirical research endeavour will lie on the interplay of energy and environmental/climate governance. This encompasses questions of nuclear power politics, discussions about (offshore) wind energy and the framing of energy derived from gas plants. The project will analyse discourses about the “adequate” political conduct concerning energy governance within different national and international, formal and informal policy output, as well as public and media discourses. In a second step, we will compare the different narratives as well as derivation of inherent norms and interests, which contribute to explain public acceptance of

and compliance with specific policies on the one hand, and contestations of policies, on the other hand.

As information sources for cross-domain fusion, we analyze media outlets, transcribed interviews with stakeholders/broader population, and social media. We use natural language processing methods to identify and extract major political narratives and frames. In addition, we will use community detection and argument mining techniques together with information diffusion analysis to capture the various actors and opinions.

That norms matter for politics is a widely shared observation [19]. Questions of norm diffusion, localization and contestations have accordingly become a topic of growing relevance within political science research, e.g. [1, 3, 6, 18]. Following more recent norms research [19] we use a broader definition of norms, which can be defined as “ideas of varying degrees of abstraction and specification with respect to fundamental values, organizing principles or standardized procedures” [13, pp. 103-104]. To explain norm acceptance on the one hand and norm contestations on the other hand, the role of norm entrepreneurs can be considered of crucial relevance [17, pp. 36]. Norm entrepreneurs “work to persuade other agents to alter their behavior in accordance with the norm entrepreneur’s ideas of appropriate behavior”, [11, pp. 36]. Research on norm diffusion has generated highly relevant empirical insights on the influence of different types of norm entrepreneurs within international organizations, as well as intra-state norm entrepreneurs, which will play an important role within this research endeavour. Crucial types of individual norm entrepreneurs, who induce preference changes, may be political or religious leaders, but also social influencers or, for

✉ Ralf Krestel
r.krestel@zbw.eu

Annegret Kuhn
akuhn@politik.uni-kiel.de

Wilhelm Hasselbring
hasselbring@email.uni-kiel.de

¹ ZBW – Leibniz Information Centre for Economics, Kiel, Germany

² Christian-Albrechts-Universität zu Kiel, Kiel, Germany

example, celebrities [14]. However, major challenges exist within existing (political science) research; so far, mostly single case studies or small N comparisons have been performed due to the manageability of empirical data.

To alleviate this problem, we propose an interdisciplinary research collaboration between political science and computer science through unsupervised and semi-supervised content analysis and fusion with the help of automated text mining methods. The use of artificial intelligence methods would enable large-scale analysis of heterogeneous textual resources. Our vision is the identification of crucial norm “entrepreneurs”, “champions” and “influencers” who drive discussion and adoption of novel policies. To identify the various actors and in the long term make predictions about their influence, we will identify and model policy and norm acceptance and adoption dependent on origin or early adopters.

Case: energy and environmental governance

As highlighted above, we have chosen the topic of the interplay of energy politics and norms of environmental/climate protection as the empirical focus; foremost, because of its high practical relevance. Moreover, we find numerous cases of highly polarized contested discussions, e.g. [9], which make the topic a productive research agenda for the study of norm diffusion and contestation. To give an example, the case of nuclear energy can be scrutinized in a historical comparative perspective. The novel framing of nuclear energy as “green energy” fostering sustainability by the EU Commission and its acceptance or refutations within different countries of the European Union serves as a fruitful and timely empirical showcase for the detection of norms and interests in the political and public discourses of different groups of actors and the influence of different norm entrepreneurs – as well as the limits of the exertion of influence by norm entrepreneurs. In line with that, discourses of the EU Commission about gas energy as an environmentally friendly type of energy, as well as its contestations, could constitute another contemporary empirical focus. Contradicting the emerging “anti-fossil fuel norms” [4, 8], the success of its acceptance might be limited. On the other hand, frames about energy provision and the norm of national security/sovereignty, which have been increasingly proliferating in various European states since the Russian invasion of Ukraine, and might, in the short or even longer run, suppress normative concerns about sustainable energy politics. To put a specific focus on the influence of mainly locally acting norm entrepreneurs a further showcase could be the normative framing of (offshore) wind energy. Wind energy has become increasingly important in Germany and beyond. The expansion of wind energy production is ac-

companied by local contestations of different stakeholder groups [10]. Contestations do, however, differ highly with regard to their extent and severity, which can, at least partly, be traced back to the influence of local norm entrepreneurs fostering or opposing normative claims about the appropriateness of wind energy.

To analyze the diffusion/contestation of specific norms we use automatic/semi-supervised content analysis. Our analysis encompasses as a first step the mapping of relevant political and legal output by using multiple heterogeneous data sources (details on the different data sources are outlined in the following section).

We compare the different discourses to answer the following questions: What are the prevalent norms for fostering a specific policy? Can predominantly similar/differing norms be found within different discourses? The level of analysis would be: political output versus public discourse, as well as comparison of different subgroups of public discourses. Building on that comparison, in a next step, we identify the influence of different types of norm entrepreneurs on public norm acceptance and contestations, respectively. We will try to figure out which types of norm entrepreneurs are especially successful in shaping the public acceptance of specific energy policies. Through which channels do norm entrepreneurs mainly influence public normative considerations about energy policy, and what are the limits of the influence of norm entrepreneurs to shape public acceptance/contestation?

To do so, we combine two types of methods: the unsupervised and semi-supervised content analyses will be supplemented by network analysis (agent-based models) to consolidate our findings about the specific influence of different types of norm entrepreneurs.

Relevant data and information sources

Analyzing only one source of information, e.g. a particular party’s manifesto, captures only a limited societal view on complex problems. Norm acceptance is such a complex problem, which is influenced by many factors and dependent on a variety of interconnected parameters. Getting a more complete picture therefore requires the joint analysis of multiple sources and *fusing* them together to an integrated whole. To this end, many sources need to be integrated. Unfortunately, it is not enough just to gather every snippet of text related to a particular norm, because every source has its own rules and assumptions, as well as technical requirements and ways of communicating. Combining and integrating various sources poses a tremendous challenge to get substantiated results on norm acceptance and diffusion.



Fig. 1 Selected information and opinions on nuclear energy from different sources. (Sources: <https://eur-lex.europa.eu>, <https://dsriver.bundestag.de>, <https://www.spiegel.de>, <https://www.theguardian.com>, <https://twitter.com>)

With the goal of covering and gathering as much information as possible, the sources to investigate can be categorized into three groups:

1. Political and legal output
2. Mass media and public statements
3. Social media and individual data

Note that all three groups not only produce text but also audio-visual content. At the initial stage, we will focus on written text and defer the inclusion of raw videos and images to future work. Luckily, videos and images often contain metadata, captions, etc., which can be analyzed, and interviews and speeches are often transcribed, resulting in additional information sources that can be analyzed using text mining methods.

Please also note that the three types of sources do not coincide with individual entities, since they could be represented in multiple data sources. Former US President Donald Trump, e.g., was part of the government but communicated not only through official channels but made also extensive use of social media. The distinction of sources and roles of communication is not always clear, and norm entrepreneurs can also be found in different source media. Fig. 1 shows some example texts of interest for automatic analysis to paint an opinion landscape regarding the highly controversial topic of nuclear energy. In the following, we describe the potentially relevant and interesting data sources in more detail.

Under *political and legal output*, we subsume everything that is emitted by governing bodies, parties, administrative entities, etc. Examples are: EU regulations (e.g. the Taxonomy Complementary Climate Delegated Act (EU Commission 2022)), national laws and regulations, political speeches (e.g. in the EU and national parliaments), court rulings and other legal output.

Mass media and public statements, on the other hand, produce their own content but also functions as intermediaries between official communication and citizens [15].

With this, they fulfill a crucial task in opinion making within democratic systems. In the context of norm acceptance and contestations, major media outlets take on the role of norm entrepreneurs. With access to newspaper archives, this role can be analyzed, and the effect on norm acceptance quantified.

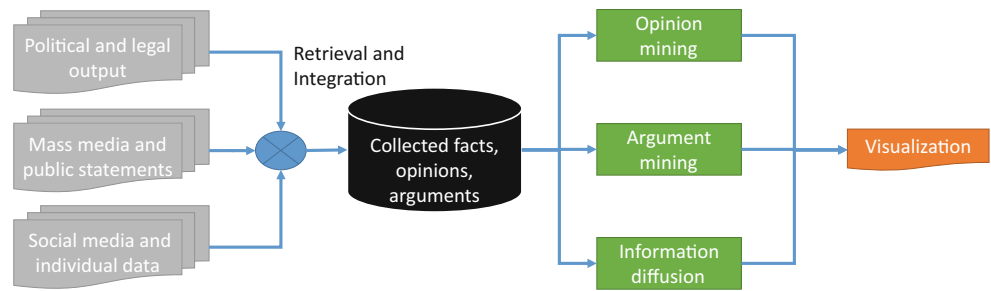
As a third group of sources, we can look at *opinions of individual citizens and social media*. Social media, such as Twitter and Facebook, are only one type of source, others include blogs and user-generated content, e.g. Wikipedia. Surveys and (transcribed) interviews are further sources of the “people’s opinion” that can be analyzed and integrated into the bigger picture.

These groups are not always easy to distinguish, e.g. politicians might use Twitter extensively to communicate (c.f. former US President Trump), and news outlets might have Facebook pages. It is therefore an additional challenge to identify the actual actors and their role within their communication network.

Cross-domain fusion

Modern text mining and content analysis methods can support political science research by enabling large-scale modelling of opinions and their diffusion, taking into account a plethora of sources. This is particularly interesting for investigating norm acceptance – a field, where large-scale studies have so far hardly been conducted due to the high costs of manual evaluation and interpretation and the difficulty of integrating multiple data sources. Traditionally, one model has been computed for one specific data source, e.g. there are studies analyzing opinions on Twitter [16] and how they are propagated through the Twitter network [2]. An especially challenging issue is the necessity of simultaneously dealing with data from various domains, as outlined in the previous section, and, instead of integrating po-

Fig. 2 Architectural overview of the different components of our proposed approach



tentially many individual models to get a holistic picture, perform *cross-domain fusion*.

A crucial step in an (semi-)automatic analysis of norm acceptance is, thus, the fusion of data across domains at different processing levels. It is obviously not enough just to collect the different data sources into one data silo, but the sources at varying stages of the natural language processing pipeline must be integrated. The whole process can be broken down into a sequence of processing steps and their associated computer science fields as follows:

- Finding and selecting sources and statements from a broad range of text types: information retrieval & data integration
- Knowledge extraction, opinion mining, and sentiment analysis: text mining
- Narratives and argument mining: computer linguistics
- Information diffusion: social network analysis
- Visualization and interpretation: human–computer interaction

Fig. 2 shows a coarse-grained architectural overview of the proposed system. The goal is to automate this pipeline as much as possible, while allowing the political scientist to interfere and guide the process based on expert knowledge.

Retrieval and integration

The stance of an individual or a group on norms manifests itself in various outlets: policy reports, contracts, political agreements, laws, news, but also in more informal ways, such as through social media, speeches, (transcribed) interviews, etc. Automatically extracting relevant statements from these different data sources is very challenging. Further, sources need to be assessed with regard to trustworthiness and authenticity. Also, the attribution of statements to a specific group or entity is often not easy. Especially in the political domain, underlying norms are not always visible in concrete statements or advocated solutions and decisions. Automatic detection, therefore, needs to have a common understanding of social and political backgrounds, e.g. in the form of a knowledge graph or other background models. In this first step, the focus is on selecting the statements and specific sources for further processing and analysis.

Interwoven with this step is the integration of many heterogeneous data types and information from the political domain into a unified representation. The identification of actors and attribution of statements needs to happen at this stage as well. The fact that individual actors usually use more than one medium to communicate makes it necessary to identify and disambiguate actors to establish some kind of profile.

Opinion mining

To populate profiles of political actors (individuals, parties, media companies, decision makers, etc.) opinions and stances need to be extracted. Opinion mining and sentiment analysis are well-researched areas when it comes to, e.g. product reviews. In the political domain, opinion mining is much more difficult [12]. This has two reasons: firstly, political opinions are typically very complex and not just binary (good or bad) but a lot more nuanced. Further, stances on a particular political issue are not independent of each other. Norms are related and dependent on each other, and it is not possible to negate one norm without also changing the stance towards other dependent norms without making a logical error. Secondly, political actors are often inclined to obfuscate their true stance or at least keep it vague. This makes opinion mining in the political domain much more challenging and, therefore, requires a deeper understanding of the utterances of political actors.

Argument mining

In the political sphere, simple sentiment analysis methods typically fail. Except for explicit voting on a well-defined issue or answering specific, clearly defined questions, norm acceptance and contestations are typically expressed in longer discussions. Here, methods from argument mining can be employed to narrow down the political views and normative stances. For complicated issues, extracting narratives around a certain norm could help to represent the various stances [5]. Methods from representation learning as well as automatic summarization could be employed to position political actors in a kind of latent argument space regarding norms and political views. Progress in embed-

ding methods and language models [7] are also promising to capture complex semantic information.

Information diffusion

Many political actors and also many citizens use social media to get information on the one hand and to spread their opinions on the other hand. This has led to powerful influencers with millions of followers and groups with high reach. Social network analysis, or more specifically, information diffusion research can identify highly influential accounts and track the spread of information through a social network. It is of paramount interest to monitor political actors and interest groups to analyze norm acceptance and how they position themselves with regard to particular political topics. A large-scale analysis of these communication structures might uncover latent relations and influence among various groups.

Visualization

All those computational models, representations, and analyses are of limited use without a proper way for political scientists to explore and investigate them. Therefore, finding a good way to visualize results backed up by their grounding sources and a clear depiction of communication structures are vital. Interpretation of the automatically generated results by political scientists is necessary, and capabilities to interact with the results, for example by browsing or exploring the representations and aggregated results, need to be supported. Even with the most advanced NLP tools, human experts need to draw the right conclusions from the data and put the results not only into the political context but also into the socio-economic context and consider other factors that might be relevant for norm acceptance. To this end, retrieval functionality needs to be provided to make the resulting model and its representation accessible.

Spatio-temporal dimension

Large-scale processing and the fusion of multiple types of sources opens up new possibilities for norm acceptance research beyond a mere increase in data volume. The global structure of information networks and the advancement in automatic translation make transferring results from one political region to others possible. Further, comparing regions and analyzing their influence on each other becomes possible. Within a globalized world, not only goods and money are no longer bound to country borders, but also opinions and norm acceptance are globally influenced. Besides the spatial analysis, the temporal aspect of changing norm acceptance over time can be investigated on a much finer level using cross-domain fusion. Especially the in-

terplay between different norms and how acceptance and contestation of particular norms lead to changing stances on other topics can over time become a new field of study for political science.

Outlook: prediction and explanation

With a good model for the influence of norm entrepreneurs on norm acceptance in place, we hope to be able to perform predictions on how novel norms will be accepted in the future. Further, predicting future norm entrepreneurs, unexpected coalitions regarding stances towards a norm, or the development of a discussion regarding a norm, might be possible beyond the field of energy politics as well.

A major challenge for our research endeavour is factors influencing the role of norm entrepreneurs with regard to acceptance/contestation of political discourse and norms, which cannot be gathered through text/discourse analysis. As described above, we try to capture the influence of these factors through a methodological triangulation by complementing our research with a network analysis. Still, further (potential) explanatory factors might stay unobserved and will have to be gathered through complementary methodological approaches.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Acharya A (2004) How ideas spread: whose norms matter? Norm localization and institutional change in asian regionalism. *Int Org* 58(2):239–275
2. Bakshy E, Hofman JM, Mason WA, Watts DJ (2011) Everyone's an influencer: quantifying influence on twitter. In: *Proceedings of the fourth ACM international conference on Web search and data mining*, pp 65–74
3. Bernstein S (2004) Legitimacy in global environmental governance. *J Int Law Int Relat* 1(1-2):140
4. Blondeel M, Colgan J, Van de Graaf T (2019) What drives norm success? Evidence from anti-fossil fuel campaigns. *Glob Environ Polit* 19(4):63–84

5. Campos R, Jorge A, Jatowt A, Bhatia S, Finlayson M (2021) The 4th international workshop on narrative extraction from texts: Text2story 2021. In: European Conference on Information Retrieval. Springer, Berlin Heidelberg, pp 701–704
6. Deitelhoff N, Wolf KD (2013) Business and human rights: How corporate norm violators become norm entrepreneurs. In: The persistent power of human rights: from commitment to compliance, pp 222–238
7. Devlin J, Chang M, Lee K, Toutanova K (2019) BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT). Association for Computational Linguistics, pp 4171–4186 <https://doi.org/10.18653/v1/n19-1423>
8. Green F (2018) Anti-fossil fuel norms. *Clim Change* 150(1):103–116
9. Hensengerth O (2015) Global norms in domestic politics: environmental norm contestation in cambodia's hydropower sector. *Pac Rev* 28(4):505–528
10. Hildebrand J, Renn O (2019) Akzeptanz in der energiewende. In: *Energiewende*. Springer, Berlin Heidelberg, pp 261–282
11. Hoffmann MJ (2003) Entrepreneurs and norm dynamics: An agent-based model of the norm life cycle. www.polisci.upenn.edu/ps-i/Pamla/Hoffmann_norms.doc. Accessed 18 July 2022
12. Krestel R, Wall A, Nejd W (2012) Treehugger or petrolhead?: identifying bias by comparing online news articles with political speeches. In: Proceedings of the 21st World Wide Web Conference, WWW 2012 Lyon. ACM, pp 547–548 <https://doi.org/10.1145/2187980.2188120>
13. Krook ML, True J (2012) Rethinking the life cycles of international norms: the united nations and the global promotion of gender equality. *Eur J Int Relat* 18(1):103–127
14. Legros S, Cislighi B (2020) Mapping the social-norms literature: an overview of reviews. *Perspect Psychol Sci* 15(1):62–80
15. Michailidou A, Trez HJ (2017) The media as public intermediaries of knowledge in Europe: from deliberation to democratic legitimation, 1st edn. Routledge, pp 171–191
16. Pak A, Paroubek P (2010) Twitter as a corpus for sentiment analysis and opinion mining. In: Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)
17. Payne RA (2001) Persuasion, frames and norm construction. *Eur J Int Relat* 7(1):37–61
18. Risse T, Risse-Kappen T, Ropp SC, Sikkink K (1999) The power of human rights: International norms and domestic change vol 66. Cambridge University Press
19. Wiener A (2017) Constitution and contestation of norms in global governance: A cycle model. Presentation at the University of Cambridge, Cambridge. vol 8

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.