Session B3

New Methods for TA - Computational Content Analysis: Experiences, Potentials and Challenges

Chairs: Titus Udrea, Daniela Fuchs, Leo Capari (Institute of Technology Assessment), Anja Bauer (Alpen-Adria-University Klagenfurt)

In order to advice politics and society on emerging technologies, their societal implications and governance, Technology Assessment practitioners frequently summarize the state of research on particular topics, reconstruct political, media and societal discourses, and continuously monitor scientific and technological developments. All these activities involve the analysis of large amounts of content, be it scientific articles, policy documents, newspaper articles, social media contributions, films or images. With this session, we aim to explore whether and how Computational Content Analysis (CA) offers methodological innovations for TA studies.

Computational CA and related methods are among the fastest expanding class of methods currently in use for textual data analysis and other forms of communication. Examples include: data/text mining, topic modelling, network analysis and bibliographic research, knowledge discovery and mapping, as well as other tools focused on extracting patterns and information from large (textual) databases. The digital age is making increasingly large amounts (and types) of data available for analysis and opening new research avenues. Using information retrieval, statistics, computational science, machine learning and database technology, Computational CA includes a strong technology emphasis. From this perspective, uncovering the practical roles these methods and tools play for TA researchers is valuable.

This session seeks to inspire knowledge exchange and dialogue between researchers that use Computational Content Analysis (CA) in the context of TA studies (or related fields). More specifically, as a methodology focused session, the aim is to elaborate and compare different empirical applications of Computational CA and discuss best practices (for example in regard to reliability, validation, accuracy, etc.). We invite presentations that reflect on experiences made with specific methods and tools of Computational Content Analysis, for example with a view on the following topics and questions:

- Literature review: How may Computational CA facilitate systematic literature reviews in TA? What role can methods such as topic modelling, network analysis or bibliographic analysis play in assessing scientific fields and discourses?
- Discourse Analysis: Can Computational CA methods and tools advance quantitative and qualitative discourse analysis? What experiences have been made in the analysis of media discourses (newspaper articles) or the analysis of Social Media (e.g. Twitter communication)? What are advantages and disadvantages compared to other approaches of discourse analysis?
- Visual content: What methods and tools are useful for the analysis of visual content (such as images)? What opportunities and challenges arise for TA studies from the automated analysis of visual content?
- Methodological and practical challenges: What (novel) methodological and practical challenges accompany the use of Computational CA in TA studies? How do practices of data collection, analysis and interpretation change? How can reliability, validity and accuracy of analysis be secured?
- Combination with other TA methods: How can Computational CA be usefully combined with other TA methods such as participatory methods or case study analysis? Which challenges arise from such mixed methods approaches?

Using Automated Text Analysis Methods to Uncover Socio-Political Narratives within Computer Modelling and Simulations

Authors: Titus Udrea, Daniela Fuchs, Leo Capari (Institute of Technology Assessment), Anja Bauer (Alpen-Adria-University Klagenfurt)

Academic literature on computer modelling and simulation has broadened significantly along with the uses of modelling tools within the policy-making process. As such, academic research output using computer modelling is intended to develop, as well as assess the impacts and potential consequences of decisions within complex policy contexts. Analysing the ever-increasing amount of computer modelling publications and assessing their policy narratives presents particular challenges.

In our paper we present a methodological framework that comprises two approaches of automated content analysis to scientific articles using different computer modelling and simulation exercises. In order to analyse the modelling and simulation research literature our empirical analysis includes three corpora constructed from abstracts of scientific articles within three distinct policy areas: energy systems modelling in the context of energy transition, the modelling of the impacts of trade policies and the modelling of nanotechnology risks.

First, we use an automated text mining technique to analyse and create a word co-occurrence map in order to explore the relationships between the main research topics and modelling clusters. Secondly, an unsupervised semantic-based text structure and relationship analysis with a latent Dirichlet allocation (LDA) topic model is used to uncover (sub)topics and specific computer modelling and simulation approaches.

This exploratory analysis first uncovers the relations between computer modelling and simulation applications and policy topics. The results also explore the relations between specific computer modelling tools/applications and (sub)topics within the socio-political narratives embedded within these specific academic publications. In our presentation we provide insights into strategies of constructing highly specialized corpora of scientific articles. We also discuss the opportunities and challenges of both automated text analysis techniques used and illustrate how they can be meaningfully combined to uncover the semantic context of texts.

A Policy Research Method Case Study: Generating and Extracting Evidence-based Policy Inferences from a large EC Framework Programme Project

Authors: Tanja Sinozica (Institute of Technology Assessment), Mete Basar Baypinarb (Istanbul Technical University), Edward Bergmanc (WU Vienna University of Economics and Business), Miklos Hornyakd, Ferenc Kruzslicze and Attila Vargaf (University of Pecs)

In 2004 the European Neighbourhood Policy (ENP) was instituted following the greatest single enlargement of the European Union (EU), to support security and peaceful relations between the EU and neighbouring countries with a unified governance approach to economic, social and political aspects of international cooperation. This paper reports on an effort to develop and test a methodology for bridging social science research and policy communities on an important policy question that concerns comity between the EU and its Eurasian, Middle Eastern and North African neighbouring countries, although the approach applies to any broad policy issue for which multiple sources and types of research evidence are present. Five evaluative elements are developed and implemented whose complementary application result in a large set of policy inferences, a strategy of implementation, and researcher insights concerning the method. This case study suggests that the recommended evidence synthesis methodology has good potential for informing policy that is comprised of multiple elements, studied by large research teams, and enacted by diverse agents. The suggested methodology requires engagement by active researchers and policy experts in the

formulation of policy options. It is put forward that improving the quality of evidence-informed policy will depend upon institutions and practices in the research and policy making communities.

Merge your data! – Mixing methods of digital content, discourse, and network analysis

Author: Nils Matzner (Alpen-Adria-University Klagenfurt)

For getting a grip on massive amounts of computer assessable data, new methodologies have been developed, such as computational discourse analysis, digital methods and data science. Critique of these methods often focusses on too simplistic online newspaper word clouds or tool-generated Twitter networks. To overcome these shortcomings, social scientists and TA practitioners may on the one hand combine various digital tools and capabilities of programming languages (such as R), and on the other hand complement these findings with qualitative social research. This paper will highlight the advantages and challenges of merging data and their analysis from various sources. While doing so, it has to be taken into account that "follow the methods of the medium" with internet and other digital data one has to understand "digital objects" (Rogers 2013) as a novel entity, that requires methodological innovation. Furthermore, it will be argue that a sophisticated mix of methods for computational data may unveil links between various arenas (e.g. online media and social media) as well as actors (e.g. journalists and influencers). The paper presents examples from emerging climate engineering discourses and networks. The examples show that terminologies in climate engineering research and policy are strongly arena-dependent.