WORKSHOP PROGRAM

Algorithms in Practice: Devices of Power and Valuation

Smådalarö Gård, May 10-12, 2015

http://www.smadalarogard.se

Welcome to our first algorithm network workshop. The aim of this first workshop is to strengthen collaboration in the social study of algorithms, start thinking about a summer school on algorithms, and to think about how we can collaborate in the longer term: as a group or perhaps in smaller constellations.

WORKING FORMAT

The workshop will be organized around four round-table discussions (the four themes below), with a number of discussants opening up each theme with a short presentation. The plan is that you, dear participant, will choose one theme from above and prepare a 10 minute speech which opens up facets of these themes for discussion.

To allow for a vivid discussion, please also prepare a couple of questions to the presenters in your theme.

THEMES

I	Sorting and valuing things with algorithms
II	Objectivity vs value-ladedness
III	Inclusions/exclusions and algorithms
IV	Valuing algorithms (How do people grapple with what is a good/bad algorithm?)

SUNDAY

19:00 Dinner

MONDAY

9:30-10.00 Welcome and introduction (Francis & Lotta)

10:00-11:45 Session I: Sorting and valuing things with algorithms (Chair: Lotta Björklund Larsen)

11:45-12:45 Lunch

12:45-13:45 Talk Walk

14:00-15:30 Session II: Objectivity vs value-ladedness (Chair: Francis Lee)

15:30-16:00 Fika

16:00-17:30 Session III: Inclusions/exclusions and algorithms (Chair: Teun Zuiderent-Jerak)

18:00-19:00 Sauna with beer (remember to bring your swimming gear!)

19:30 Dinner

TUESDAY

9:00-11:00 Session IV: Valuing algorithms (Chair: Johan Nilsson)

11:00-11:15 Fika

11:15-11.30 Summing up, Steve Woolgar

11:30-12:00 Whats next… (Francis & Lotta)

12:00 Lunch

13:00 Departure

Session 1: Sorting and valuing things with algorithms

Big data analytics via algorithms

Magnus Boman (text written with Daniel Gillblad for a book chapter TBC), KTH
mab@kth.se

Abstract: It is now reasonable to define Big Data and the analytics that go with it in terms of a seemingly significant shift in what data sizes and velocities are practical to work with, along with a renewed realization that data is extremely valuable. Computational power and storage capacity have grown to the point where working with huge data sets is feasible with relatively cheap technology. More than that, computation and storage are becoming more scalable, and an investment in a machine of fixed capacity that becomes obsolete is becoming a thing of the past. We have since the early 00s seen the development of tools and middleware that harness the computational power of data centers with tens of thousands of machines, giving us the opportunity to scale our analytics algorithms to data sizes of exabytes and beyond. Not surprisingly, these developments coincide with the avalanche of data that we are amassing today. Earlier driven mainly by the staggering amount of information that humans leave behind when publishing on and using Internet services, our digital traces, we are currently seeing more and more information being collected from machines, sensors, and devices with little human interaction, increasing the speed of data collection even further. Techno-positivists claim that this is a trend that will extend into the next five years, boosted by infrastructural developments like 5G and Internet of Things, a claim that screams for scrutiny under critical perspectives, not least because of the related claim that our digital infrastructure is becoming our societal infrastructure.

Data Massage and Algorithmic trials in Governance of Danish Education

Helene Ratner, Copenhagen Business School
hr.mpp@cbs.dk

Stakeholders around Danish education are currently involved in developing new database devices (Ruppert, 2012) with the aim of facilitating systematic evaluation of the quality of education. A central feature of these devices is granular access to data. Speaking of this as a shift of data use from 'controlling' to 'massaging' individual primary schools, civil servants at the Ministry of Education envision that granular access to data affords dialectic of information and intervention, in which database devices evaluate and anticipate individuals' and institutions' performance. The development of new database devices, however, also occasions what I refer to as 'algorithmic trials': situations in which decisions about algorithms become a point of contention or negotiation. I briefly discuss two such algorithmic trials: one involves negotiating a distinction between research and politics when researchers and civil servants debate how to visualize and evaluate data on pupils' well-being. The other involves deciding on 'good' pupil performance when the ministry changes the evaluation criteria of pupils' performance in national tests.

Learning machines and their passion for error

Adrian Mackenzie, Lancaster University
a.mackenzie@lancaster.ac.uk

In very many settings today, algorithms that learn are highly valued. What does it mean for an algorithm to learn, and to learn in so many different settings (stock markets, scientific fields, medicine, media and entertainment, etc.)? What kind of pedagogy, training, testing and examination accompanies this learning? Typically, such algorithms are trained during a 'supervised learning' phase, and then tested for their ability to 'generalize' to hitherto unseen cases. The power of such 'learning machines' to classify, rank or predict depends less on an innately powerful algorithm shredding data like a food processor than on disciplined alignments between the data that the algorithm encounters, penalised treatments of variations and mis-classifications, and a thoroughly test-based optimisation of algorithmic movements. In many ways, this mode of algorithmic disciplining seems to owe much more to disciplinary techniques of power than it does to mathematics, computer science or statistics. If that is the case, and if we are seeing a generalization of algorithmic power, then this disciplinary mode of formation of algorithmic learning (a formation in which algorithms are more like delinquent juveniles than sovereign abstractions) may have powerful effects on the forms of sorting, ordering, classification and regulation assembled in contemporary social fields. At the core of this disciplinary formation of machine learners lies an irreducible potency of error, which is both the target of much optimisation and the margin of indeterminacy whose ongoing presence animates a seemingly inexorable intensification of algorithmic processing. A small, almost toy demonstration of facial recognition algorithms, `kittydar`, illustrates both the power and problems of this sorting power. On the one hand, after its training phase has finished, `kittydar` is mostly automatic in its sorting of images of cat faces. On the other hand, the work that prepares and repairs this capacity to sort means that any appearance of learning and any performance of automatic face recognition involves multiple agents acting at different times.

How Do Algorithms Interact? The Comparative Sociology of High-Frequency Trading

Donald MacKenzie, University of Edinburgh
d.mackenzie@ed.ac.uk

As the themes of these workshops make clear, algorithms sort and value people and things. They include some things and people and exclude others. The focus of my talk is on a situation in which the sorting and valuing is done not just by an individual algorithm, but by algorithms interacting with each other (and with human beings as well).

The talk will briefly explain the main mechanism by which trading algorithms interact, the so-called 'limit order book'. It will then sketch the differences between how algorithms interact in three markets: a) US shares; b) US futures; c) foreign exchange. I will argue that algorithms are not just devices of power, but also exist in and are shaped by fields of power (fields of which they themselves form part), and that this can be seen by comparing how algorithms interact in markets that superficially seem similar.

Session 2: Objectivity and value-laden-ness'

'Why we love algorithms'

Philip Roscoe, University of St Andrews
pjr10@st-andrews.ac.uk

A casual observer of algorithms can only be struck by how much is put into, and expected from, these mindless computerised processes. Bitcoin, for example, promises a libertarian techno-utopia, sound money, and even community (Maurer, Nelms, & Swartz, 2013). It may not be a universal appeal, but those who like algorithms love algorithms.

I will suggest that our delight in algorithms has less to do with perceived objectivity, and more to do with perceived omniscience; that the appeal of algorithms is embedded in an entirely separate intellectual tradition from the formal-rationality (Weber, 1978) associated with objective numbers and bureaucratic management systems. Algorithmic authority has charismatic roots and stems from a theological tradition: Hayek's (1973) appropriation of natural ordering vested providential power in the market, understood as an algorithm, evolutionary in nature, and all knowing. For Hayek market superiority was based on its computational superiority; 'cyborg economics'(Mirowski, 2002) suggests that algorithmic power is superior, and that the market is simply a special case of algorithms more generally (Dennett, 1995). For Dennett, algorithms are mindless labourers, working tirelessly on our behalf to evaluate a world of possibilities. All knowingness, not objectivity, is the primary attribute of the algorithm.

I offer two ways of thinking through this claim. The first is via theories of trust. Hawley (2011) distinguishes 'rich, interpersonal trust' from mechanical 'reliance' by the notion of making and meeting commitments. This requires ability and effort, as well as a degree of self-knowledge, agency-filled activities no-longer beyond the reach of machine-learning algorithms. But trust alone cannot account for our enchantment with algorithms. May (2011) has suggested that love flows from the sense of 'ontological rootedness' – that we with love those people or things which give us a feeling of place in the world. Love of the divine is a paradigm case. We must strive to love, but in the impossibility of such love being requited, we must be prepared for injustice, cruelty and abandonment. Algorithms such as the crypto-currency 'blockchain' seem to have such characteristics. In other words, my talk is a preliminary attempt to sketch out the theological underpinnings of the algorithm.

Dennett, D. C. (1995). Darwin's Dangerous Idea. London: Allen Lane: The Penguin Press.
Hawley, K. (2011). Trust: A Very Short Introduction. Oxford: Oxford University Press.
Hayek, F. A. (1973). Law, Legislation and Liberty. London: Routledge.
Maurer, B., Nelms, T. C., & Swartz, L. (2013). "When perhaps the real problem is money itself!": the practical materiality of Bitcoin. Social Semiotics, 23(2), 261-277.
May, S. (2011). Love: A History. New Haven and London: Yale University Press.
Mirowski, P. (2002). Machine Dreams: Economics Becomes a Cyborg Science. Cambridge: Cambridge University Press.
Weber, M. (1978). Economy and Society. Berkeley: University of California Press.

Algorithms and/as public concerns. The algorithmic in ecologies of sensation

Willem Schinkel, Erasmus University Rotterdam
Schinkel@fsw.eur.nl

The main aims of my talk are twofold: 1) to try to specify what I deem to be of most social scientific interest in studying 'algorithms'; 2) to outline some of the public issues at stake in the increasing ubiquity of digital algorithms.
The first aim involves an effort to specify what we study, when we study algorithms. Here, I propose that we focus on what Tarleton Gillespie has called 'the algorithmic'. The point is here to understand the distributed nature of algorithms. Rather than mere calculative codes, I argue that algorithms are more fruitfully conceived as distributed devices that help structure and sustain ecologies of sensation, i.e., socio-technical ecologies that continually sense and monitor themselves.
The second aim leads me to zoom in on the questions of public relevance that such algorithmic ecologies of sensation entail. Which political options are blotted out by algorithmic mediation? Which organizational relocations of knowledge, power and accountability are involved? And, finally, what economy thrives on the ubiquity of algorithmic ecologies of sensation? My illustrations of some of these points are based on a study of Google's search algorithms, and they stem from a more encompassing research interest in ubiquitous monitoring and ecologies of sensation.

Session 3: Inclusions/exclusions and algorithms

The Algorithm Will Include You Now: What's Digital Payment All About?

Bill Maurer, University of California at Irvine
wmmaurer@uci.edu

This talk will set out the main business arguments for financial inclusion via mobile and digital payment services in the global South, as they have evolved over the past decade. What began as an argument for revenue collectors about formalizing "informal" economic practices and financial flows for tax purposes, later morphed into a full-throated argument for "financial inclusion" whose business case was murky at best. In the face of this, providers started changing tactics, seeking new ways to use transactional data generated by electronic payment. The talk will reflect on these shifts, from the service providers' perspective, while opening up for discussion the very notions of "inclusion" and "exclusion" in emerging algorithmically-defined market spaces.

Enumeration, Resistance, and the Jedi: Implications of Big Data in Official Statistics

Baki Cakici, Goldsmiths, University of London
b.cakici@gold.ac.uk

The Jedi census phenomenon, a movement that protested the inclusion of the religion question on the census survey in several English speaking countries, attracted more than 390 000 citizens in England and Wales in 2001. Although the protest was a form of census resistance, the UK National Statistician reported that it had also strengthened the census results as it led to increased participation among groups that did not traditionally return their census forms. In this discourse, it was possible to interpret resistance itself as a data point to strengthen the produced knowledge.
A similar shift is underway as National Statistical Institutes in Europe increasingly turn to highly complex computational methods such as big data analytics to improve their data collection and processing capabilities. In these methods, individual citizens are not consulted about their presence or absence as the algorithms figure it out (accurately or otherwise). Participation becomes increasingly passive, to the point where the possibility to opt out becomes extremely unlikely, sometimes disappearing completely. Dissent and subversion remain possible, but they are increasingly moved to digital domains which require technological expertise to navigate successfully (e.g. anonymisation services, proxies, VPNs). As with the Jedi protest, the very fact that an activist movement objects to the data collection generates new data points that can be used to enumerate them, often with the help of social media data.
In my presentation I discuss the relationship between algorithmic enumeration and subjectivity using examples from two census-related Big Data pilot projects at a European National Statistical Institute.
· A clustering algorithm (DBSCAN) to identify frequently visited locations based on Twitter data.

· A machine learning method (logistic regression) to identify unoccupied houses based on smart meter data.

Through these cases, I pose the following questions: Which subject positions open up when algorithms process population data, and which positions are excluded? What are the implications of enacting a population that is enumerable by its data traces? What forms does resistance take in response to a data-driven enumeration methodology?

Scaling up and down – (big) data and algorithms in the field of security

Dennis Broeders, Erasmus University Rotterdam / Dutch Scientific Council for Government policy
dennis.broeders@gmail.com

The strength of big data is – like the saying goes – in numbers. Many of the perceived benefits of big data analysis depend on scale in the sense that conclusions drawn from big data are usually applicable to the aggregate level and reveal major shifts and trends that can be seen through data correlations. The level of analysis and the 'bigness' of data compensate for the fact that the data are 'dirty'. However, algorithmic outcomes are, or need to contribute to, "actionable knowledge". In most cases the analysis is not done for the sake of analysis. When governments start using big data and algorithms they usually do so to inform and/or execute policy. In many countries governments are scouting out the possibilities for the use of algorithms in policy, not in the least in the fields of safety and security. There are – at least - two relevant dimensions to consider here. The first is the public-private dichotomy/continuum (which is no longer a strict dichotomy in terms of organisation, and increasingly less so in term of data, but is in terms of the (power) relation to their relative 'subjects'). Dealing with citizens rather than consumers requires different safeguards for the use algorithmic probability in policy; this goes double for the field of security. Government is a monopolist – exit is hardly a serious option – and especially in the field of security 'actionable knowledge' may ultimately mean that someone gets a knock on the door from the authorities. That requires different safeguards than the algorithmic predictions from, say, Amazon.com. The second axis, or continuum, is the scale on which the actionable knowledge is to take effect: the aggregate, the group and the individual. There is a difference between using algorithms to track epidemics, to profile groups and to pinpoint individuals. These various levels of collectiveness and degrees of identity require different degrees of protection. Where the promise of big data is grounded in the idea that the best results come from massive data and a lack of interest in the individual (and instead a focus on trends in the aggregate), many governments are ultimately interested in groups and individual cases. One might fear that a political belief in the power and 'accuracy' of big data in the aggregate is transported to a lower scale by 'translating' aggregate findings groups or even the individual – a fallacy of scale.

Session 4: Valuing algorithms

Assessing adaptive design drug trials: An inquiry into the valuation of algorithmic endeavours to produce knowledge

Claes-Fredrik Helgesson, Department of Thematic Studies - Technology and Social Change, Linköping University
claes-fredrik.helgesson@liu.se

How can a drug trial be assessed prior to it being conducted? Empirically, this study focus on how researchers, companies, and regulators work with this question as regards a fairly new breed of trials, namely adaptive design clinical trials (ADTs). An ADT is algorithmic in the sense that it includes planned opportunities for altering one or several aspects of the study based on analyses of interim data. The possibility of alterations sets such trials apart from conventional randomised controlled trials (RCTs). The question of what is a good trial design is always topical. Yet, the valuations involved in assessing an ADT becomes more convoluted when the design itself includes a number of possible paths the actual trial can take depending on rule-based decisions guided by interim analyses.

The increased discussion and deployment of adaptive design trials appears to be linked to a number of transformed articulations of what matters in trials as well as new ways to establish if a given design is considered to be a good one or not. Simulations of proposed designs are among the tools used to establish how a given adaptive design trial might behave in different scenarios. The broader argument that I want to develop with this case study is that the emergence of new unfathomable algorithms like adaptive design trials provides opportunities for re-ordering what is at stake and that the arena where this re-ordering is done is in the devising of their assessment.

Valuing Algorithms: How do people grapple with what is a good/bad algorithm...?

Farzana Dudhwala, Said Business School, Oxford
farzana.dudhwala@sbs.ox.ac.uk

To the lay person with no knowledge of how an algorithm works, the valuation of an algorithm is not based on technical know-how. What, then, is involved in making a decision about the value of an algorithm? At least in cases where an algorithm is the result of a self-quantification process, I'd like to proffer the idea that perhaps an algorithm's value is determined by intuition and a credible narrative. For example, if using a mood tracking application that determines levels of anxiety and depression based on selections of mood on a scale (using an algorithm to determine eventual indication of anxiety and depression) and this determination does not match up with an individual's own intuition about how they are feeling then perhaps they might deem the algorithm to be a bad one. However, I would also argue that there is a certain threshold beyond which the algorithm must go in order for it to be classed as bad. Within the realms of reasonable credibility, it might be the case that the algorithm is deemed to be a good one and perhaps it is believed that the mood that has become different to what one may have intuited. It is this tension that I would like to discuss.

Interpretability: Learning to Listen to Algorithms

Nick Seaver, University of California at Irvine
nseaver@uci.edu

A recent theme in critical discussion about algorithms is the notion that they have become inscrutable to the people who make them. Neural networks, machine learning, and elaborate technical infrastructures make it hard to know exactly why specific outputs appear, even for "insiders" who once were the objects of calls for transparency. If we used to think algorithms were obscure because they were secret, we can now add to this the more fundamental obscurity of complexity. Nonetheless, the designers and builders of algorithms, like "outsider" users, do try to make sense of their systems. Drawing on several years of ethnographic fieldwork with the developers of algorithmic music recommendation systems, I detail the production of "interpretability" — a term used by the coders of neural networks and builders of recommender systems to describe how algorithmic outputs make sense. I suggest that insiders and outsiders alike have developed an "acoustemology" of algorithms, in which they come to understand how algorithms work by listening to them.

Algorithm Behavior

Teun Zuiderent-Jerak, Department of Thematic Studies – Technology and Social Change
Linköping University
teun.zuiderent-jerak@liu.se

When establishing a new network, one often starts to see the topic of that network everywhere. Then, usually about halfway through the network, the question all of a sudden gets raised, "what is actually not a … [insert: 'bio-object', 'technology of personalization', 'algorithm']". I propose raising this question a little earlier, by exploring the difference between my work on guidelines within the sociology of standardization and the social study of algorithms. Although there are many similarities between the sociology and anthropology of standardization and that of algorithms, I also think it may be helpful to point to the differences between them – or rather to the possibly quite specific place of algorithms within the study of standards.

In my work, together with Esther van Loon, on a guideline for dealing with problem-behavior in elderly care, we found that the standard was targeted at increasing certain types of reflexivity through being highly prescriptive. Not merely did it seriously restrict the use of sedative medication; it also prescribed a clearly outlined stepped approach for a reflexive process that professionals dealing with problem-behavior had to follow. Through this combination of restricting dominant modes of acting (sedating) and prescribing desired modes of reflection (especially the steps where the 'problem' of the problem-behavior gets unpacked), the guideline was able to change persistent dynamics in long-term care. It did so by making the standard itself highly visible and present during the work on problem-behavior: otherwise professionals would quickly slip back into repressive modes of dealing with the issue.

Based on this case I will raise a few questions for the social study of algorithms:

How does the invisibility of algorithms influence their possibilities for acting at a distance? What are the possible relationship between algorithms and reflexivity? Are they always at odds? Or can algorithms 'frame' reflexivity just like guidelines can? What is the possible benefit of treating algorithms as a highly specific case of standardization?

Spectacular Algorithms

Malte Ziewitz, Cornell University
mcz35@cornell.edu

Search engine optimization (SEO) consultants have created algorithm studies networks long before we did. Concerned with helping their clients rank at the top of search engine results pages, these mostly self-trained professionals have been facing many of the challenges social scientists are struggling with today. How to make sense of seemingly inscrutable algorithms? What counts as a "good" algorithm? And how far can "algorithm talk" get us anyway in view of urgent practical and political concerns? Drawing on long-term fieldwork in the UK search marketing industry, I shall suggest that there are a few things we can learn from SEO consultants about using and valuing algorithms in analytic practice. Specifically, I shall suggest a move from treating algorithms as a largely uninterrogated analytic resource to understanding how they figure in practical reasoning – as "algorithmic reasoning." This shift from rules to reasoning, from data to events, from empiricism to experimentation will not only illuminate the work of search marketing professionals, but also help us value algorithms in our own work.

PARTICIPANTS

Adrian Mackenzie, Lancaster University	a.mackenzie@lancaster.ac.uk
Baki Cakici, Goldsmiths	b.cakici@gold.ac.uk
Bill Maurer, University of California at Irvine	wmmaurer@uci.edu
Claes-Fredrik Helgesson, Tema T, Linköping University	claes-fredrik.helgesson@liu.se
Dennis Broeders, Erasmus University Rotterdam	dennis.broeders@gmail.com
Donald Mackenzie, Professor, University of Edinburgh	d.mackenzie@ed.ac.uk
Eva Danielsson, Tema T, Linköping University	eva.danielsson@liu.se
Farzana Dudhwala, Said Business School	farzana.dudhwala@sbs.ox.ac.uk
Francis Lee, Tema T, Linköping University	francis.lee@liu.se
Helene Ratner, Copenhagen Business School	hr.mpp@cbs.dk
Ivanche Dimitrievski, Tema T, Linköping University	ivanche.dimitrievski@liu.se
Johan Nilsson, Tema T, Linköping University	johan.nilsson@liu.se
Lotta Björklund Larsen, Tema T, Linköping University	lotta.bjorklund.larsen@liu.se
Magnus Boman, KTH-Royal Institute of Technology	mba@kth.se
Malte Ziewitz, Cornell University	mcz35@cornell.edu
Nick Sievers, University of California at Irvine	nseaver@uci.edu
Patrick Vonderau, Stockholm University	patrick.vonderau@ims.su.se
Philip Roscoe, University of St Andrews	pjr10@st-andrews.ac.uk
Steve Woolgar, Oxford University, UK and Linköping University	steve.woolgar@liu.se
Teun Zuiderent Jerak, Tema T, Linköping University	teun.zuiderent-jerak@liu.se
Willem Schinkel, Erasmus University Rotterdam	schinkel@fsw.eur.nl