



VALTIOTIETEELLINEN YHDISTYS  
STATSVETENSKAPLIGA FÖRENINGEN  
FINNISH POLITICAL SCIENCE ASSOCIATION

## **2. Digitalization of research objects and research methodologies in political science: What should, could and should not be done?**

### **Puheenjohtajat:**

Pertti Ahonen, [pertti.ahonen@helsinki.fi](mailto:pertti.ahonen@helsinki.fi)

Matti Nelimarkka, [matti.nelimarkka@hiit.fi](mailto:matti.nelimarkka@hiit.fi)

Mari Marttila, [mari.marttila@hiit.fi](mailto:mari.marttila@hiit.fi)

### **Työryhmän kuvaus:**

In late 1990s and early 2000s, the notion of information society was used in examining and framing the impact that information technology had on society. Today, the notion of digitalization is used in reflecting upon the same topic, but technologies have greatly evolved and matured. Commonly referenced examples of digitalization include changes in employee-employer relations (e.g. Uber, Wolt), automation and platformization (e.g. Airbnb, Facebook). Digitalization has also influenced the overall society, e.g., the role of education institutions and easily accessible education through massive open online courses (MOOCs). In this workshop, we ask what are the impacts of digitalization upon political science. In particular, our interests relate to two dimensions:

Methodologically, digitalization influences data collection and analysis. The field of computational social science takes advantage of novel data sources or new methods that can be used for data analysis. Novel data sources refer to (often online) digital trace data provided by users, such as Twitter or mobile phones as trackers. From an analytical perspective, the question is how we can use algorithmic tools, such as machine learning or simulation, to examine social phenomena. There is an emerging trend to apply these tools in political science research. Moreover, there are signs that empirical political science research that has calmly flown in the academic mainstream, such as the examination of texts of political party programs, increasingly takes into account the possibilities that computational methods offer.

New political phenomena also open up for empirical researchers to examine. First, digitalization creates new avenues to speak and act, importantly various social media platforms. These avenues have interested political scientists for some time, such as political campaigns and protests. Another direction for empirical research would be to apply science and technology studies framework. In that area, tools and environments are seen as constructed and the question is why certain choices have been made in the construction process.

In this workshop, we invite papers that address the question “What should, could, and should not political scientists study given digitalization and by means that include digital methodologies?”.

- The “shoulds” can arise, for instance, from pressing concerns to tap new digital methods to examine social media discourse concerning immigration, refugees, terrorist acts and frozen or ongoing political and military conflicts, and to draw inferences concerning networks of actors under examination.
- The “coulds” derive, for instance, from the enabling capacity of the digital methodologies to examine data in substantially larger proportions than entrenched methods, and to engage in new kinds of quantitative and quantitative-assisted qualitative examination.
- Finally, the “should nots” are situated where political scientists have to seriously consider to which types of political power they render services on the one hand and on the other to which political emancipatory efforts they dedicate themselves.

Each paper should examine a case that illustrates the actual (realized and proven) or potential (possibly and promising) influence of digitalization upon theoretical, methodological or empirical aspects of political science research. Those whose papers are accepted for presentation should deliver a full manuscript not later than two weeks before the conference, meaning a deadline of 25 February 2016. We prefer English contributions for the inclusion of other than Finnish-speakers, but we also accept proposals for papers in Finnish or Swedish.

## **Abstraktit:**

### **Määrällisen ja laadullisen tutkimusotteen yhdistäminen Verohallinnossa**

Verohallinnon sisäinen kehittämistoiminta pyrkii muun muassa helpottamaan asiakkaiden asiointiprosesseja. Konkreettisten kehittämiskohteiden tunnistaminen vaatii asiakkaiden kohtaamien, asiointiin liittyvien hankaluuksien ymmärtämistä. Määrällisten ja laadullisten lähestymistapojen yhdistäminen on tässä suhteessa osoittautunut varsin toimivaksi lähtökohdaksi. Verohallinnossa käytetään sekä määrällisiä että laadullisia analyysejä rinnakkain samojen ongelmien syiden löytämiseksi. Nämä eri lähestymistavat yhdistyvät suoraan muun muassa tekstinlouhinnan menetelmien kautta. Eräs hedelmälliseksi osoittautunut tekstinlouhinnan menetelmä on structural topic modeling (STM). Verohallinnossa on kehitetty R-kieleen pohjautuvia työkaluja laajojen tekstiaineistojen automaattisen analysoinnin helpottamiseksi. Esittelen tuloksia veronmaksajien ja Verohallinnon virkailijoiden välisen kirjeenvaihdon analysoinnista STM-menetelmällä.

Jarno Tuimala, Verohallinto, jarno.tuimala@vero.fi

## **Svenska folkpartiet och vänster-högerskalan, 1906-2016: En kvantitativ analys**

Cross-temporal and cross-sectional research on political parties and their policy positions has typically relied on data derived from content analyses of electoral manifestos (i.e. the MRG/CMP/MARPOR data) or from expert surveys (e.g. the CHES data). These approaches have, however, recently been challenged by automated content analytical techniques such as, e.g., WORDFISH or WORDSCORES. This paper shows how WORDFISH, a scaling algorithm to estimate policy positions based on word frequencies in texts (Slapin & Proksch 2008), can be used when studying policy positions of Finnish parties. The main focus is on the Swedish People's Party in Finland (Svenska folkpartiet i Finland; SFP) and its positions on the left-right dimension. The general purpose with the paper is to complement recent studies (Sundberg 2006; Isaksson 2011) on the left-right placement of SFP by examining party platforms as well as electoral manifestos, by focusing on socioeconomic as well as sociocultural aspects of the left-right dimension, and by comparing the positions of SFP with the positions of the leading parties in the Finnish party system. (If the paper is accepted: paper in Swedish, abstract + oral presentation in English.)

Andreas Fagerholm, Åbo Akademi, andreas.fagerholm@abo.fi

## **A Topic Modeling Approach to the Coproduction of Political Communication: the Case of Romania's 2014 Presidential Elections**

The aim of this paper is to examine how the coproduction of online-based political communication as a form of citizen engagement translates into offline participation. I do so by analyzing the case of Romania's 2014 presidential elections, where two relevant trends emerged: firstly, social media debuted as an alternative and highly impactful campaign tool. Secondly, social media became a space for documenting and disseminating information on irregularities in the voting process, and a channel for organizing large-scale street protests in response to these following the first round of elections. Between rounds, the typically active/passive campaign roles of candidates and voters were reversed. Voters became active suppliers of a new set of salient issues. In turn, candidates Klaus Iohannis and Victor Ponta took a reactive position in the runoff and developed their own response strategies. In this study I apply topic modelling to three types of sources: candidates' electoral programmes, text posts on their respective Facebook campaign pages, and a corpus compiled from public comments posted by users on these materials between rounds. In the latter case, I consider only the textual content of comments. User specific meta-data is not collected due to ethical concerns. The purpose of the analysis is threefold. Firstly, I examine candidates' electoral programmes to identify the themes and issues they proposed during the active stage of their involvement in the communication process. Secondly, I analyze user comments to determine which issues were imposed by voters as active communicators, and how they were expressed. Finally, I examine how issues identified at this point translated into candidates' campaign materials between rounds, and the degree to which their original themes persisted or were replaced by the new ones.

Laura Elena Sibinescu, University of Helsinki, Department of Political and Economic Studies, laura.sibinescu@helsinki.fi

## **Suomi 24 - Muodonantoa aineiston sisällöille: Citizen Mindscapes -hankkeen haasteita/in Finnish or in English**

Esityksen aiheena ovat otsikon ilmaiset haasteet erittäin laajan aineiston hankinnassa rahoitettua tutkimushanketta varten.

Krista Lagus, University of Helsinki, Consumer Research Center, krista.lagus@helsinki.fi

## **Negotiating the Analog Mainstream with Digital Methods in Hand Visions from the Examination of Party Programs and Government Political Programs**

Using ongoing research on Finnish party programs and government "master" political programs, the presenter asks, how scholars turned digital could negotiate the research mainstream.

Pertti Ahonen, University of Helsinki, Department of Political and Economic Studies, pertti.ahonen@helsinki.fi

## **Natural language processing in policy analysis – machine-learning the structure of EU energy policy**

Policy analysis relies on a mapping of relevant policy instruments and related practices. Modern policy produces vast amounts of text, including roadmaps, policy briefs, background papers - in amounts not accessible to a human reader. This constrains our ability to identify sources of influence, trajectories of policy ideas, and the overall big picture.

This paper explores the use of machine learning methodology and natural language processing for understanding the structure of a policy sector, represented as a big data corpus of policy documents. The method is topic modelling, where each document consists of multiple topics, represented as word probability distributions. The method is unsupervised: the topics are unknown, and the success of the analysis hinges on whether the topics make sense to a human reader – whether they are semantically coherent. The paper discusses the relationship between social scientific theories (discourses, frames, institutions) and the results of such modelling.

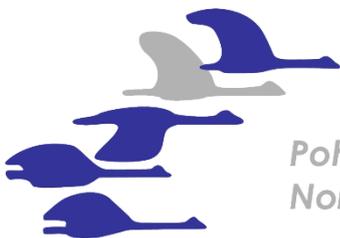
This paper analyses the structures of European Union energy policy, using a policy corpus of 615 documents and 4.5 million words. These documents were published by EU officials, but are not necessarily official documents. The documents were published between 2005 and 2015 and represent a snapshot of what EU energy policy is means now.

Arho Toikka, University of Helsinki, Department of Political and Economic Studies, arho.toikka@helsinki.fi

## **An automated Discourse Quality Index (DQI) classifier to measure deliberation automatically**

Political scientists have recently studied deliberation in decision-making processes. One approach in this research has been to classify contributions based on the deliberative characteristics on that work. This approach is however limited due to the resources related to properly conduct this classification, e.g. using the Discourse Quality Index (DQI). However, recently social scientists have explored how machine learning algorithms can be used to support in classifications. This work presents how these can be used to explore deliberation. Our results show, that conducting analysis in DQI approach is difficult and leads to poor accuracy in the prediction. However, we present an alternative, reduced DQI model and observe much better accuracy.

Matti Nelimarkka, University of Helsinki/Aalto University/HIIT, [matti.nelimarkka@hiit.fi](mailto:matti.nelimarkka@hiit.fi)



*Pohjoismaiden hallinnollinen liitto Suomen osasto ry*  
*Nordiska administrativa förbundet Finska avdelningen rf*