

Ethics of AI and Algorithmic Reason

Department of War Studies

Claudia Aradau

Department of Digital Humanities

Tobias Blanke

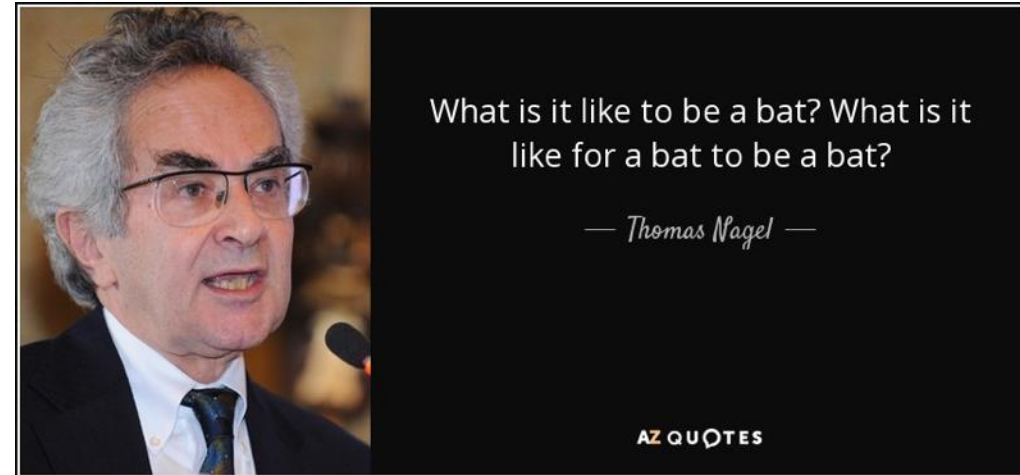
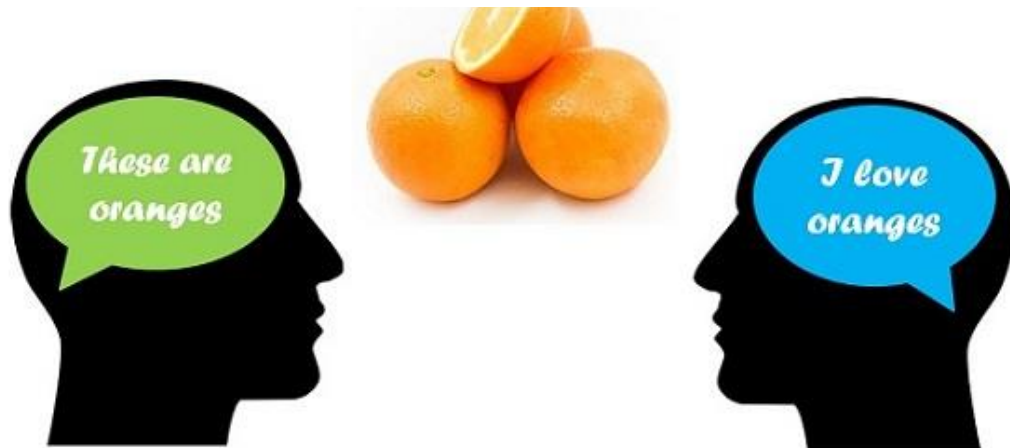
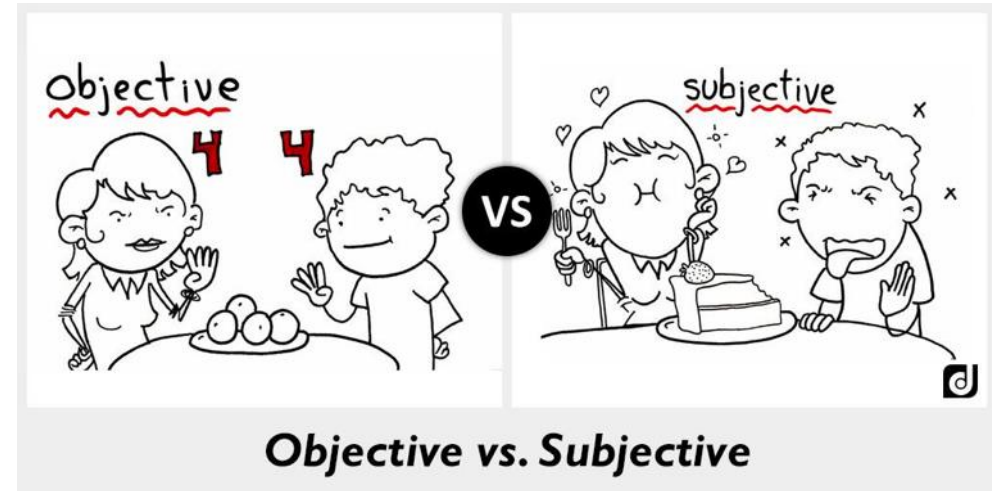
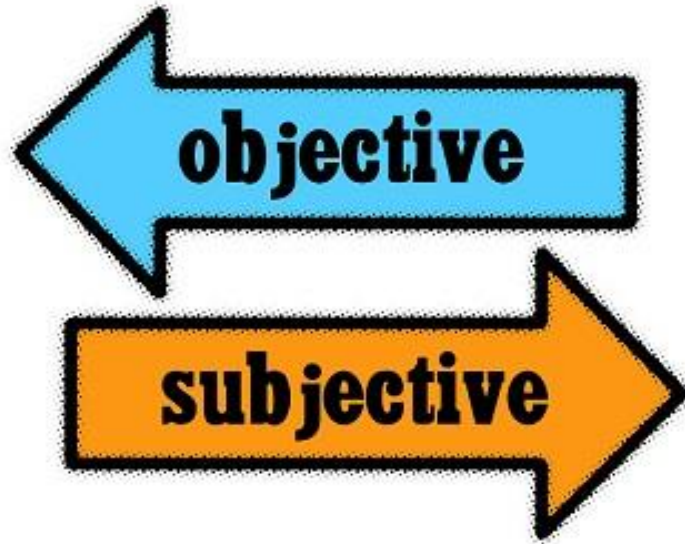
The logo for King's College London, featuring the text "KING'S" in a large, white, serif font, "College" in a smaller, white, cursive font, and "LONDON" in a white, serif font below it. The text is set against a red square background. Below the text are two horizontal white lines.

KING'S
College
LONDON

Overview

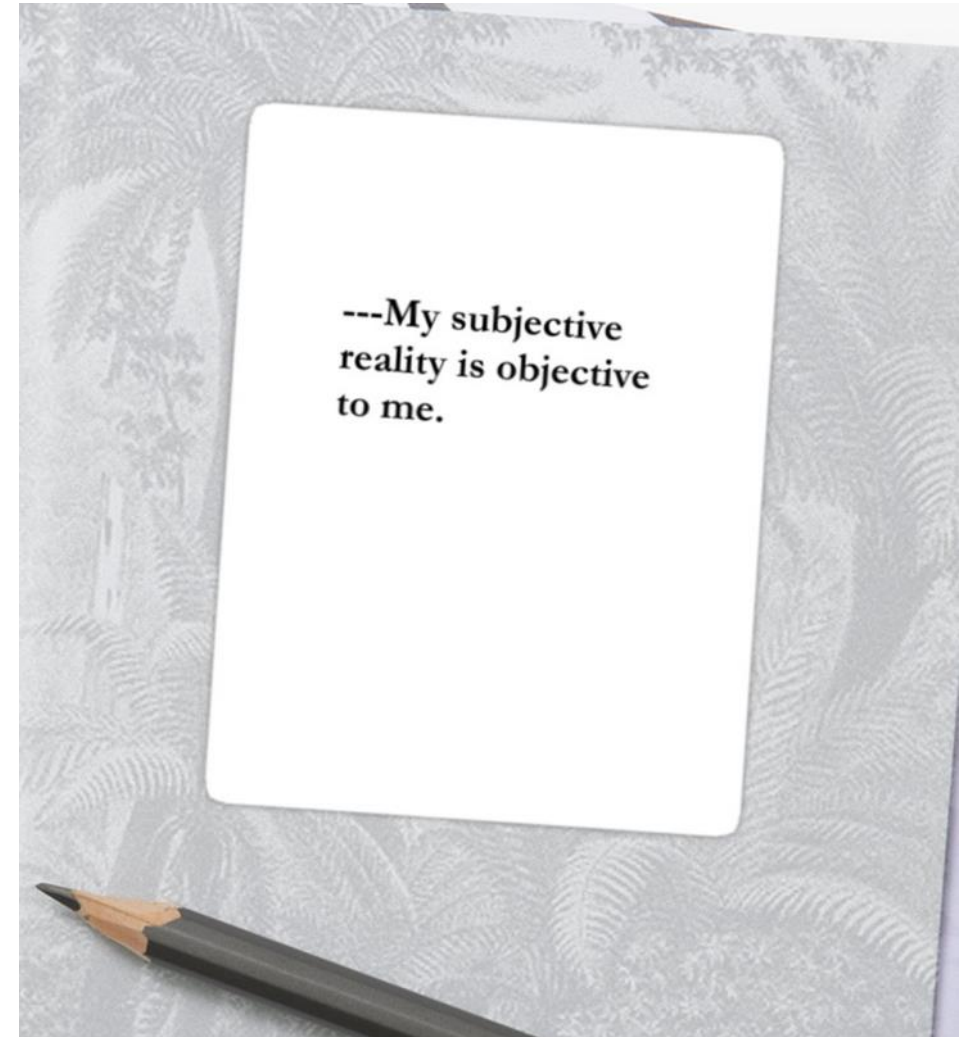
- **Ethics and its relevance to AI**
- **Ethical Subjectivity**
- **Failed Subjectivity**

Philosophy as Dichotomy of Subjectivity and Objectivity



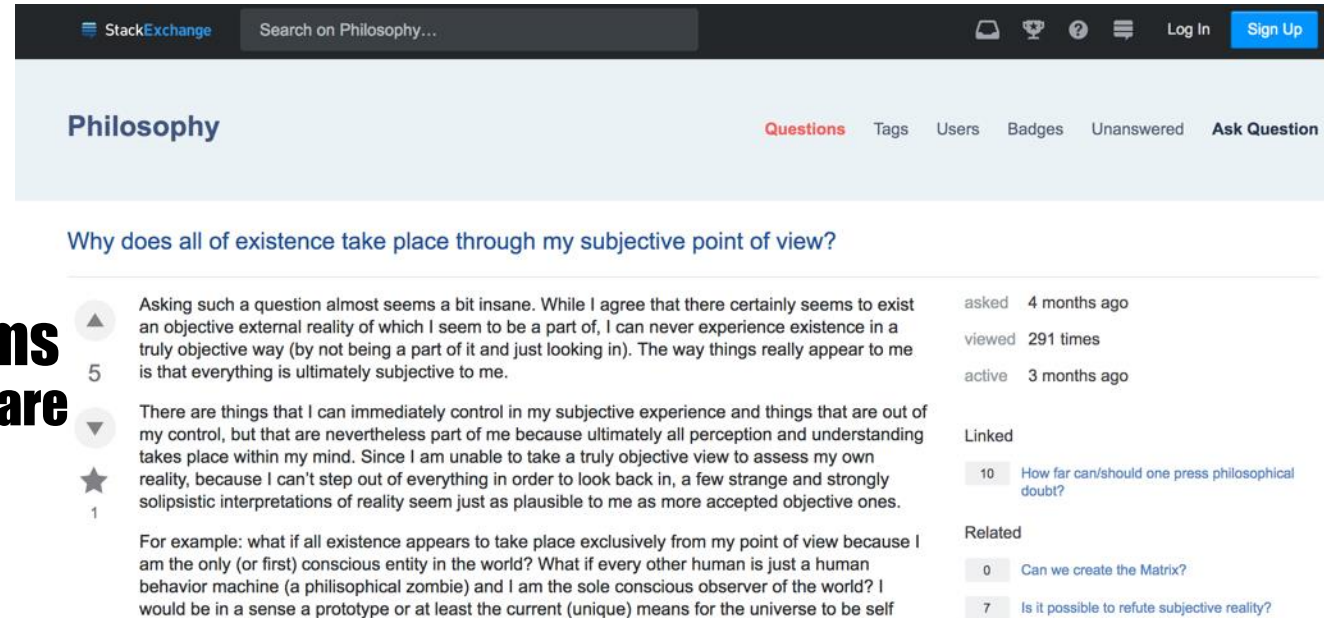
Ethics a celebration of subjectivity

- **Code of Conduct**
- **Applied ethics**
 - **bioethics, business ethics, environmental ethics, etc. etc. etc.**
- **Moral agents or ethical subjects**
 - **Ethical development by technology?**



Some 2000 year old challenges (in Europe)

- **Celebration of freedom**
 - **But what is immoral?**
- **Morality breaks traditions and customs**
 - **Religion is more than morality, as these are historical stories that include immoral actions too like the use of stoning in the testament**
- **Our own morality includes the claim that all other rational people should endorse it**
 - **In heterogeneous societies there seems to be no universal code or at least not a direct one**



The screenshot shows a StackExchange page for the 'Philosophy' tag. The question title is 'Why does all of existence take place through my subjective point of view?'. The page includes a search bar, navigation links (Questions, Tags, Users, Badges, Unanswered, Ask Question), and a list of answers. The top answer, by user '5', has 5 votes and discusses the subjective nature of existence. The second answer, by user '1', has 1 vote and discusses the possibility of being the only conscious entity in the world. The right sidebar shows related questions with their respective scores and titles.

StackExchange Search on Philosophy... Log In Sign Up

Philosophy Questions Tags Users Badges Unanswered Ask Question

Why does all of existence take place through my subjective point of view?

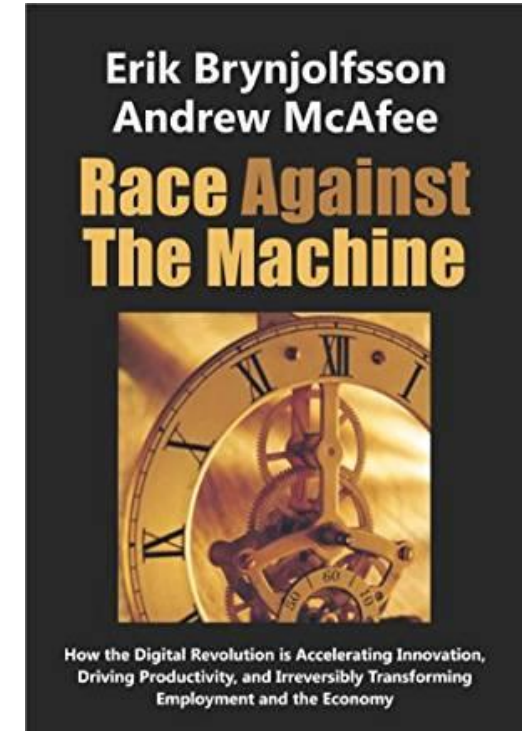
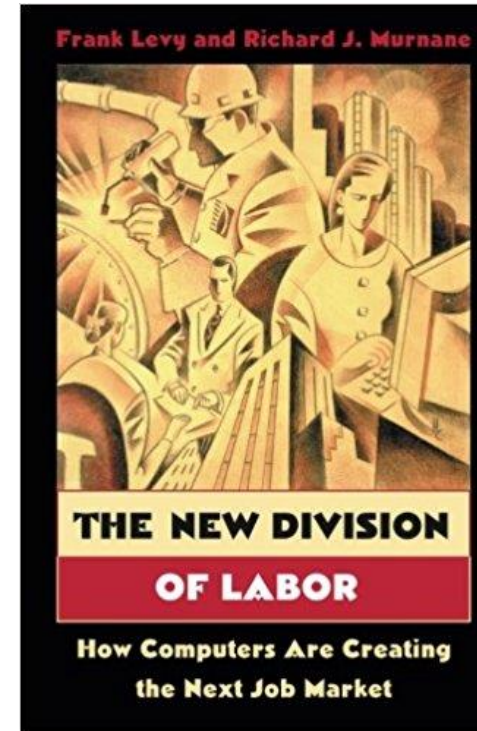
▲ 5 Asking such a question almost seems a bit insane. While I agree that there certainly seems to exist an objective external reality of which I seem to be a part of, I can never experience existence in a truly objective way (by not being a part of it and just looking in). The way things really appear to me is that everything is ultimately subjective to me. asked 4 months ago viewed 291 times active 3 months ago

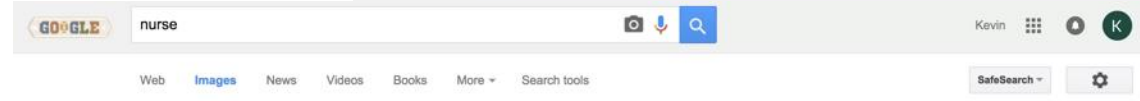
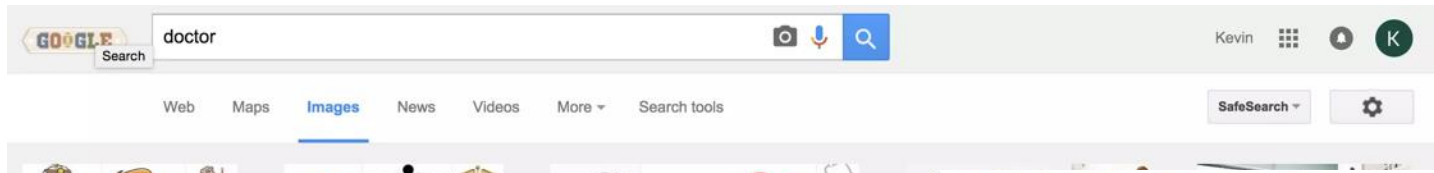
▼ 1 There are things that I can immediately control in my subjective experience and things that are out of my control, but that are nevertheless part of me because ultimately all perception and understanding takes place within my mind. Since I am unable to take a truly objective view to assess my own reality, because I can't step out of everything in order to look back in, a few strange and strongly solipsistic interpretations of reality seem just as plausible to me as more accepted objective ones. Linked 10 How far can/should one press philosophical doubt? Related 0 Can we create the Matrix? 7 Is it possible to refute subjective reality?



Ethics and AI

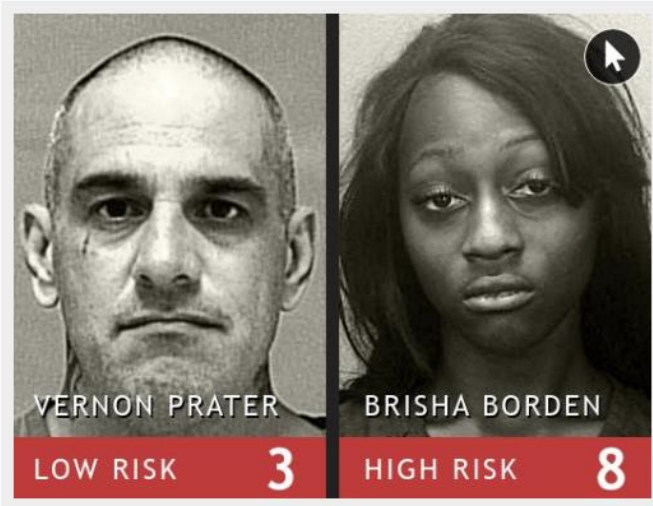
- **(Super-)intelligence taking over the world**
 - **Terminator**
 - **Drones**
- **The stuff of PhD theses: If a fully automated car loses control and has to decide whether to roll over a young woman pushing a stroller or two elderly women**
- **Automation**
 - **Complete new areas of work are affected**
- **Data control and privacy**
- **Bias**





https://mostlysignssomeportents.tumblr.com/post/144859196220/algorithmic-risk-assessment-hiding-racism-behind

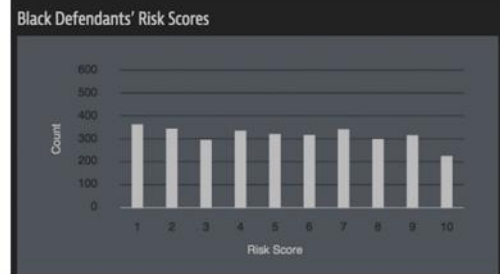
Algorithmic risk-assessment: hiding racism behind "empirical" black boxes



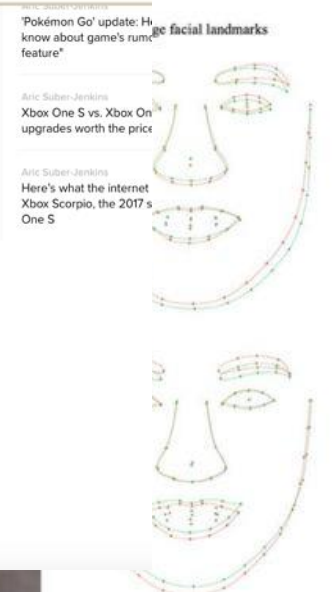
Courts around America and the world increasingly rely on software based risk-assessment software in determining bail and sentencing; the systems require the accused to answer more than a hundred questions which are fed into a secret model that spits out a single-digit "risk score" that courts use to decide who to lock up, and for how long.



Tech.Mic
U.S. Courts Are Using Algorithms Riddled With Racism to Hand Out Sentences



White defendants were routinely given lower threat scores than black defendants.



Don't see reply. Use

My pix C
<http://cr>
<http://bo>
[@docton](#)

GPG key
GPG key
3A77 F8



EQUALS CHANGE BLOG

Can computers be racist? Big data, inequality, and discrimination

<https://www.fordfoundation.org/ideas/equals-change-blog/posts/can-computers-be-racist-big-data-inequality-and-discrimination/>



Technology Program Officer, Internet Freedom



<https://vimeo.com/145335290>

Civil Rights and Algorithms

At the beginning of 2014, as an answer to the growing concerns about the role played by data mining algorithms in decision-making, US President Obama called for a 90-day review of big data collecting and analysing practices.



https://www.whitehouse.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf

House of Lords Report

- **Principles:**
 - **Common good and benefit of humanity**
 - **Intelligibility and fairness**
 - **Data rights or privacy of individuals, families, or communities**
 - **The autonomous power to hurt, destroy, or deceive human beings should never be vested in artificial intelligence**
- **Need to look at near-future scenarios rather than doomsday**

*"As soon as it works, no one calls it AI anymore ..."*¹

You wake up, refreshed, as your phone alarm goes off at 7:06am, having analysed your previous night's sleep to work out the best point to interrupt your sleep cycle.² You ask your voice assistant for an overview of the news, and it reads out a curated selection based on your interests.³ Your local MP is defending herself—a video has emerged which seems to show her privately attacking her party leader. The MP claims her face has been copied into the footage, and experts argue over the authenticity of the footage.⁴ As you leave, your daughter is practising for an upcoming exam with the help of an AI education app on her smartphone, which provides her with personalised content based on her strengths and weaknesses in previous lessons.⁵

On your way to work, your car dashboard displays the latest traffic information, and estimates the length of your journey to the office, based on current traffic conditions and data from previous journeys.⁶ On arrival, you check your emails, which have been automatically sifted into relevant categories for you.⁷ A colleague has sent you several dense legal documents, and software automatically highlights and summarises the points most relevant to a meeting you have later.⁸ You read another email, sent by your partner, asking if he can borrow your bank login details to quickly check something. On closer inspection you decide it is probably a fake, but still, you hesitate before deleting it, wondering briefly how the spammers captured his writing style so unerringly.⁹

You have other things to worry about though, as you head to a hospital appointment. However, after a chest x-ray, you are surprised when the doctor sits you down immediately afterwards, explaining that you look to have a mild lung infection—you had expected it to take weeks before the results came back.¹⁰

Your relief is short lived—a notification on your phone warns you of suspicious activity detected on your bank account, which has been automatically stopped as a result.¹¹ You call the bank, and someone called Sarah picks up, and helps you order a replacement card. Except, you soon realise, Sarah is not human at all, just a piece of software which sounds just like a real person.¹² You are a little unnerved you did not realise more quickly, but still, it got the job done, so you do not particularly mind.

After a quick detour to the local supermarket, where the products on the shelves have all been selected automatically based on previous customer demand, current shopping trends and the likely weather that day, you drive home.¹³ On your way back, your car detects signs that you are feeling slightly agitated, and chooses some music you have previously found relaxing.¹⁴ After dinner, you and your partner watch a film suggested by your TV, which somehow strikes just the right note for both of your normally divergent tastes.¹⁵ After dozing off, your house, predicting you are asleep by now, turns off the bathroom light and turns on the washing machine, ready for another day.¹⁶

AI Ethics as career

Secure | <https://www.cs.ox.ac.uk/efai/>

Talis



Home Towards a Code of Ethics for AI Is AI Ethics Special? Developing Codes of Ethics for AI Resources



Elon Musk donates \$10M to keep AI beneficial

Ethics in Artificial Intelligence

This is the website for a project based in the [Computer Science Department, University of Oxford](#), 'Towards a Code of Ethics for Artificial Intelligence Research', which is being carried out by Professor Mike Wooldridge, Professor Peter Millican, and Dr Paula Boddington. It is one of a [group of projects](#) funded by the Future of Life Institute in 2015 for a global research programme aimed at [keeping AI beneficial to humanity](#), with a donation of \$10,000,000 from Elon Musk. We were very excited to be one of 37 projects granted an award.

Artificial Intelligence

DeepMind's new AI ethics unit is the company's next big move

Google-owned DeepMind has announced the formation of a major new AI research unit comprised of full-time staff and external advisors

PENTAGON WILL EXPAND AI PROJECT PROMPTING PROTESTS AT GOOGLE



This talk focuses on Ethical Subjectivity

What would it mean
to create ethically
behaving AI? – Food
for Thought
Think about:

AI's decision are not subjective enough as
they are too objective. It has always tried to
reproduce the subjectivity of experts.



Ethical Subjectivity for the Digital Human

Critique of the datafication of culture

Data citizens



Enabling the full participation in the datafied society

- **Big Data**
- **Network devices including the Internet and social media**

Critical Citizenship involves not just the rights but also to claim rights

Datafication of Culture



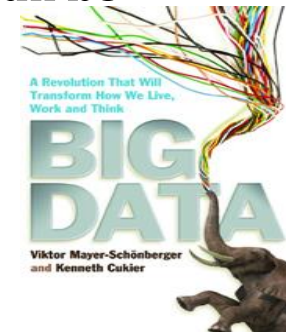
'To datafy a phenomenon is to put it in a quantified format so that it can be tabulated and analysed' (Mayer-Schönberger and Cukier, 2013: 78)

Datafication

- **process whereby the world is captured in data**
- **& reconstituted into new forms of value**

Insights on

- **what we think, how we feel, what we respond to and in what way, where we go, what we do, who we interact with, what we listen to, what we read, what we like, who we like it with, who we like, and so on**



Data Citizens: Our Data Ourselves



Home > Hacks > Kings College OpenBSD Hack

Kings College OpenBSD Hack

New project

See the full event details

5 projects

Data Discovery by @robo_python, @jennylea, @_JDa_, @SaileshPatel
Visualise social data and mobile use

```
android :1238  
twitter :147  
palo :68  
facebook :373  
spotify :1886
```

Data Mining by Adam Ferguson, Jenny Lea, Sailesh Patel, James Daly
Displaying Data from the SQLite database created in the MobileMiner app.

Improvements to Mobile Miner by Finnbar Keating and Tamara Herbert
We made a list!

MacMiner/PyMine by Team Norge

Tobias Blanke
Mark Coté
Jennifer Pybus
Giles Greenway



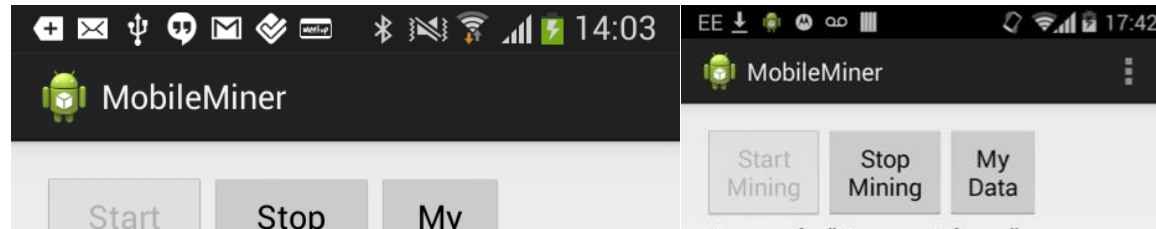
Arts & Humanities
Research Council

KING'S
College
LONDON

Research Devices

Mining Mobile Youth Cultures

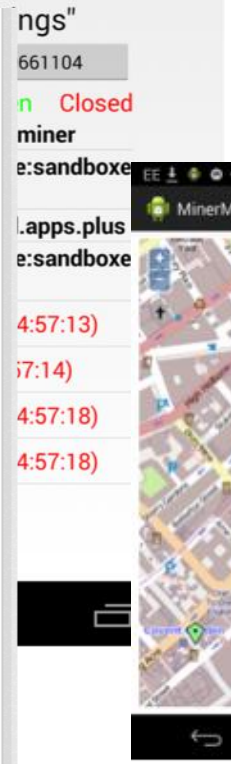
Tobias Blanke, Giles Greenway, Jennifer Pybus and Mark Cote
 Department of Digital Humanities, King's College London, United Kingdom
 Email: tobias.blanke@kcl.ac.uk, giles.greenway@kcl.ac.uk, jennifer.pybus@kcl.ac.uk, mark.cote@kcl.ac.uk



2d
 Social Data Project.
The game "TheLine" is making rather frequent use of an internet connection.

Search data ... Go » Filters

protocol	ip	port	opened	closed
tcp	7.1.1.7	3002	2014-07-09T13:25:06	2
tcp	5.4.12.10	443	2014-07-09T13:21:33	2
tcp	7.1.1.7	3002	2014-07-09T13:20:03	2
tcp	5.4.12.10	443	2014-07-09T13:11:48	2
tcp	7.1.1.7	3002	2014-07-09T12:41:47	2
tcp	7.1.1.7	3002	2014-07-09T12:11:46	2
tcp	5.4.12.10	443	2014-07-09T12:09:45	2
tcp	7.1.1.7	3002	2014-07-09T12:09:42	2
tcp	7.1.1.7	3002	2014-07-09T11:56:59	2
tcp	7.1.1.7	3002	2014-07-09T11:52:17	2
tcp	5.4.12.10	443	2014-07-09T11:46:35	2
tcp	7.1.1.7	3002	2014-07-09T11:36:44	2
tcp	6.2.12.10	443	2014-07-09T11:02:37	2
tcp	5.4.12.10	443	2014-07-09T10:46:23	2



Abstract—In this short paper we discuss our work on co-research devices with a young coder community, which help investigate big social data collected by mobile phones. The development was accompanied by focus groups and interviews on privacy attitudes and aims to explore how youth cultures are tracked in mobile phone data.

The project thus identifies BSD as a distinct subset of big data, one particularly relevant to arts and humanities researchers.

I. INTRODUCTION

Our social and cultural world is transformed through the unprecedented growth in the data we generate about ourselves. We exist in a data-centric society characterised by our information-rich environment and a 'quantified self' [1]. In order to better understand these changes, we have engaged in a project to investigate 'Big Social Data (BSD)', and the role

II. BACKGROUND — REALITY MINING

Sandy Pentland is widely recognised for his idea of using mobile phones for 'reality mining' [3]. In 2004, he and his colleagues analysed 350,000 hours of mobile phone data and have since captured in 'living laboratories' human culture on an unprecedented scale [3]. They investigated mobile phone data from the Ivory Coast to track commuter behaviour on public transport. The aim was to find patterns that would help reduce commuting time. This experiment is typical for a host of reality mining projects that discover patterns of real-life regularities in terms of mobility, lifestyle choices, opinions,



Original Research Article

Hacking the social life of Big Data

Jennifer Pybus¹, Mark Coté² and Tobias Blanke³

Abstract

This paper builds off the *Our Data Ourselves* research project, which examined ways of understanding and reclaiming the data that young people produce on smartphone devices. Here we explore the growing usage and centrality of mobiles in the lives of young people, questioning what data-making possibilities exist if users can either uncover and/or capture what they would otherwise be controlled by. We outline the MobileMiner app we created to consider how gaining access to one's own data not only augments the agency of the individual but also empowers the collective user. Finally, we discuss the data making that transpired during our hackathon. Such interventions in the otherwise closed processes of datafication are meant as a preliminary investigation into the possibilities that arise when young people are given back the data which they are normally structurally precluded from accessing.

Keywords

Data, data making, datafication, hacking, mobiles, youth

Big Data & Society
 July–December 2015: 1–10
 © The Author(s) 2015
 Reprints and permissions:
 sagepub.com/journalsPermissions.nav
 DOI: 10.1177/2053951715616649
 bds.sagepub.com



Empowering Data Citizens

Towards a Mobile Social Data Commons

Giles Greenway*, Leonard Mack†, Tobias Blanke*, Mark Cote* and Tom Heath†

*Department of Digital Humanities, King's College London, United Kingdom

†Open Data Institute, London, United Kingdom

Email: tobias.blanke@kcl.ac.uk, giles.greenway@kcl.ac.uk, mark.cote@kcl.ac.uk

Abstract—This paper discusses how born-digital cultural material can be opened up for research. We focus in particular on the grey area between private mobile phone data and its publication and use for research and beyond. We report on the results of the 'Empowering Data Citizens' (EDC) project, which is a collaboration between King's College London and the Open Data Institute. The work builds on the project Our Data Ourselves (<http://big-social-data.net/>), which studies the content we generate on our mobile devices, what we call big social data (BSD), and explores the possibilities of its ethical storage.

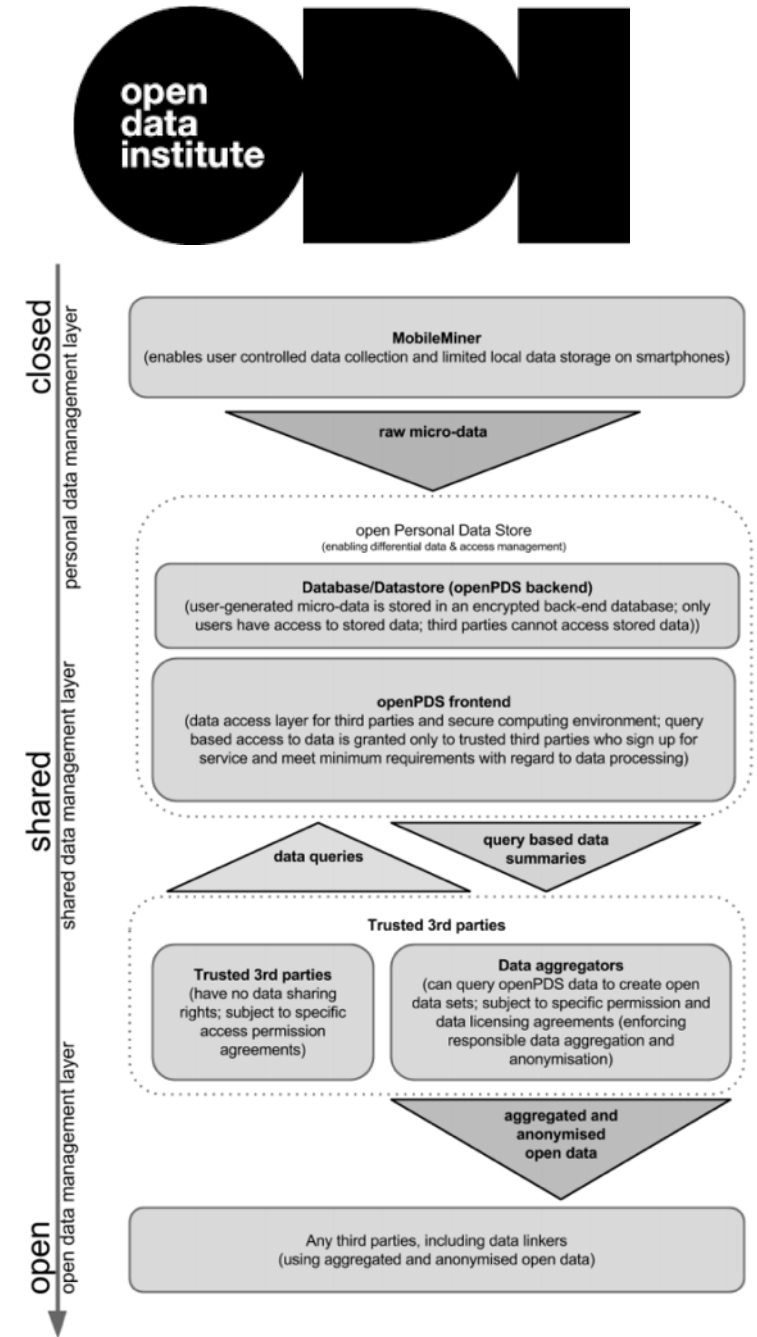
I. INTRODUCTION

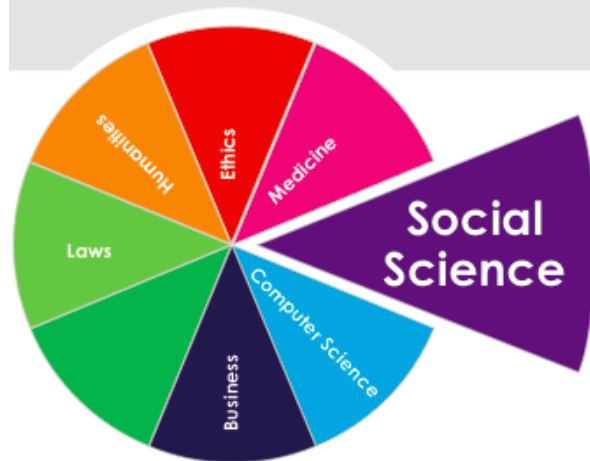
Our project addresses a basic research question: How do we transform BSD into open data, and in turn, empower the end users of mobile devices and cultivate new data communi-

anonymisation technologies for publishing cultural data. This approach will cultivate open data cultures, for example, by presenting the potential surplus from integrating it with other linked data resources such as the concept ecosystem of DBpedia [1].

The main vector of our research is in approaching born-digital cultural content via the model of open data. Open data refers to data available for anyone to use for any purpose and free of cost. Open data should be in formats that are interoperable, that is, it can be linked, and thus easily shared, in a standard and structured format for easy reuse. The key deliverable of our project is the cultivation of an ethical environment of openness for this kind of important born-digital content for cultural analysis.

Greenway, G., Mack, L., Blanke, T., Cote, M., Heath, T. 'Towards a mobile social data commons'. In *Big Data (Big Data)*, 2015 IEEE International Conference on. IEEE, pp. 1639–1642, 2015.





The members of SoBigData have a consolidated experience in users' behaviours analysis, social media analytics, human mobility analytics, graph mining, urban data analytics, digital health for personalized precision medicine.

Stakeholders



Human Mobility Analytics

Social mobility behaviours analysis
Social and demographic indicators
Public Traffic management
Public transportation planning



Text and Social Media Mining

Topic annotator and text analytics tools
Customer- and domain-specific text mining
Bespoke interactive visualisations
Graph mining
Semantic analysis



Social Network Analysis

Real-time social media monitoring and analytics
Social data aggregation and visualisation
Community discovery methods
Social networks from time series
News and financial behavior



Social Data

Social Mining for Smart Cities
Users' behaviours analysis
Social and demographic indicators
Health and well-being data analysis
Analytical Platforms for Social Mining
Ethical Data Mining
Visual Analytics for Social Mining

What SoBigData can offer to you:

The RI will take care of the legal, ethical, methodological, and infrastructural issues arising from working with social data, in order to enable data scientists to focus on research itself. It will provide access to the following key types of social data:

- Mobile and sensor data
- Social networks data
- Social media data, including Twitter, Facebook, and FourSquare content, organised into topic- and problem-specific social media virtual collections
- Mobility data, e.g. London Transport Oyster Card records and vehicular GPS trajectories
- Open social data and relevant Linked Open Data resources
- Other social data (such as one of the largest databases of Pinterest records)

Horizon 2020 SoBigData
For ethically sensitive scientific discoveries and advanced applications of social data mining to the various dimensions of social life, as recorded by 'big data'
Multi-disciplinary including digital humanities and social sciences

<http://sobigdata.eu/>

SoBigData@KCL

Lead on the training and education:

1. **Summer schools: Lipari summer schools on computational social science and London summer school on FakeNews**
2. **Datathons: NervousNet Zurich and Estonia social impact**
3. **Gender and diversity issues in data science: R ladies and Artificial Intelligence for Gender Minorities**
4. **Open teaching materials (video lectures, virtual machines, SWIRL and Notebooks)**



nervousnet

R
Ladies



LONDON

MA Big Data in Culture and Society

The screenshot shows the King's College London website. At the top left is the King's College London logo. To its right is a search bar labeled "Search website". Below the logo is a navigation menu with links: Prospective Students, Student Services, Research & Innovation, Our faculties, Giving to King's, Alumni Online, and About King's. Below this is a secondary navigation bar with links: Undergraduate, Postgraduate, More courses, International students, Why King's?, Accommodation, Student life, and Visit King's. The main content area features a large image of three students talking, with the text "Big Data in Culture & Society MA" overlaid. Below the image is a dark navigation bar with links: Overview, Entry requirements, Course detail, Locations, Career prospects, Testimonials, Fees & funding, Next steps, and a red "Apply" button.

Overview

Our Big Data in Culture & Society MA recognises the growing importance of Big Data in contemporary society and addresses the theory and practice of Big Data from an arts and humanities perspective.

Key information

Application status

Open

Duration

- i) Theorising Big Data**
- ii) Big Data in Practice: Co-laboratories, Tools & Methods**
- iii) Social and Cultural Analytics**
- iv) Big Data Law and Ethics**
- v) Dissertation**

<http://www.kcl.ac.uk/study/postgraduate/taught-courses/big-data-in-culture-and-society-ma.aspx>

Acts of digital parasitism: data, humanitarian apps and platform economies

EVENTS

DIGITAL ECOSYSTEMS OF REFUGEE MOBILITY

Location Anatomy Museum (6th Floor) King's Building Strand Campus

When 08/06/2017 (09:30-17:00)

Contact Claudia Aradau (claudia.aradau@kcl.ac.uk).

In a 2016 report on 'Connecting Refugees', the UNHCR argued that 'a lack of connectivity constrains the capacity of refugee communities to organise and empower themselves, cutting off the path to self-reliance' (UNHCR, 2016: 8). Connectivity thus articulates the simultaneous improvement of refugees' lives and the transformation of humanitarianism by ecosystem for refugees. In fostering digital humanitarian organisations, NGOs and corporate new modes of algorithmic governmentality.

This workshop proposes to explore the social these developments. According to Thomas B algorithmic governmentality is focused not so but on relations (Rouvroy & Berns, 2013). Th how datafied and digital relations emerge bot humanitarian organisations, and between not the digital ecosystems. It also proposes to co world as human/nonhuman entanglements, a through refugee practices on social media, bi communications in the refugee ecosystems.



LISIS
Laboratoire
Interdisciplinaire
Sciences
Innovations
Sociétés

INRA
SCIENCE & IMPACT

ANR
PROJET FINANCE PAR L'ANR
PROJET FUNDÉ BY THE ANR

**UP
EM**
UNIVERSITÉ
PARIS-EST
MARNE-LA-VALLÉE

**INTERNATIONAL
CONFERENCE**

**Governing by Prediction?
Models, data and algorithms
in and for governance**

Paris, Musée des Arts et Métiers,
11-13 September 2017

Data Translation Zone: KCL & Tactical Tech Community Partnership



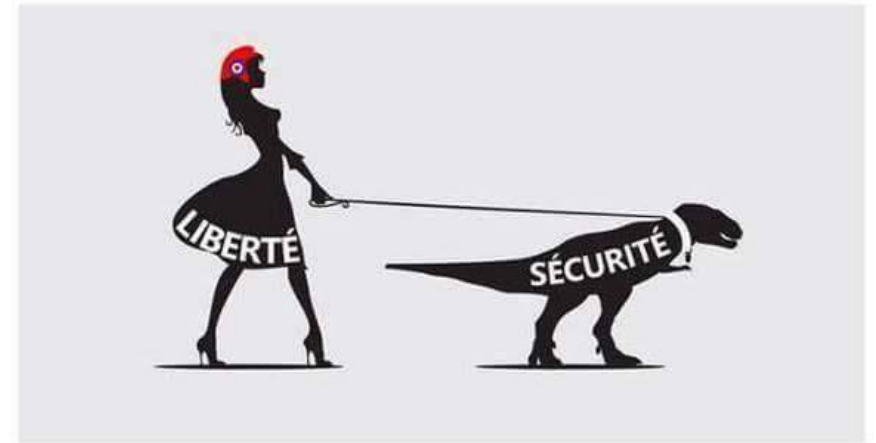
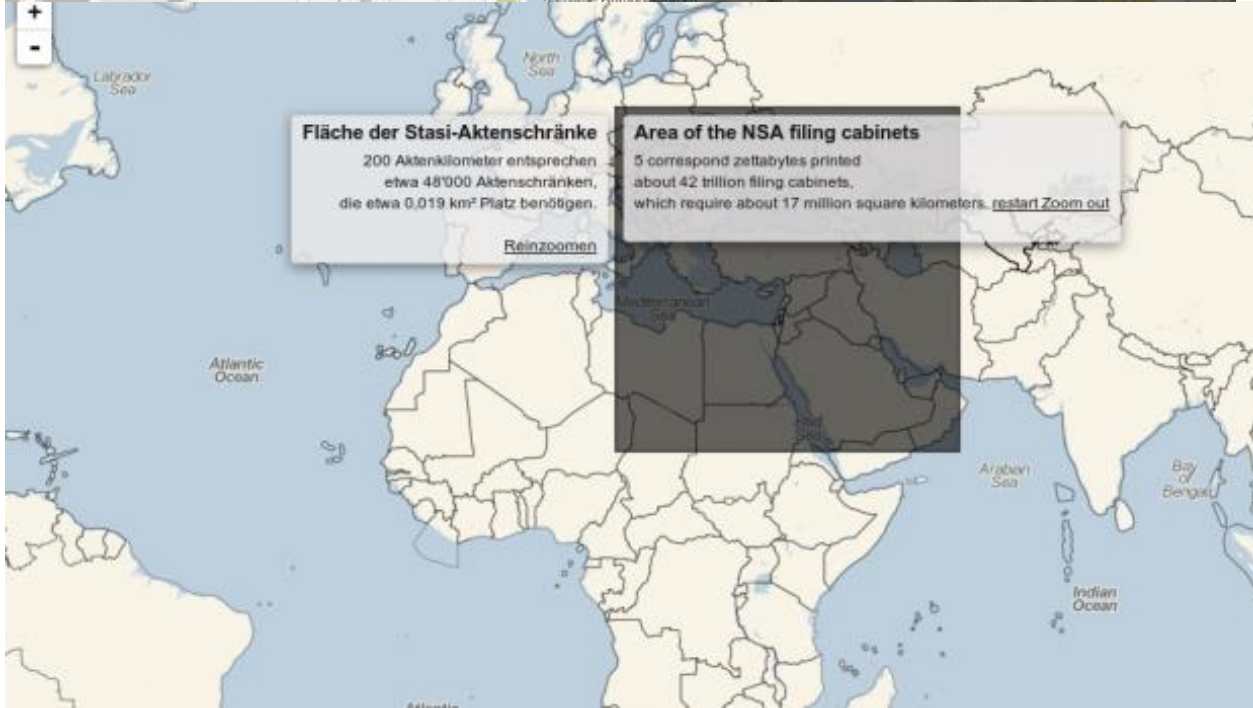
KING'S COLLEGE LONDON: HACKING THE MOBILE ECOSYSTEM

Monday, 30 October, 2017 - 2:00 pm - 5:00 pm

Location: [The Glass Room London](#)

In this workshop, led by Dr Giles Greenway of the Digital Humanities Department at King's College London, participants will analyse Android apps using a suite of tools packaged in a custom virtual machine running in Oracle VirtualBox. By downloading and extracting Android packages from the Google Play Store, they will be able to determine the third parties involved, and the permissions requested from the user. Even non-coders will be able to learn what data apps transmit by decompiling and analysing their source-code. There will also be some basic examination of live network traffic.

The Glass Room is an immersive 'tech store with a twist' that disrupts our relationship with technology and encourages visitors to make informed choices about their online life.



Big Data Security Assemblage

- **Three Dimensions**

- **Inverting the DIK hierarchy and data as a complex epistemic object**
- **Division of labour in constructing digital security devices and computers enrolled in socio-technical assemblages**
- **There are no 'unreasonably effective' algorithms. Critical approaches to security and surveillance need to engage with the methods and routine practices of Big Data-security analytics.**

- **Claudia Aradau and Tobias Blanke. "The (Big) Data-security assemblage: Knowledge and critique." *Big Data & Society* 2.2 (2015)**

The screenshot displays the SAGE Journals website interface. At the top, the SAGE Journals logo is on the left, and navigation links for 'Browse', 'Resources', and 'My Tools' are on the right. A search bar is also present. Below the navigation is a dark banner with the journal title 'Big Data & Society'. Underneath the banner is a secondary navigation bar with links for 'Home', 'Browse', 'Submit Paper', and 'About'. The main content area features the article title 'The (Big) Data-security assemblage: Knowledge and critique' by Claudia Aradau and Tobias Blanke. It includes the publication date 'First Published October 9, 2015', the article type 'Research Article', and a 'Check for updates' button. There are buttons for 'Download PDF' and 'Article information'. At the bottom of the article page, the 'Abstract' section is visible, starting with the text: 'The Snowden revelations and the emergence of 'Big Data' have rekindled questions about how security practices are deployed in a digital age and with what political effects. While critical scholars have drawn attention to the social, political and legal challenges to these practices, the debates in computer and information science have received less analytical attention. This paper proposes to take seriously the critical knowledge developed in information and computer science'.

Critique of AI decision-making

Inter-subjectivity

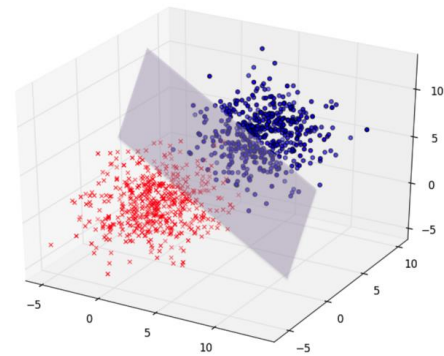
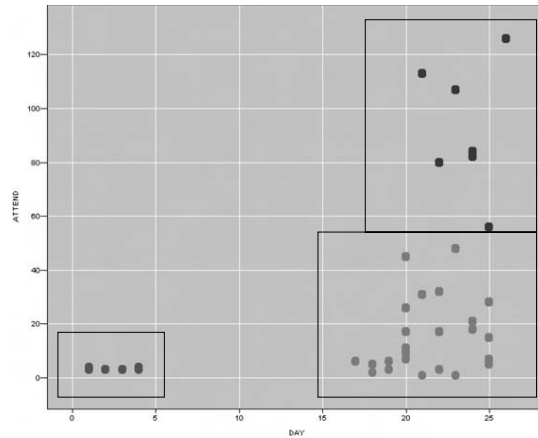


Figure 6-6 *An unsupervised learning algorithm was used to cluster the calls into similar groups. This resulted in the identification of three distinct clusters of calls, based on the number of participants and the day of the month that the conference occurred. Additional information suggested possible operational differences between the clusters. (Screenshot of Two-Step output taken by the author is from Clementine 8.5; SPSS, Inc.)*



Aradau, C. & Blanke, T., Politics of prediction: security and the time/space of governmentality in the age of big data, European Journal Of Social Theory. 20, 3, 2017.

Self and other

Understand usage of capabilities
Use Case Class Mapping



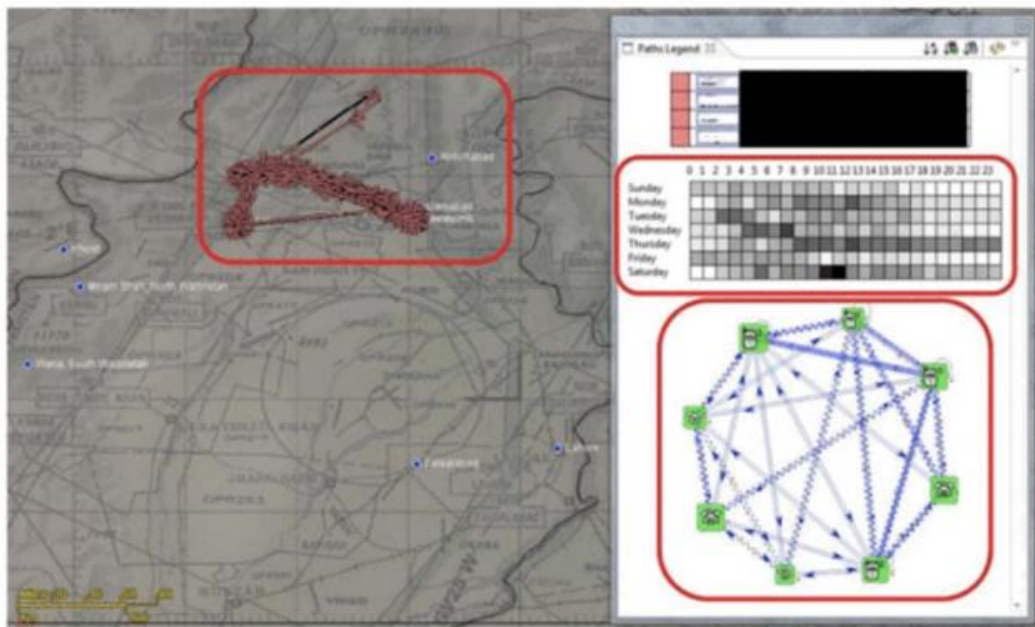
- Developed jointly between GCHQ and NSA to understand
 - the benefits of our current capabilities
 - where respective strengths and weaknesses exist
- Provides a clear set of drivers for architectural evolution
 - Missing capabilities
 - Suboptimal use of capabilities
 - Opportunities for collaboration and reuse



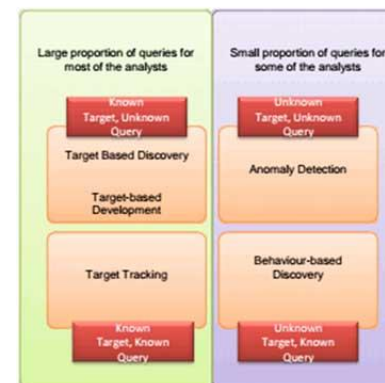
U.S. GOVERNMENT DESIGNATED PROMINENT AL JAZEERA JOURNALIST AS "MEMBER OF AL QAEDA"

Aradau, C. & Blanke, T., Governing others: Anomaly and the algorithmic subject of security European Journal of International Security, 2018

From GSM metadata, we can measure aspects of each selector's **pattern-of-life**, **social network**, and **travel behavior**



Understand usage of capabilities
Use Case Class Mapping



- Developed jointly between GCHQ and NSA to understand
 - the benefits of our current capabilities
 - where respective strengths and weaknesses exist
- Provides a clear set of drivers for architectural evolution
 - Missing capabilities
 - Suboptimal use of capabilities
 - Opportunities for collaboration and reuse

Avoid becoming an outlier