## Session B1

### “Limits to … predictive power” – policy advice in the age of computational modelling

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

“Computer statt Politik” (computers instead of politics) - already in the early 1970s the German magazine “Der Spiegel” questioned governmental efforts to rationally predict, plan and govern by the means of computational power. Since then, the role of computational modelling and simulation (CMS) in informing and guiding policies and politics has steadily increased. Technical improvements - both with regard to data basis (for example via more accurate measuring methods) as well as enhanced computing power allow for the development and calculation of ever more complex models. To name just a few examples, computer models are used in risk governance to determine the toxicity of chemicals; computer simulations inform political strategies for energy transitions on EU, national and regional levels, and computer models suggest potential effects of preferential trade agreements between countries. Yet, CMS in policy advice are not uncontested: The most prominent example might be the accusation of the politicization of IPCC climate models. Moreover, questions of interpretability and informative values remain challenging for those directly involved in advising policy makers or societal actors. Recently, debates on computational modelling in policy advice have gained new impetus by advances in Artificial Intelligence (AI) and Big Data. Based on abundant and just-in time data, learning machines are tasked with developing models for the prediction and management of a variety of situations and decisions, be it the occurrences of crimes or terrorist attacks, the most efficient use of social assistance, or successful strategies in foreign policy. Such developments fortify concerns about algorithmic biases and the risks of autonomous decision-making for democracies.

For Technology Assessment, computational modelling in policy advice is of interest in two ways: First, computational modelling, big data and artificial intelligence are objects of analysis in TA, relating to the question how the increasing digitization affects knowledge production and advisory systems and, in a broader view, societal decision-making and democracy. Second, TA practitioners are themselves users, and in some cases even developers, of computer models and simulations. For this session, we invite presentations that reflect on either of the two dimensions, CMS as an object of TA analysis or CMS as a tool of TA analysis, along the following themes and questions:

- (Hidden) assumptions, values and biases in models and algorithms: What kind of and whose assumptions and values are (intentionally or unintentionally) inscribed in computational modelling approaches – and therefore may rather remain unchallenged in broader discourses - and how? What biases are produced or maintained by models and algorithms and how can they be disclosed?
- Models at the interface between science (or specifically TA) and politics: How are model assumptions and results negotiated and communicated between different actors (scientists, stakeholders, policy-makers)? What kind of challenges do arise for TA when using modelling approaches, simulation results or data analytics in policy advice? How can accountability and transparency of policy advice and decision-making processes be ensured in the view of ever more complex models and algorithms?
- Models between democracy and technocracy: What opportunities and barriers does CMS provide for democratic deliberations of technology visions? In what way are different actor groups and their perspectives included in modelling practices? How is such openness with regard to actors mirrored in respective discourses around CMS? Do new developments in AI, Big Data and machine learning foster technocratic tendencies in policy advice?

## The socio-political narratives of computer models: Insights from the modelling of trade policies, energy systems and the risks of nanotechnologies

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

Computer models have become increasingly prominent technologies to support public policies on a wide range of complex and societally controversial questions. Models fulfil a variety of functions, from the identification and analysis of societal problems to the examination of different policy instruments and the assessment of the impacts and costs of planned and implemented policies, yet they also serve to justify and legitimize public action. In this role, models are not neutral tools, simply providing orientation and answers to (exogenously given) societal or political questions but have performative effects for socio-political discourses. Modellers adopt and reframe policy questions, they make relevance decisions regarding parameters, factors and scenarios to include, and, in some cases, derive recommendations and options for societal and political actions.

In our presentation, we explore and discuss the (re)production of socio-political visions and narratives in computational modelling and simulation in three areas: energy system modelling in the context of energy transition, the modelling of the impacts of trade policies and the modelling of nanotechnology risks. Methodologically, we build on quantitative and qualitative analysis of scientific articles presenting modelling and simulation exercises and results. By the means of bibliometric analysis and text mining approaches we explore the main subjects and socio-political semantics of modelling in the three areas. We identify, for example, geographic priorities, technological foci, and dominant policy discourses (e.g. climate change, risk or free trade). Our findings further illustrate the entanglement of epistemic and social order by indicating that specific modelling approaches are strongly interlinked with specific social and political issues.

We deepen this overview analysis with a qualitative analysis of selected key articles from each area. The qualitative analysis allows insights in how policy issues and missions are adopted and transformed in modelling exercises, how societal visions and ideas find their ways into models and scenarios and how ultimately modelling results are used to support or challenge policy narratives. We find quite distinct patterns across and within the three areas, from rather implicit and abstract reference to existing social and political discourses (for example the risk assessment discourse in the context of nanomaterials) to differentiated strategies of boundary work (for example when adopting normative missions in the context of energy transition but refraining from clear support of particular policy options) to deriving explicitly normative claims and recommendations from modelling and simulations. Overall, with our analysis of scientific modelling discourses, we illustrate how the 'politics of models' does not only concern their use at the science-policy interface, but is already inscribed in their development, application and scientific exploitation.

## Ambivalences of legibilization. Computational social science and its effects on policy advice

*Author: Florian Eyert (Weizenbaum Institute for the Networked Society)*

Computational modeling and simulation (CMS) plays an increasingly important role in processes of social ordering, from individual court decisions to environmental policy. With the advent of "algorithmic governance" such models are diffusing from the inner circle of science to broader societal applications. A central, but so far understudied part of this development is the rise of computational social science (CSS) as a computational modeling tool in both public policy and private management. The contribution proposes to not understand CMS and CSS as entirely new phenomena, but to take into consideration their historic continuitity by conceptualizing them as projects of legibilization (Scott 1998): attempts to render social processes intelligible through

simplification and quantification with the goal of "government at a distance" (Miller and Rose 1990). Against this background it is then possible to trace the epistemic shifts that do occur with the emergence of CSS, such as a new focus on prediction (Hofman, Sharma, and Watts 2017), a new empiricism and inductive probabilism and a new interdisciplinary epistemic culture.

The rise of CSS models in policy advice is characterized by ambivalences. On the one hand, they seem to foster new optimism in possibilities for steering and planning in society due to a new level of model complexity and even enable new modes of critique (Ensign et al. 2018). On the other hand, rather than merely solving technical problems, computational models open up new risks concerning data protection (Pohle 2012), bias, epistemic opacity (Humphreys 2009) and disregard for situational and qualitative forms of knowledge. The paper will contrast two approaches to CSS in order to analyze how they deal with these ambiguities: machine-learning-based techniques (e.g. in predictive policing (Seo et al. 2018)) and agent-based simulation (e.g. in migration policy (Simon et al. 2018)). Both are contemporary methods used in CSS, but in some respects they constitute orthogonal modes of knowledge production with their particular limits to predictive power. Estimating future developments through an extrapolation of past data at the cost of a conservative bias and the risk of hidden assumptions in data sampling and labeling can be opposed to predicting the future through the application of deductive transition rules to modeled representations of society, capturing process complexity but at the cost of reduced realism and a vulnerability towards hidden assumptions in the modeling of causal mechanisms and initial distributions.

What unites these approaches is the fact that the normative judgements of CSS modellers can easily be obscured behind what appears to be an objective practice (Jasanoff 2011), thus creating an additional ambivance of knowledge and worldview that is difficult to untangle. On the basis of Habermas' three models of political consulting (Habermas 1969) one can approach this ambivalence between scientific knowledge and politics. While the technocratic and the decisionist views favor one over the other, the pragmatic view is based on mutual dialogue and negotiation. The paper presents first results from interviews with computational social scientists in order to shed light on the field's internal perspective on these ambivalences and negotiations. It concludes with some provisional recommendations: In order to realize the potential of such a pragmatist, cooperative relation in policy advice, CSS should take into account existing social science theories, be integrated with qualitative forms of empirical research and embrace the tradition of participatory modeling. The task for technology assessment and related disciplines is to further broader societal debate about the extend and modalities of CMS in policy advice and to constructively analyze the ambivalences of legibilization.

*References*

*Ensign, Danielle et al. (2018): "Runaway Feedback Loops in Predictive Policing". In: Proceedings of Machine Learning Research 81, pp. 1–12. url: https://arxiv.org/abs/1706.09847.*

*Habermas, Jürgen (1969): "Verwissenschaftlichte Politik und öffentliche Meinung". In: Technik und Wissenschaft als »Ideologie«. Frankfurt a.M.: Suhrkamp, pp. 120–145.*

*Hofman, Jake M., Amit Sharma, and Duncan J. Watts (2017): "Prediction and explanation in social systems". In: Science 355(6324), pp. 486–488.*

*Humphreys, Paul (2009): "The philosophical novelty of computer simulation methods". In: Synthese 169(3), pp. 615–626.*

*Jasanoff, Sheila (2011): "The Practices of Objectivity in Regulatory Science". In: Social Knowledge in the Making. (Ed. by) Charles Camic, Neil Gross, and Michèle Lamont. Chicago: University of Chicago Press, pp. 307–338.*

*Miller, Peter and Nikolas Rose (1990): "Governing economic life". In: Economy and Society 19(1), pp. 1–31.*

*Pohle, Jörg (2012): Social Networks, Functional Differentiation of Society, and Data Protection. url: http://arxiv.org/pdf/1206.3027v1.pdf.*

*Scott, James C. (1998): Seeing like a state: How certain schemes to improve the human condition have failed. New Haven: Yale University Press.*

*Seo, Sungyong et al. (2018): "Partially Generative Neural Networks for Gang Crime Classification with Partial Information". In: AAAI ACM Conference on Artificial Intelligence, Ethics and Society. url: http://paleo.sscnet.ucla.edu/SeoEtAl-AIES-2018.pdf.*

*Simon, Miranda et al. (2018): "A data-driven computational model on the effects of immigration policies". In: Proceedings of the National Academy of Sciences 115(34), pp. 7914–7923.*

## The AMS Algorithm in Austria: Locating Biases via a Mathematical Framework

*Author: Paola Lopez (University of Vienna)*

Starting in 2019, the unemployment office in Austria (Arbeitsmarktservice Österreich, in short AMS) has been using a predictive model (Labour Market Opportunity Model, in German: Arbeitsmarkt-Chancen-Modell) to algorithmically classify job-seekers into three groups with different eligibility for support according to their predicted chances on the labour market.

This predictive model was developed especially for this purpose under the leitmotif of efficiency, and it became known in the media under the name AMS algorithm primarily with the publication of its accompanying method paper: It can be inferred from the paper that the personal data entry Gender: Female results in an automatic deduction of points, which means that a woman can be assigned to a less eligible group solely on the basis of her gender. Further potential point deductions according to personal data and individual features, such as age, parenthood, disability or nationality, lead to an intersectionally compounded disadvantage.

This is just one example of a worrying trend to legitimize socio-political decisions with highly questionable implications and effects using technological means that appear objective and neutral. The various algorithmic techniques that are being used in the US justice system to predict offenders' risk of recidivism are another and well-known example. Both these algorithmic decision aid systems use personal data, as well as data related to the individual history of the members of a highly vulnerable and dependent target group (job-seekers and criminal offenders, respectively) to predict the probability of a certain event (job placement and risk of recidivism, respectively). Based on this probability, measures are imposed that can have profound effects.

Thus, if biases are inscribed and encoded in these predictive systems, then these systems can be regarded as performative self-fulfilling prophecies which, in a first step, depict the reality of discrimination in the digital sphere, then, in a second step, normatively reinforce it as a supposedly objective fact and finally, in a third step, return it to the social sphere via the resulting measures.

This paper identifies three tangible building blocks of the AMS algorithm from a mathematical perspective. The predictive model and thus the categorization of job-seekers depends on and changes with each of the three building blocks which are all based on human decisions and thus on human assumptions, so that such a predictive model can be neither objective nor neutral. Using these three building blocks and the mathematical conceptual structure of the predictive model as an analytical framework, it becomes clearly visible where intrinsic biases are located and how this model in its current use can reproduce and reinforce discriminatory realities. Finally, the paper argues that this framework can be utilized to analyze a large class of predictive models (namely: all machine learning classification algorithms that are based on logistic regressions with set thresholds) in order to identify and locate biases and predisposed assumptions that can be harmful when applied to vulnerable target groups.

## Algorithmic profiling of job seekers in Austria: how to make austerity politics effective

*Authors: Doris Allhutter, Astrid Mager (Austrian Academy of Sciences), Florian Cech, Fabian Fischer (TU Wien), Gabriel Grill (University of Michigan)*

In October 2018, the Public Employment Service Austria (Arbeitsmarktservice, AMS) announced plans to roll out an algorithmic profiling system. Based on a statistical model of employment seekers' prospects on the job market from previous years, the system is designed to classify present and future job seekers into three categories. The first category consists of those with high chances to find a job within half a year, the second one includes those with mediocre prospects on the job market, and the third category consists of those clients with a bad outlook of employment in the next two years. The declared goal is to focus support on job seekers part of the second category, because for them "active labor market programs" (ALMP) are expected to have the most impact on improving chances to find employment. In contrast, job seekers from the first category are expected to find employment without much support, while job seekers from the third category are deemed as basically unemployable, no matter whether they participate in ALMPs or not. The algorithmic system should thus enable an efficient allocation of ALMPs by investing in job seekers for who ALMPs are expected to be most cost-effective for the AMS. In turn several representatives have referred to the third category as one for the "hopeless". This classification practice with its choices of properties for the statistical model, prompted a public debate on algorithmic bias, stigmatization and discrimination of people predicted to have a high risk of long-term unemployment (Johnson et al. 2018). Women, for example, are assessed as less employable in one published model than men. Similarly, job seekers from outside of the European Union are listed as having lower chances on the labour market than citizens of EU member states. This poses crucial questions in terms of the implicit politics of the profiling system: What are the implications of using retrospective data to calculate future chances of individuals on the job market? What variables and categories are considered in the statistical modeling and which ones are neglected? How does the system influence the job market and clients' prospects and which further implications can be expected in the practices of the AMS?

Based in critical data studies (Iliadis and Russo 2016, Rieder and Simon 2016, Striphas 2015, Van Dijck 2014), surveillance studies (Gandy 2016) and research on fairness, accountability and transparency in machine learning (Sandvig et al 2014, Veale et al. 2018) this talk will discuss the inherent politics of the AMS profiling system. An in-depth analysis of the available technical documentations and public statements will focus on three main areas of concern:

- First, the AMS claims the system just captures the "harsh reality" of the labour market and predicts job seekers' chances by taking into account existing inequalities. Our analysis problematizes claims that the system is neutral and objective. We argue it merely approximates the labour market's current state based on chosen data sources, attributes and methods reflecting value-laden judgements of the system designers. This claim of the AMS denies the system's role in shaping the job market and the clients' prospects.
- Second, we will discuss the distribution of agency and accountability within the socio-technical assemblage, particularly between the algorithmic profiling system and the case workers. This has implications for questions of accountability and the capacity for situated evaluations of the clients.
- Third, the concealed technical functionality of the system and its social consequences raise questions of transparency. Our talk will advance discussions on transparency and accountability requirements for deploying algorithmic systems in public agencies.

Using the example of the Austrian profiling system, our analysis will make a contribution to broader discussions on fairness, accountability and transparency in algorithmic systems and machine learning; both in the public sector and in other sensitive areas.

*References*

*Iliadis, A., & Russo, F. (2016). Critical data studies: An introduction. Big Data & Society, 3 (2), 1–7. Retrieved 2018-03-13, from https://doi.org/10.1177/2053951716674238 doi: 10.1177/2053951716674238*

*Johnson, T., Scoppetta, A., & Buckenleib, A. (2018). Tackling long-term unemployment through risk profiling and outreach: a discussion paper from the employment thematic network. (Technical Dossier No. 6). ESF Transnational Platform. Retrieved from https://publications.europa.eu/en/publication-detail/-/publication/ee87b6da-6d21-11e8-9483-01aa75ed71a1/language-en*

*Gandy, O. H. (2016). Coming to terms with chance: Engaging rational discrimination and cumulative disadvantage. Routledge.*

*Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014). Auditing algorithms: Research methods for detecting discrimination on internet platforms. Data and Discrimination: Converting Critical Concerns into Productive Inquiry, 1–23.*

*Striphas, T. (2015). Algorithmic culture. European Journal of Cultural Studies. 18, 395–412 http://journals.sagepub.com/doi/10.1177/1367549415577392*

*Rieder, G., & Simon, J. (2016). Datatrust: Or, the political quest for numerical evidence and the epistemologies of Big Data. Big Data & Society, 3 (1), 2053951716649398. Retrieved 2018-03-13, from https://doi.org/10.1177/2053951716649398 doi: 10.1177/2053951716649398*

*Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. Surveillance & Society, 12(2), 197–208.*

*Veale, M., Van Kleek, M., & Binns, R. (2018). Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making, 1–14. https://doi.org/10.1145/3173574.3174014*