

SYMPOSIUM

estimating interest groups' policy positions through content analysis: a discussion of automated and human-coding text analysis techniques applied to studies of EU lobbying

adriana bunea^{a,*} and *raimondas ibenskas*^b

^aDepartment of Political Science, University College London, 31 Tavistock Square, London WC1H 9QU, UK

E-mail: a.bunea@ucl.ac.uk

^bDepartment of Politics and International Relations, University of Southampton, Highfield Campus, University Road, Southampton SO17 1BJ, UK

E-mail: R.Ibenskas@soton.ac.uk

*Corresponding author.

doi:10.1057/eps.2016.15; published online 27 May 2016

Abstract

The promises and pitfalls of automated (computer-assisted) and human-coding content analysis techniques applied to political science research have been extensively discussed in the scholarship on party politics and legislative studies. This study presents a similar comparative analysis outlining the pay-offs and trade-offs of these two methods of content analysis applied to research on EU lobbying. The empirical focus is on estimating interest groups' positions based on their formally submitted policy position documents in the context of EU policymaking. We identify the defining characteristics of these documents and argue that the choice for a method of content analysis should be informed by a concern for addressing the specificities of the research topic covered, of the research question asked and of the data sources employed. We discuss the key analytical assumptions and methodological requirements of automated and

human-coding text analysis and the degree to which they match the identified text characteristics. We critically assess the most relevant methodological challenges research designs face when these requirements need to be complied with and how these challenges might affect measurement validity. We also compare the two approaches in terms of their reliability and resource intensity. The article concludes with recommendations and issues for future research.

Keywords EU lobbying; policy position estimates; text analysis; automated and human-coding techniques

Content analysis is widely recognised in social sciences as a well-established 'research technique for making replicable and valid inferences from texts [...] to the contexts of their use' (Krippendorff, 2004: 18). In recent years, political science has witnessed a considerable development of different automated (computer-assisted) content analysis techniques, applied mainly to the field of legislative studies and party politics research (Laver *et al*, 2003; Slapin and Proksch, 2008; Diermeier *et al*, 2012). These techniques treat words as data and infer ideological or policy positions of political actors based on words frequency in a political text. Recently, quantitative text analysis was also applied to EU lobbying research to examine interest groups' policy positions and lobbying success (Klüver, 2009, 2013) or their framing strategies (Boräng *et al*, 2014; Klüver and Mahoney, 2015; Klüver *et al*, 2015). By 'automated text analysis' we refer to supervised and unsupervised text-scaling algorithms conducted with the help of statistical softwares,¹ whereas by 'human-coding' we refer to text analysis that relies on human coders for extracting, categorising and quantifying the information of interest.² While both approaches use numerical values to express the variables of interest (e.g., policy positions), the process through which these numbers are generated

differs: one method relies on computer codes and text-scaling algorithms, while the other employs human judgement for the interpretation of text and the assignment of relevant text blocks to pre-established conceptual categories and coded variables that are clearly stated in a coding protocol (codebook).

The promises and pitfalls of automated and human-coding content analysis techniques applied to examine political texts have been extensively discussed in the scholarship on party politics and legislative studies (Benoit and Laver, 2007; Budge and Pennings, 2007; Lowe and Benoit, 2013). In this article, we conduct a similar comparative analysis outlining the pay-offs and trade-offs of automated and human-coding content analysis in the context of research on EU lobbying. Our article is motivated by Grimmer and Stewart's (2013) pertinent observation that automated methods 'are no substitute for careful thought and close reading and require extensive and problem-specific validation' (267). We build on our previous research that provided a comparative assessment of the main unsupervised text-scaling algorithm used in the scholarship on EU lobbying (namely *Wordfish*, see Klüver, 2009) with the human-coding method we have employed to estimate interest groups' positions expressed in the European Commission's consultation on the reduction of CO₂

emissions from passenger cars (2007) (see Bunea and Ibenskas, 2015). Our previous study presented a more elaborated methodological discussion supported by statistical evidence of automated and human-coding techniques applied to EU lobbying research. With the present article, we aim to address a broader audience by providing a more general and less technical discussion of key methodological and substantive points to be considered by scholars of EU lobbying.

Our discussion fits well the context of a fast developing yet still maturing empirical literature on EU lobbying and interest groups that presents important opportunities for methodological innovation while also running the temptation to indiscriminately borrow analytical tools from across a broad spectrum of fields of research (Bunea and Baumgartner, 2014). We propose a discussion of the conditions under which such 'methodological transfers' can suitably take place. We limit our discussion to the issue of estimating groups' positions based on their formally submitted position documents. The justification for this is twofold. First, EU open consultations are a relevant data source in the study of EU lobbying strategies and policy influence (Klüver, 2013; Bunea, 2014; Rasmussen and Carroll, 2014). It is therefore important to identify the most efficient and effective methods to analyse this valuable data source. Second, we decided to trade width for depth in our analysis and focus on one automated text analysis method only. We leave for future research the task of critically assessing other computer-assisted text analysis techniques that were applied to other key fields of EU lobbying research such as, for example, interest groups' frames of policy issues and political debates (see Boräng *et al*, 2014).

We first identify the defining characteristics of interest groups' position documents in the EU policymaking. We contend that the choice for a method of

'... the choice for a method of content analysis should be informed by a concern for addressing the specificities of the research topic covered, of the research question asked and of the data sources employed'.

content analysis should be informed by a concern for addressing the specificities of the research topic covered, of the research question asked and of the data sources employed (in line with Roberts, 2000). We then discuss the key analytical assumptions and methodological requirements of automated and human-coding text analysis and the degree to which they match these text characteristics. We critically assess the most relevant methodological challenges research designs face when these requirements need to be complied with and how these challenges might affect measurement validity (see also Eising, 2016). We also compare the two approaches in terms of their reliability and resource intensity. We conclude with recommendations and issues for future research. We place a greater analytical focus on discussing the automated text analysis because of its novelty in the literature and because its complex methodological assumptions make it less intuitive and easy to grasp for non-methodologists.

INTEREST GROUPS' POLICY POSITION DOCUMENTS IN THE CONTEXT OF EU POLICYMAKING

Interest groups' policy positions documents represent a valuable data source.

These texts were used, for example, to research lobbying in the context of US federal bureaucratic rule-making (Nelson and Yackee, 2012), and judicial and legislative politics (Caldeira and Wright, 1990; Evans, 1996). In the literature on European lobbying, they were employed to study interest mobilisation, lobbying strategies and lobbying success at the national (Rasmussen, 2015) and EU level (Klüver, 2013; Bunea, 2013; Rasmussen and Alexandrova, 2012). In most studies, researchers employed human-coding content analysis to estimate groups' positions and levels of lobbying success (see Eising *et al*, 2015), although automated approaches were also applied (Klüver, 2009).

In the context of the EU system of governance, interest groups' position documents present a set of salient characteristics that differentiates them in some fundamental ways from ideological texts, such as party manifestos or legislative speeches. We consider these characteristics to be key in informing a researcher's decision on what content analysis technique best fits his/her research design. We observe thus the following characteristics:

- (1) The *substantive content of texts* is technical in nature: these documents usually communicate technical and factual information, corresponding to what the literature describes as a technocratic policy-making system in which expert knowledge and policy-specific information are key when private actors interact with policy-makers and express their opinions on concrete policy matters. The level of technicality varies across texts, according to the degree of expert knowledge and specialisation an organisation possesses. Two relevant observable implications follow from this. First, these texts refer to complex policy realities and provide

'... interest groups' position documents present a set of salient characteristics that differentiates them in some fundamental ways from ideological texts such as party manifestos or legislative speeches'.

information that usually corresponds to *several issues* and *policy dimensions*. Second, the information conveyed with the help of *numerical values* can have an important substantive meaning, and it sometimes represents the key instrument for discerning between policy positions.

- (2) *The structure* of documents allows a relatively easy and straightforward identification of *policy issues* in relation to which organisations express positions. Generally, these documents follow closely the structure of the Commission's consultation call that clearly identifies the policy issues on which stakeholders' positions are requested. This differs from ideological texts, especially party manifestos, whose authors are less constrained with regard to the issues they address in texts.
- (3) *The intended purpose of communication* can vary across texts, although all documents are formally addressed to the European Commission. Whereas some texts are used by organisations to affect policy outcomes by formulating specific demands and transmitting technical, specific, policy-relevant information, other texts do not express a specific position but consist instead of more general statements about a policy area or aspects of EU decision making.

- (4) Their *authorship* varies greatly in terms of types of organisational actors formulating them (i.e., business, environmental NGOs, local authorities, Euro-federations, national associations, individual organisations). This also implies that the level of texts' technicality varies according to the degree of expert knowledge and specialisation that an organisation possesses in a policy area.
- (5) Their *text format* is not uniform and it can vary in terms of length, terminology, writing style and language. Some documents are longer than others, some use more technical terms while others adopt a laymen vocabulary.
- (6) Policy documents are often written in several *different languages* in accordance with the legal framework allowing the use of any of the EU twenty-four official languages when participating in the policy-making process.
- (7) The relevant *unit of analysis* for these texts varies across documents and can range from one sentence to several paragraphs. A qualitative inspection of these documents reveals that organisations use anything in between one or several sentences to one or several text paragraphs to express their position on one policy issue.

Having identified these defining characteristics of policy position documents, we now turn and discuss how they affect the applicability of automated and human-based content analysis.

CONTENT ANALYSIS: AN OVERVIEW

AUTOMATED TEXT ANALYSIS

Content analysis techniques are classified into two broad approaches: automated

and human based. Automated content analysis uses computers to either classify the content of texts into specific topics or scale them to extract actors' positions. In contrast, the human-coding approach is based entirely on 'the use of people as coders, with each using a standard codebook and coding form to read, view, or otherwise decode the target content and recode his or her objective and careful observations on pre-established variables' (Neuendorf, 2002: 52). Both classes of content analysis are applied to address various research questions and imply the use of different analytical techniques. One of their most prominent use concerns the measurement of political actors' positions, a fundamental quest for the study of politics (Benoit and Laver, 2006: 14).

While we recognise that there is a wide range of both supervised and unsupervised text-scaling algorithms that are employed for conducting automated text analysis (Slapin and Proksch, 2014: 129–30; Skalski, 2002; Alexa and Zuell, 2000), we review only one particular method, namely *Wordfish*, for reasons outlined in the introduction. Before we show how its methodological rigours match the characteristics of groups' position documents, we briefly introduce this method.

Wordfish is a text-scaling algorithm that uses word frequencies to place documents within a uni-dimensional policy space (Slapin and Proksch, 2008). The method assumes that word frequencies follow a Poisson distribution, which has a single parameter (λ) representing both the mean and the variance. The functional form of the model is:

$$y_{ij} \sim \text{Poisson}(\lambda_{ij})$$

$$\lambda_{ij} = \exp(\alpha_i + \psi_j + \beta_j^* \omega_i),$$

where y_{ij} is the frequency of word j in the document of the actor i , α is a set of document fixed effects, ψ is a set of word

fixed effects, β is the estimate of a word-specific weight that captures the importance of the word j in discriminating between actors' positions and ω is an estimate of the position of the actor i . By having fixed effects for actors and words, the method accounts for the possibility that some actors have longer documents and some words are used more often than others by all actors. The model is estimated using an expectation maximisation algorithm.

Wordfish provides essential *methodological opportunities* that hold across fields of research: high measurement reliability, time efficiency, compatibility with conducting large- n research and comparatively lower levels of intensity of data collection efforts relative to approaches relying on human-coding.³ Relative to other text-scaling algorithms (e.g., *Wordscores*), *Wordfish* has more straightforward underlying assumptions and it does not require any rescaling method or raw scores. It is therefore a relatively user-friendly and transparent algorithm. Most importantly, *Wordfish* does not require the existence of two reference texts expressing and marking the most extreme positions in relation to the analysed dimension (as required by *Wordscores*). This is particularly relevant in the field of EU lobbying because such reference texts that 'can be estimated with confidence from independent sources or assumed uncontroversially' (Laver *et al*, 2003: 313) are not available. Whereas for party manifestos, the external validation of the reference texts was done with the help of expert surveys, no such data sources are currently available on interest groups' policy position documents.

These are all noteworthy opportunities. Yet, the empirical application of this method also raises a set of research design challenges that we discuss in the next section. We consider each challenge to be of equal importance. Assessing empirically the relative impact of each

methodological trade-off on the measurement validity of position estimates is a complex task that goes beyond the purpose of the present article.

HUMAN-CODING

Several human-coding techniques have been applied to the measurement of political actors' policy positions, with the Comparative Manifestos Project being perhaps the best-known project. In the literature on lobbying no such massive data-coding project exists but several human-coding techniques were used (see Yackee and Yackee, 2006; Eising *et al*, 2015; Vannoni, 2016). The methodological assumptions of human-coding text analysis are less strict and elaborate. They consist mainly of three fundamental principles: (1) the coding of text needs to be based on a codebook detailing all steps of the coding process, the content of interest, the unit of analysis, the coding format and the categories of variables to be coded based on the analysed texts; (2) the training of coders on how to systematically apply this codebook; and (3) conducting and reporting an inter-coder reliability test to evaluate the reliability of the coding scheme (Hayes and Krippendorff, 2007). This approach has no specific strict requirements regarding the quality of analysed texts in terms of their comparability, authorship, format, informative content or the use of numerical values to convey a message (see also Voltolini, 2016). These aspects are all taken into account before the codebook is developed and then addressed in the content of the coding protocol as the researcher finds necessary.

In this study we build on the human-coding we applied in our previous research (Bunea and Ibenskas, 2015). To code groups' positions, we first examined the EU official documents associated with the examined consultation to identify the policy issues on which organisations

adopted positions. We then identified the positions of organisations on each issue based on their position documents by first reading the text sections/paragraphs providing the relevant information, and then by coding the position based on one (or two) key sentences that explicitly expressed it. Each position was coded in the data set. Our approach thus differs from the classical approach that usually implies a systematic coding of all sentences in a paragraph/sub-section followed by their distribution into positive and negative categories and the analysis of the aggregate distribution of these positive and negative values (see the Comparative Manifestos Project).

MEASUREMENT VALIDITY: METHODOLOGICAL ASSUMPTIONS AND CHALLENGES

While it is generally agreed that automated content analysis outperforms human-coding in terms of reliability and resource intensity, the scholarship still debates about how the two approaches perform in terms of their validity (Slapin and Proksch, 2014: 138). Automated content analysis is based on a number of assumptions, whose violation may affect the validity of its results. Slapin and Proksch identify the following methodological assumptions that need to be satisfied when applying automated content analysis:

(1) The method assumes that the employed algorithm captures the *dimensionality of the policy space* analysed appropriately. As text-scaling is applied to capture an underlying, latent variable that cannot be directly observed or measured (i.e., ideology, policy position), having a well-defined prior knowledge and clear definition/identification of what

'Automated content analysis is based on a number of assumptions, whose violation may affect the validity of its results'.

dimension(s) is/are estimated with the applied algorithm is a strong methodological imperative.

- (2) The method assumes that the analysed texts *provide information* that is directly linked to and relevant for the underlying policy dimension(s) that is examined with the text-scaling algorithm.
- (3) An automated approach requires that all analysed texts are similar (and thus comparable) in terms of their authorship, text generation process, targeted audience and communication purpose (e.g., convey a political ideology to the electorate, state a position in a political debate). The texts should use a similar terminology and be generated within similar institutional/organisational settings by similar types of actors (e.g., political parties, MPs). The documents must be written in the *same language*. This is a crucial assumption since the estimates of computerised content analysis are based on the relative frequencies with which different words are used across texts. Word frequencies have a substantive meaning and allow distinguishing the substantive differences between texts in relation to the underlying, latent variable that is researched.
- (4) *Only words are data* and their frequency in a text provides substantive information about the latent variable analysed. The observable implication of this assumption is that substantive information conveyed with the help of numerical values, graphs or figures

is usually removed from the analysed text, or if kept, it is not given appropriate weight in the estimation algorithm because usually such information is not frequently repeated in the text.

Below, we discuss the extent to which the specific characteristics of interest groups' policy documents match these methodological assumptions and how this could create several methodological challenges. We also discuss the extent to which human-coding can address these potential challenges.

CAPTURING APPROPRIATELY THE DIMENSIONALITY OF THE EU POLICY SPACE

A key aspect that may affect the validity of the results of automated text analysis is whether the number of dimensions identified to describe the political/policy space is appropriate and does not discard relevant information (Benoit and Laver, 2012: 199). Text-scaling methods extract the positions of actors in a low- (usually single-) dimensional space, although the existing methods differ with regard to how many dimensions they derive. The *Wordfish* method places actors in a single-dimensional space. The creators of *Wordfish* make a very explicit argument in this respect: 'first define the dimensions ex ante and, second, use only documents that contain information relevant to that dimension. Defining the dimension implies being transparent about what information is being used' (Slapin and Proksch, 2008: 712). Current applications of *Wordfish* to analyse groups' position documents follow this methodological requirement by assuming that the entire text of a position document provides information about one policy dimension only: '[s]ince all documents discuss only the Commission initiative for reducing CO₂ emissions from cars, one can assume

uni-dimensionality and, thus, the complete texts were used for the analysis' (Klüver, 2009: 541).

We argue that generally the requirement for a uni-dimensional space can raise a serious methodological challenge when analysing interest groups' position documents because these texts usually provide information about a policy space that can have (and usually does have) more than a single dimension. These texts provide information about a *large number of issues* and they are *authored by diverse organisations*. While a large number of issues does not necessarily imply a multi-dimensional space, the differences between stakeholders in terms of their interest type, organisational form and national origin usually result in a complex 'bundling' of policy issues that then translates into a multi-dimensional policy space. In addition, the practical realities of EU policymaking suggest that policy events revolve around several issues that are treated by decision makers individually and not collapsed into one policy dimension. A brief reading of different Commission's calls for consultations reveals that in their policy practice, the European bureaucrats design detailed consultation documents in which they formulate specific questions aimed at asking for stakeholders' policy feedback on several, specific and distinct issues.

The scholarship has recently developed multi-dimensional text-scaling algorithms (Diermier *et al*, 2012), and one such approach was applied in the scholarship on interest groups to examine policy frames (Klüver and Mahoney, 2015). Nevertheless, prior knowledge about the number and substantive meaning of policy dimensions is important when interpreting the results of these analyses (cf. Benoit and Laver, 2012). For example, as one prominent characteristic of interest groups' position documents is that they provide *information with the help of technical terminology*, one cannot assume

that the underlying space that structures groups' positions is based on the classic dimensions of 'left-right', 'liberal-conservative', or 'pro- anti-EU'. Instead, a good knowledge of the dimensions specific to the examined legislative proposal is required.

Human-coding is more flexible in this respect as it allows identifying the positions of stakeholders on specific issues, which can then be used to examine and interpret the dimensionality of the policy space using multi-dimensional data analysis techniques. For example, we identified ten policy issues corresponding to the consultation on the reduction of CO₂ emissions from passenger cars based on the questions asked by the Commission in the consultation document and groups' position documents. With the help of Specific Multiple Correspondence Analysis we found that these issues correspond to two policy dimensions, each describing a different regulatory regime aimed at improving environmental standards by regulating car producers or consumers' behaviours (Bunea and Ibenskas, 2015).

INFORMATION CONTENT

A second challenge in applying automated content analysis to interest groups' position documents is that these may include information that is not directly linked to and relevant for the underlying policy dimension(s). This results from two characteristics of these documents: their *technical terminology* and the *variation in the intended purpose of communication*.

First, the use of technical terminology, numbers and figures in interest groups' policy documents means that, different from ideological texts such as party manifestos or legislative speeches, the frequency with which a word is used does not differentiate between actors' positions on the latent policy dimension(s). These differences are instead marked by using key technical terms or numerical values,

which are not given appropriate weight in the estimation algorithm because usually such information is not frequently repeated in the text. The automated content analysis of such technical documents is therefore unlikely to uncover the substantive policy differences between positions, and would instead capture differences based on other grounds such as different styles of writing. Moreover, applying automated content analysis usually requires the removal of all numeric values even if they play a key role in differentiating between interest groups' positions.

Second, the variation in the purpose of communication is also problematic because it means that some documents (or parts of them) do not provide information related to specific issues and are therefore uninformative for uncovering policy positions. This observation supports the argument that a successful application of automated text analysis needs to be embedded in a thorough reading of texts: '[i]ndeed a deep understanding of the texts is one of the key advantages of the social scientist in applying automated methods' (Grimmer and Stewart, 2013: 270).

In contrast, human-coding can effectively deal with both challenges. First, human coders are able to differentiate between actors' positions on the basis of both key words and numerical values. In the consultation on CO₂ emissions, one key issue was the time frame for reaching the emissions reduction target. Two substantially different positions were expressed: pro-environment organisations asked for 2012 to be the deadline for reaching the target, while car producers advocated for this deadline to be 2015. Numerical values were key in identifying the positions differentiating between them. The removal of numbers from texts (that is usually required in automated text analysis) would have prevented in this case an accurate identification of positions.

Second, human-coders are also able to discard irrelevant information or uninformative documents when coding interest groups' positions. For example, several organisations that participated in the consultation on CO₂ emissions did not express any specific positions in their documents. Thus, one could reasonably argue that their position documents were not informative with respect to the underlying policy dimension(s) analysed (Bunea and Ibenskas, 2015).

THE TEXT-GENERATING PROCESS

Another challenge in applying automated content analysis to study interest groups' policy documents arises because the process through which their content is generated varies substantially. These differences result from the variation in the *authorship* and the *intended purpose of communication* of these documents. As already mentioned, a very diverse set of organisations are involved in EU policy-making. They differ in terms of organisational settings and lobbying capabilities, and benefit to different degrees from the presence of staff members specialised in the formulation of policy position documents. This instead affects the level of technicality of their documents. Differences in the text-generating processes also imply that these texts are not comparable units of analysis amenable to text-scaling algorithms. An additional challenge arises when trying to integrate in one analysis documents generated by private and public actors that have substantially different organisational environments. This can constitute a relevant challenge when attempting to estimate lobbying success by looking at which positions (expressed in position documents) are translated into policy outcomes (stated in the legislative proposals or final legislative acts). By their very nature, groups' position documents and EU official texts are not comparable texts generated

'... a successful application of automated text analysis needs to be embedded in a thorough reading of texts'.

within similar institutional settings, which implies that they cannot be included in the same automated text analysis.

Also, organisations use policy documents for different communication purposes. Some employ them to transmit expert knowledge and specialised information as part of their attempts to exert policy influence. Others use them to indicate their stakeholder status, to contest the legitimacy of the initiated policy initiative or to show their constituency that they are performing their EU representational mandate. Therefore, interest organisations differ from each other in more fundamental ways than parties or MPs that represent among themselves a more homogenous group of political actors, that generate political texts in more similar institutional environments and are subjected to similar 'institutional constraints' (Proksch and Slapin, 2012).

Human-coding content analysis does not have any specific requirements regarding the text-generating process and is thus better equipped to deal with documents that were generated by different types of actors in different organisational settings.

VARIATION IN THE LINGUISTIC REGIME

A further challenge to applications of automated text analysis is posed by the assumption that the analysed texts are comparable. This aspect is key in the context of EU policymaking because in this system of governance policy position documents are written in different languages. The observable implication of this

is that text-scaling algorithms are applied to texts written in one language only (most often English). Texts written in any other languages must be excluded from the analysis since text-scaling algorithms require a unique linguistic regime.⁴ This implies a trade-off between discarding some texts and implicitly reducing the number of observations, and not conducting the text-scaling analysis at all. Discarding data points can however potentially bias estimates of the population of organisations lobbying within a certain policy/consultation event, estimates of the size of lobbying sides and coalitions, and affect subsequent causal inferences made about the aggregate levels of lobbying success estimated for both individual organisations and lobbying sides within an event or lobbying venue. In contrast, human-coding can address the issue of documents' diverse linguistic regime by allowing the development of codebooks in different languages.

WORDS AS UNIT OF ANALYSIS

Lastly, the very basic assumption that words are the most appropriate unit of analysis for investigating the content of position documents is challenging. The scholarship theorising content analysis recommends that in a research design the relevant unit of analysis should be decided based on the context of research: '[t]he key in selecting a unit of analysis is not to assume that one's population of text is comprised a priori of clearly-distinguishable text-blocks. On the contrary, it is the researcher's responsibility to divide this population into blocks – blocks that can be uniquely identified according to the contextual variables required for addressing the research question at hand' (Roberts, 2000: 268). To pay attention to the context of research focusing on EU policy documents means to take into account their technical nature. For such technical documents, the frequency of

words is less informative in terms of estimating a policy position since this frequency does not have an ideological meaning. Instead, considering that the relevant unit of analysis can range between one or several key sentences that express the policy positions to one/several paragraphs that convey this position constitutes a more reasonable approach. Some organisations express their position in a very succinct manner with the help of one or two clauses, while others employed one or two full paragraphs to specify their preference. Deciding the most appropriate unit of analysis in this context emphasises again the absolute importance for an in-depth knowledge of the texts.

RELIABILITY AND RESOURCE INTENSITY

Automated content analysis performs well in terms of assuring high measurement reliability. On the basis of the same sample of documents and applying the same scaling algorithm, one usually gets the same estimates of interest groups' positions across several different measurements. There is one caveat though: these estimates depend on the words used in the text-scaling analysis. What counts as an 'informative' word and what is deleted because is considered 'uninformative'⁵ is decided by the researcher, and therefore it is potentially subject to change and can vary across research designs. This could potentially affect the reliability measure of positions estimates.

In contrast, human-coding is more prone to systematic coding error (Mikhaylov *et al*, 2012). Its successful application depends to a large extent on the quality and intensity of the training received by human coders (Neuendorf, 2002). This aspect is particularly relevant and potentially challenging in the context of research on EU lobbying and

policymaking, which, as already mentioned, is rather technical and therefore requires a solid and thorough knowledge on behalf of coders of the analysed policy events. At the same time, the reliability of human-coding may be improved as policy issues are relatively well defined in the consultation documents of the European Commission.

Automated content analysis also performs better in terms of resource intensity. While the preparation of documents for analysis and mastering the methods of computerised content analysis may be quite labour intensive, content analysis based on human coders requires the deployment of much more substantial human and financial resources. The methodological implication of this is that human-based text analysis can be less compatible with conducting large-*n* research designs. The analysis of only a selected number of cases from the entire universe of possible cases means that the reliability of human-based content analysis may be lower. This makes case selection a key element in any research design relying on human-coding and requires a clear and well-crafted case-selection criteria, alongside an open discussion and recognition of the limits to generalisability of research findings.

CONCLUSION

A solid content analysis research design requires a good match between the assumptions of the analytical method and the characteristics of the texts that constitute its object of analysis (Roberts, 2000). In the EU context, interest groups' policy position documents present a set of text characteristics that challenge the ability of the existing methods of automated text analysis to uncover valid policy positions of interest groups expressed in these texts. These documents tend to use technical language; they refer to

multiple issues; they are authored by diverse organisations that use these documents for different communication purposes; and they are often written in several languages. Human-based content analysis is better equipped to account for these text characteristics, and therefore it represents a recommended method for studying these documents even if it requires more resources and its results can be less reliable. We consider that this explains perhaps why European Commission's policymakers themselves adopt this approach when examining and analysing interest groups' contributions to public consultations instead of opting for automated analyses.

Our discussion suggests several recommendations for the substantive and methodological research on lobbying in the EU. First, scholars who consider using automated content analysis should examine the extent to which the aforementioned characteristics are present in the particular sample of texts they want to examine. We argue that the use of automated content analysis techniques should be considered as a feasible methodological option only when the analysed policy documents are less technical in their substantive content and terminology used, the number of policy issues is low, the organisations authoring them are relatively homogenous in their organisational characteristics and they use their documents for the same goals (e.g., to express their policy positions on the issues related to a legislative proposal), and only when all (or almost all) documents are written in the same language.

More systematic research is needed to establish the variation of these text characteristics and conditions under which automated techniques can be applied across different policy areas and consultation events. The present analysis drew extensively on the consultation on one environmental consultation to identify the characteristics of interest groups'

position documents. A systematic analysis of a large and representative number of consultations may find that some of these methodological issues are more present than others. Second, future research should also examine how much each of the aforementioned text characteristics affect the validity of position estimates derived from automated content analysis. Finally, building on the insights about which of these document features are most present and provide the most important challenges to the validity of the estimates of interest groups' positions, the scholars should upgrade the existing techniques and develop new methods of automated content analysis to account for these characteristics. Nevertheless, even when such methods are developed, human-based content analysis will remain an important tool in the study of open consultations in the European Union. As Grimmer and Stewart (2013: 270) suggest, '[r]ather than replace humans, computers amplify human abilities. The most productive line of inquiry, therefore, is not in identifying how automated methods can obviate the need for researchers to read their text. Rather, the most productive line of inquiry is to identify the best way to use both humans and automated methods for analyzing texts'.

'A solid content analysis research design requires a good match between the assumptions of the analytical method and the characteristics of the texts that constitute its object of analysis'.

Acknowledgements

The authors are grateful to Helene Helboe Pedersen and Rainer Eising for their excellent comments on previous versions of this study, as well as to the participants in the 2014 ECPR workshop on methods to study interest groups, for their great insights and distinguished company. They would also like to thank Jon Slapin and Ken Benoit for introducing them to the world of automated text analysis. Both authors gratefully acknowledge the financial support received from the European Commission's Marie Skłodowska-Curie Actions Programme through their individual Intra-European Fellowships while writing this article (A. Bunea: Grant no. 622661; R. Ibenskas: Grant no. 330446).

Notes

1 See Grimmer and Stewart (2013: 268) for an excellent overview of automated content analysis methods for political texts. Supervised and unsupervised text-scaling algorithms are the two main approaches used for ideological scaling.

2 This distinction differs thus from the classical dichotomy of qualitative versus quantitative content analysis (Krippendorff, 2004: 87–98). Our analysis is based on the fundamental assumption that '[a] content analysis has as its goal a numerically based summary of a chosen message set. It is neither a gestalt impression nor a fully detailed description of a message or a message set' (Krippendorff, 2004: 87–89).

3 We note though that the application of automated content analysis techniques usually requires a careful preparation of analysed texts in terms of removing uninformative words, numbers, figures and punctuation marks. This step can also be labour intensive and relies exclusively on the efforts of researchers.

4 See however the very recent attempts made to develop 'automated multilingual content analysis techniques' by Proksch *et al.* (2015).

5 Slapin and Proksch (2014: 137) refer to these as 'stopwords', that is, words that have no ideological content such as 'prepositions and conjunctions'.

References

- Alexa, M. and Zuell, C. (2000) 'Text analysis software: Commonalities, differences and limitations: The results of a review', *Quality and Quantity* 34(3): 299–321.
- Benoit, K. and Laver, M. (2006) *Party Policy in Modern Democracies*, London: Routledge.
- Benoit, K. and Laver, M. (2007) 'Estimating party policy positions: Comparing expert surveys and hand-coded content analysis', *Electoral Studies* 26(1): 90–107.
- Benoit, K. and Laver, M. (2012) 'The Dimensionality of Political Space: Epistemological and Methodological Considerations', *European Union Politics* 13(2): 194–218.
- Boräng, F., Eising, R., Klüver, H., Mahoney, C., Naurin, D., Rasch, D. and Rozbicka, P. (2014) 'Identifying frame: A comparison of research methods', *Interest Groups and Advocacy* 3(2): 188–201.
- Budge, I. and Pennings, P. (2007) 'Do they work? Validating computerized word frequency estimates against policy series', *Electoral Studies* 26(1): 121–129.
- Bunea, A. (2013) 'Issues, preferences and ties: Determinants of EU interest groups' preference attainment in the environmental policy area', *Journal of European Public Policy* 20(4): 552–570.
- Bunea, A. (2014) 'Explaining interest groups' articulation of policy preferences in the European commission's open consultations: An analysis of the environmental policy area', *Journal of Common Market Studies* 52(6): 1224–1241.
- Bunea, A. and Baumgartner, F.R. (2014) 'The state of the discipline: Authorship, research designs, and citation patterns in studies of EU interest groups and lobbying', *Journal of European Public Policy* 21(10): 1412–1434.
- Bunea, A. and Ibenskas, R. (2015) 'Quantitative text analysis and the study of EU lobbying and interest groups', *European Union Politics* 16(3): 429–455.
- Caldeira, G.A. and Wright, J.R. (1990) 'Amici Curiae before the supreme court: Who participates, when and how much?' *Journal of Politics* 52(3): 782–806.
- Diermeier, D., Godbout, J.F., Bei, Y. and Kaufmann, S. (2012) 'Language and ideology in congress', *British Journal of Political Science* 42(1): 31–55.
- Eising, R. (2016, this issue of *European Political Science*) 'Studying interest groups: Methodological challenges and tools', *European Political Science*, doi: 10.1057/eps.2016.14.
- Eising, R., Rasch, D. and Rozbicka, P. (2015) 'Institutions, policies and arguments: Framing activities in the EU multi-level system', *Journal of European Public Policy* 22(4): 516–533.
- Evans, D. (1996) 'Before the roll-call: Interest group lobbying and public policy outcomes in house', *Political Research Quarterly* 49(2): 287–304.
- Grimmer, J. and 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.
- Hayes, A.F. and Krippendorff, K. (2007) 'Answering the call for a standard reliability measure for coding data', *Communication Methods and Measures* 1(1): 77–89.
- Klüver, H. (2009) 'Measuring interest group influence using quantitative text analysis', *European Union Politics* 10(4): 535–549.
- Klüver, H. (2013) *Lobbying in the European Union: Interest Groups, Lobbying Coalitions and Policy Change*, Oxford: Oxford University Press.
- Klüver, H. and Mahoney, C. (2015) 'Measuring interest group framing strategies in public policy debates', *Journal of Public Policy* 35(2): 223–244.
- Klüver, H., Mahoney, C. and Opper, M. (2015) 'Framing in context: How interest groups employ framing to lobby the European commission', *Journal of European Public Policy* 22(4): 481–498.
- Krippendorff, K. (2004) *Content Analysis: An Introduction to Its Methodology*, Thousand Oaks, CA; London: SAGE Publications.
- Laver, M., Benoit, K. and Garry, J. (2003) 'Extracting policy positions from political texts using words as data', *American Political Science Review* 97(2): 311–331.
- Lowe, W. and Benoit, K. (2013) 'Validating estimates of latent traits from textual data using human judgment as a benchmark', *Political Analysis* 21(3): 298–313.
- Mikhaylov, S., Laver, M. and Benoit, K. (2012) 'Coder reliability and misclassification in the human coding of party manifestos', *Political Analysis* 20(1): 78–91.

- Nelson, D. and Yackee, S.W. (2012) 'Lobbying coalitions and government policy change: An analysis of federal rulemaking', *Journal of Politics* 74(2): 339–553.
- Neuendorf, K.A. (2002) *The Content Analysis Guidebook*, London: SAGE Publications.
- Proksch, S., Lowe, W. and Soroka, S. (2015) 'Automated multilingual content analysis: A new approach to estimating conflict from parliamentary speeches across languages', Paper presented at the fifth annual conference of the *European Political Science Association*, Vienna, 25–27 June.
- Proksch, S.O. and Slapin, J.B. (2012) 'Institutional foundations of legislative speech', *American Journal of Political Science* 56(3): 520–537.
- Rasmussen, A. (2015) 'Participation in written government consultations in Denmark and the United Kingdom: System and actor-level effect', *Government and Opposition* 50(2): 271–299.
- Rasmussen, A. and Alexandrova, P. (2012) 'Foreign interests lobbying Brussels: Participation of non-EU members in commission consultations', *Journal of Common Market Studies* 50(4): 614–631.
- Rasmussen, A. and Carroll, B.J. (2014) 'Determinants of upper-class dominance in the heavenly chorus: Lessons from European union online consultations', *British Journal of Political Science* 44(2): 445–459.
- Roberts, C.W. (2000) 'A conceptual framework for quantitative text analysis on joining probabilities and substantive inferences about texts', *Quality and Quantity* 34(3): 259–274.
- Skalski, P.D. (2002) 'Computer Content Analysis Software', in K.A. Neuendorf (ed.) *The Content Analysis Guidebook*, London: SAGE Publications, pp. 225–235.
- Slapin, J.B. and Proksch, S.O. (2008) 'A scaling model of for estimating time-series party positions from texts', *American Journal of Political Science* 52(3): 705–722.
- Slapin, J.B. and Proksch, S.O. (2014) 'Words as Data: Content Analysis in Legislative Studies', in S. Martin, T. Saalfeld and K.W. Strom (eds.) *Oxford Handbook of Legislative Studies*, Oxford: Oxford University Press, pp. 126–144.
- Vannoni, M. (2016, this issue of *European Political Science*) 'Studying preference attainment using spatial models', *European Political Science*, doi: 10.1057/eps.2016.13.
- Voltoini, B. (2016, this issue of *European Political Science*) 'Framing processes and lobbying in EU foreign policy: Case study and process tracing methods', *European Political Science*, doi: 10.1057/eps.2016.18.
- Yackee, J.W. and Yackee, S.W. (2006) 'A bias towards business? Assessing interest group influence on the US bureaucracy', *Journal of Politics* 68(1): 128–139.

About the Authors

Adriana Bunea is a Marie Curie Research Fellow at University College London, Department of Political Science. Her published work examines different aspects of EU lobbying and the European Commission's stakeholders' consultation regime.

Raimondas Ibenskas is a Lecturer in Politics and International Relations at the University of Southampton. His published work examines the causes and consequences of party change across European countries.

comment on – estimating interest groups’ policy positions through content analysis: a discussion of automated and human-coding text analysis techniques applied to studies of EU lobbying

anne s. binderkrantz

Department of Political Science, Aarhus University, Bartholins Allé 7,
Aarhus C DK-8000, Denmark
E-mail: ASB@ps.au.dk

In recent years, scholars of interest groups have increasingly turned to the use of automated text analysis. This has proved useful in mapping the policy positions of groups, the different frames of communication used by groups and also in establishing the extent to which groups are successful in obtaining their political goals. The use of automated text analysis has raised optimism about the options for doing large *n*-studies with reasonable demands for time consumption and high reliability. It has also opened new avenues for approaching the crucial question of gauging the relative success of different groups. According to the contribution by Bunea and Ibenskas in this special issue, this optimism is – at least to some extent – misguided. On the basis of their review of the advantages and disadvantages of different methods for text analysis they conclude that: the

documents typically used in automated text analyses have a range of characteristics that inhibits the use of the method in that they: ‘tend to use technical language; they refer to multiple issues; they are authored by diverse organizations that use these documents for communication purposes; and they are often written in several languages’.

This conclusion is – at least in my reading – akin to saying that interest group scholars should turn to other methods in answering questions about interest group politics and leave the use of automated text analysis to scholars working in other parts of the discipline. After all, can we imagine any plausible context within the interest group field with no diversity in types of organisations participating or a possibility that multiple issues are covered in the same documents? And if we can, surely such settings are not the most

crucial for mapping the political role of interest groups. Restricting the use of automated text analysis to situations, where none of these contextual problems are present, will therefore mean that the potential for using the methods are miniscule. In my opinion, dismissing the use of automated text analysis will, however, be a premature conclusion. Pointing to limitations does not mean that these cannot conceivably be met. Neither does it necessarily mean that good scholarship cannot be conducted even in the presence of limitations. For example, because there are limitations to all approaches to studying interest group influence, the use of a range of different approaches in scholarly work is crucial to advance our collective understanding of the subject.

A fruitful strategy for future work drawing on automated text analysis would – in line with the final remarks made by Bunea and Ibenskas – focus on assessing the relative importance of these obstacles in different situations and on possible ways to overcome them. At least two main avenues may be pursued: the utilisation of increasingly sophisticated methods for automated text analysis and

clever combinations of automated and human coding. First, the potential of computer-assisted text analysis has clearly not been exhausted yet. For example, existing software packages may be utilised in mapping multiple dimensions in interest group positions as expressed in policy documents. Second, automated text analysis can never fully be a substitute for human coding. Although computers are highly reliable in grouping similar texts, humans remain crucial for validating the interpretation of such groupings – or for engaging in supervision of the coding process. A crucial concern is, for example, what similarity or difference in words across different documents signifies. Similarity may indicate that documents discuss the same issues, that they convey the same policy positions, use the same framing or seek to target the same audience. Future research may thus focus on clever ways of combining the strengths of human and automated coding as well as on the development of more sophisticated methods of analysis. What cannot be avoided in scholarship using automated coding is careful attention to the pitfalls that have now explicitly been laid out.

About the Author

Anne Skorkjær Binderkrantz is Professor in the Department of Political Science, Aarhus University. Her main research interests are interest groups, political elites and public governance. Her work has appeared in the *British Journal of Political Science*, *European Political Science Review* and *Governance*.

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.