

Big Data and the Digital Human

**Big Data Research in the Department of
Digital Humanities**

Mark Cote

@markcote

Tobias Blanke

@tobias_blanke

The logo of 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 large, white, serif font, all set against a red background. Below the text are two horizontal white lines.

KING'S
College
LONDON

Digital Human

Understanding the Human Realities of the Technical Objects of Big Data

Social and cultural dimensions of *datafication*



Datafication

- process whereby the lifeworld 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

***Who* has the access and capacity to create new forms of value through the critical and creative recombination of SoBigData?**

- data as *polyvalent* resource

Techno-Cultural Method: Interdisciplinary and Applied

Explores an array of tools,
methods and practices in big
data.

Designed for non-data
scientists

Co-laboratories

Simondon on Datafication

A Techno-Cultural Method

Mark Cost, Jennifer Pybus

Abstract

This article proposes the techno-cultural workshop as an innovative method for opening up the materiality of computational media and data flows in order to increase understanding of the socio-cultural and political-economic dimensions of datafication. Building upon the critical, creative hacker ethos of technological engagement, and the collective practice of the hackathon, the techno-cultural workshop is directed at humanities researchers and social and cultural theorists. We conceptually frame this method via Simondon as a practice-led opportunity to rethink the contested relationship between the human, nature and technology, with a view to challenging social and cultural theory that ignores the human reality of the technical object. We outline an exemplar techno-cultural workshop which explored mobile apps as it is an opportunity to use new digital tools for empirical research, and to technical objects and elements for better understanding their social and cultural dimensions. We see political efficacy in the techno-cultural method not only in augmenting critical and creative agency, but as a practical exploration of the concept of data technicity, an inexhaustible relationality that exceeds the normative and regulatory utility of the data we generate and can be linked across into collective capacities to act.

Keywords: Datafication; hacker; digital materiality; data technicity.

Introduction

Hackathons have been riding the wave of the data deluge, growing from a niche event for programmers and software developers to an established techno-cultural phenomenon with foci ranging from social justice issues to corporate profit. Related but not limited to the hackathon is the subject of the hacker, defined by the 1993 Internet Users' Glossary as a person that "delights in having an intimate understanding of the internal workings of a system, computers and computer networks in particular" (Malkin and Parker 1993: 26). This paper draws on both the practice of the hackathon and subject of the hacker to ask after the data deluge, what? That

KING'S
College
LONDON

Search website

Prospective Students Student Services Research & Innovation Our faculties Giving to King's Alumni Online About King's

Undergraduate Postgraduate More courses International students Why King's? Accommodation Student life Visit King's

Big Data in Culture & Society MA

Overview Entry requirements Course detail Locations Career prospects Testimonials Fees & funding Next steps **Apply**

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

Research and Teaching

From mobiles to IoT to the Cloud

From social media to search

From social-cultural data analytics to network viz.

From Big Data to Algorithms to AI

i. Search/Planning

ii. Reasoning and Knowledge representation

iii. Perception

iv. Actuation

v. Natural Language Processing

Machine Learning

MA Big Data in Culture and Society



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>

Data and Methods of the Digital Human


Digitized Humans: Archive Graphs

Born-digital humans: Mobiles


European Holocaust Research Infrastructure: Archives as Research Data

Secure | <https://www.ehri-project.eu>

Apps | [Talis](#)



Contact Home



EUROPEAN HOLOCAUST
RESEARCH INFRASTRUCTURE



ABOUT EHRI | NEWS & EVENTS | EHRI TRAINING | EHRI PUBLICATIONS | ONLINE RESOURCES

SEARCH

LATEST NEWS

"Standing up to Scrutiny" - Successful EHRI Workshop at the United States Holocaust Memorial Museum

NEW EHRI Blog: Forced Labourers and the Water Works Camps in the Lublin District



LONDON METROPOLITAN ARCHIVES COLLECTIONS CATALOGUE

1000 YEARS OF LONDON'S HISTORY IN DOCUMENTS, PHOTOS, MAPS AND FILMS

Introduction | Search | Family History | Indexes | Subject Guides | Library | Images & Film | Online Resources | Help

LOGIN

Enter your library card number

OR

Enter your password

LOGIN

Don't have a History Card? Register here

[Forgot my password](#)

[SEARCH HISTORY](#)

HISTORY CARD REGISTRATION

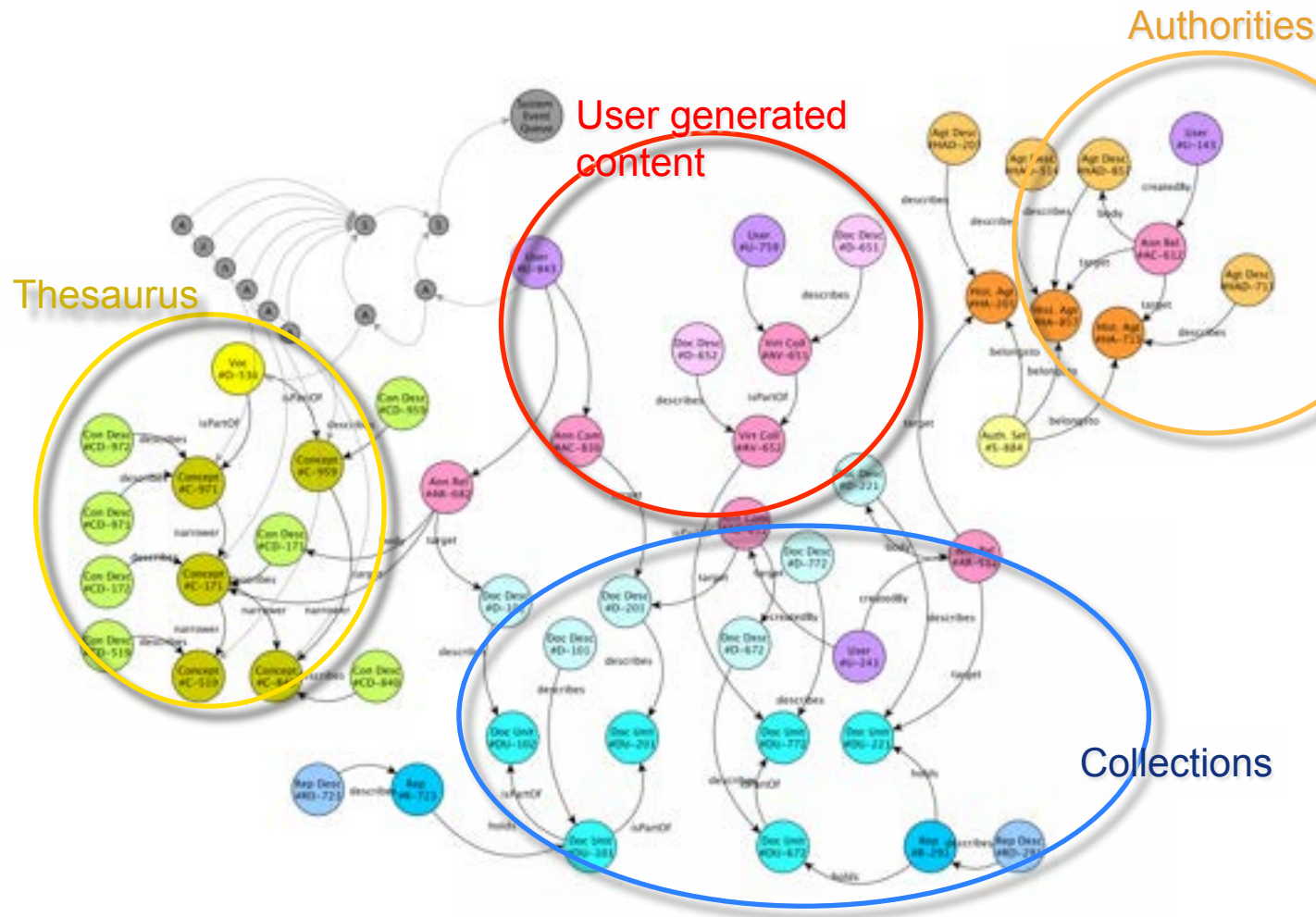
London Metropolitan Archives contains over six centuries of archives detailing the history of London and Londoners for the last seven years. To use the original documents in these collections, you must register for a City of London History Card. The registration process should take little more than five minutes to complete, and a card can be issued to you immediately by one of our staff.

WHO CAN REGISTER FOR A HISTORY CARD?

Anyone over the age of 16 can apply for a History Card.


WHY DO I NEED ONE?

Archive Graph



www.emeraldinsight.com/doi/abs/10.1108/LHT-07-2015-0070?mobileUi=0&journalCode=lht


Apps Talla




Library Hi Tech

ISSN: 0737-8831
Online from: 1983

Accepted Articles | Current Issue | Available Issues | Most Cited | Most Read | ToC Alert | RSS | Add to favorites




Clarivate Analytics
Social Sciences Citation Index (SSCI®)





Scopus®
Publish open access in this journal

[Previous Article](#)

 Save

Developing the collection graph

Article Options and Tools

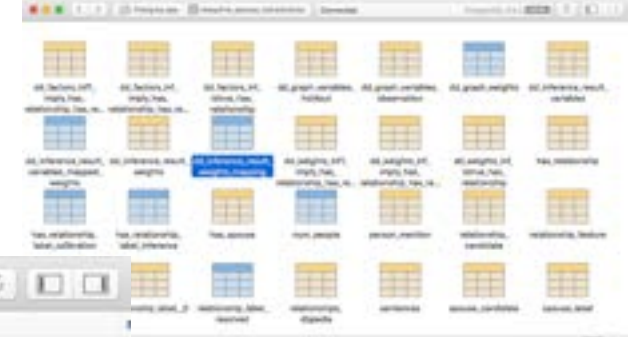
View:  HTML |  PDF

References (25) | Cited by (Crossref, 1)

[Add to Marked List](#) | [Download Citation](#) | [Track Citations](#) | [Requests & Permissions](#)

Author(s): Tobias Blanke (King's College London, London, UK)

Archives and Big Data: Complexity challenge



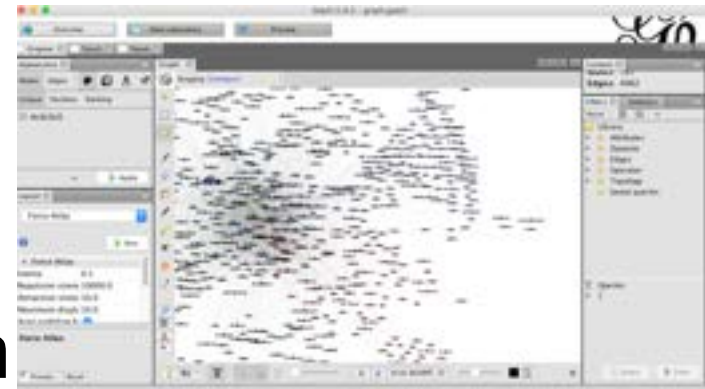
- **A deep di**
- **Ultimatel
organisa**
- **US Holoc**
- **From 120**
- **A 15-fold**

	Postgres.app	deepdrive_spouse_tobiasblanke	sentences	Connected.	PostgreSQL 9.5.2	LOGOUT
doc_id	sentence_index	sentence_text	tokens	lemmas	pos_tags	
RQ-50.010.0007_trs_en.txt	1	b' NAME: WILLIAM G. BATES INTERVIEWER: ED SHEEHY DATE: NOVEMBER 7, 1978 C...	{b,'NAME:',WILLIAM,G.,BA TES,INTERVIEWER,,ED,S...	{b,'name:,WILLIAM,G.,BAT ES,INTERVIEWER,,ED,SH...	(NNP,:NPN,NPN,NPN,NPN, NP,:NPN,NPN,NPN,:NPN,	
RQ-50.010.0007_trs_en.txt	2	I live at 2569 Windwood Court, Atlanta, Georgia 30360.	{I,live,at, 2569,Windwood,Court,,...}	{I,be,at, 2569,Windwood,Court,,...}	(PRP,VBP,IN,CD,NNP,NPN, ,"NPN","NPN,CD,...)	
RQ-50.010.0007_trs_en.txt	3	I was born September 29, 1922.	{I,was,born,September, 29,,1922,...}	{I,be,bear,September, 29,,1922,...}	(PRP,VBD,VBN,NPN,CD,"C D,...)	
RQ-50.010.0007_trs_en.txt	4	And at the time of the liberation, I was 22 years old.	{And,at,the,time,of,the,liber ation,,I,was,22,yeas,ol...	{and,at,the,time,of,the,liber ation,,I,be,22,yea,old,...}	(CC,IN,DT,NN,IN,DT,NN," PRP,VBD,CD,NNS,JJ,...)	
RQ-50.010.0007_trs_en.txt	5	My prospective profession at that time was really indefinite, because I had started colleg...	{My,prospective,profession, at,that,time,was,really,in...	{my,prospective,profession, at,that,time,be,really,infel...	(PRP \$JJ,JJ,NN,IN,DT,NN,VBD,RB...	
RQ-50.010.0007_trs_en.txt	6	At present I am a real estate regional manager for Chrysler Realty Corporation here in Atlanta.	{At,present,I,am,a,real,estat e,regional,manager,for,Ch...	{at,present,I,be,a,real,estat e,regional,manager,for,Ch...	(IN,JJ,PRP,VBP,DT,JJ,NN,J J,NN,IN,NPN,NPN,NPN,RB...	
RQ-50.010.0007_trs_en.txt	34	A: We heard some noises.	{A:,We,heard,some,noises, ...}	{a:,we,hear,some,noise,...}	(DT,:PRP,VBD,DT,NNS,...)	
RQ-50.010.0007_trs_en.txt	7	At the time of the liberation I was with the 830th Amphibian Truck Company and I was...	{At,the,time,of,the,liberat ion,I,was,with,the,830th,Amp...	{at,the,time,of,the,liberat ion,I,be,with,the,830th,Amp...	(IN,DT,NN,IN,DT,NN,PRP,V B,DT,NN,NN,NPN,NPN,NPN,NPN...	
RQ-50.010.0007_trs_en.txt	8	And in my travels in Bavaria, which is in the southern part of Germany, we liberated Dach...	{And,in,my,travels,in,Bavari a,,which,is,in,the,south...	{and,in,my,travels,in,Bavari a,,which,be,in,the,south...	(CC,IN,PRP \$NNS,IN,NPN,"WDT,VB...	
RQ-50.010.0007_trs_en.txt	9	Q: How did you first hear about these camps, Mr. Bates?	{Q:,How,did,you,first,hear, about,these,camps,,Mr...	{Q:,how,do,you,first,hear, about,these,camp,,Mr.,B...	(NPN,:WRB,VBD,PRP,RB,VB IN,DT,NNS,"NPN,NPN,...)	
RQ-50.010.0007_trs_en.txt	10	A: Actually, we had heard about the various camps, but had no idea where they were.	{A:,Actually,,we,had,heard, d,about,the,various,camp...	{a:,actually,,we,have,hea r,about,the,various,camp..."	(DT,:RB,"PRP,VBD,VBN,IN, DT,JJ,NNS,"CC,VBD,DT...	
RQ-50.010.0007_trs_en.txt	11	At this particular time we were moving a good bit and we were really advancing from a tow...	{At,this,particular,time,we,w ere,moving,a,good,bit,an...	{at,this,particular,time,we,b e,move,a,good,bit,and,we...	(IN,DT,JJ,NN,CC,PRP,VBD,VBD, DT,JJ,NN,CC,PRP,VBD,RB...	
RQ-50.010.0007_trs_en.txt	196	How do you feel about that?	{How,do,you,feel,about,tha t?,}	{how,do,you,feel,about,tha t?,}	(WRB,VBP,PRP,VB,IN,DT,...)	
RQ-50.010.0007_trs_en.txt	12	I really think it was just a matter that we happened to be passing this general directio...	{I,really,think,it,was,just,a, matter,that,we,happened...	{I,really,think,it,be,just,a, mter,that,we,happen,to,be...	(PRP,RB,VBP,PRP,VBD,RB,D T,NN,IN,PRP,VBD,TG,VB,V...	
RQ-50.010.0007_trs_en.txt	13	Q: Did your unit know about this?	{Q:,Did,your,unit,know,abo ut,this?,}	{Q:,do,you,unit,know,about this?,}	(NPN,:VBD,PRP \$NN,VB,IN,DT,...)	
RQ-50.010.0007_trs_en.txt	14	A: No, I don't think so.	{A:,No,,I,do,n't,think,so,}	{a:,no,,I,do,not,think,so,...}	(DT,:UH,"PRP,VBP,RB,VB, NN,IN,DT,NN,CC,PRP,VBD,RB...	

place, time,

Research Use Case No 1: Names and Networks. Chances of survival

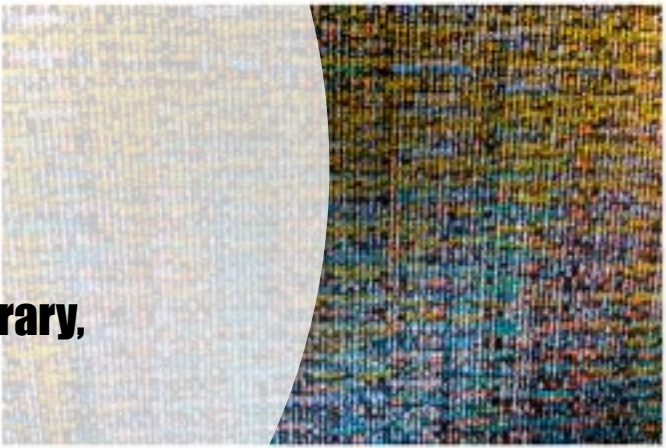
- **Aim:** Specifying which networks contributed to survival.
- **Method:** Network visualisation.
- **Sources:** Digitised archive of the Jewish Council of Amsterdam and Friesland; Digital Monument to the Jewish Community the Netherlands.



Born digital big data and approaches for history and the humanities

This AHRC-funded network will bring together researchers and practitioners from a range of stakeholder groups, to discern if there is a

- **Partners: SaS and KCL (lead)**
- **Sussex, Oxford, Warwick, Cambridge, Sheffield, National Archives, British Library, MediaLab, Waterloo, Glasgow, Leeds**



one at the Institute of Historical Research, University of London, and one at the University of Cambridge, the network will address the current state of the field; establish the most appropriate tools and methods for humanities researchers for whom born-digital material is an important primary source; discuss the ways in which researchers and archives can work together to facilitate big data research; identify the barriers to engagement with big data, particularly in relation to skills; and work to build an engaged and lasting community

SEARCH

Search... 

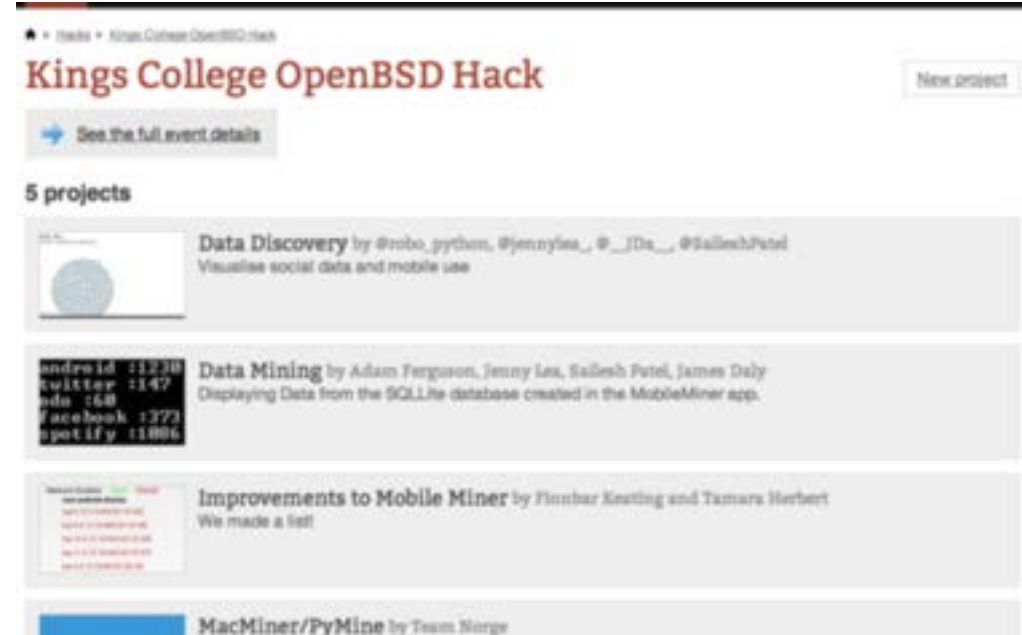
RECENT POSTS

A year of big data

FUNDED BY THE AHRC



Born digital humans: Our Data Ourselves



Tobias Blanke
Mark Coté
Jennifer Pybus
Giles Greenway



Arts & Humanities
Research Council



Research Devices & Outputs

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

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.

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, etc. from born-digital traces of human life.

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 it plays in digital culture.

Original Research Article

Hacking the social life of Big Data

Jennifer Pybus¹, Mark Cote² 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 data controllers such as Facebook monetize and share about themselves with third-parties. We outline the MobileMiner, an app we created to consider how gaining access to one's own data not only augments the agency of the individual but of the collective user. Finally, we discuss the data making that transpired during our hackathon. Such interventions in the enclosed 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

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

SOCIETY

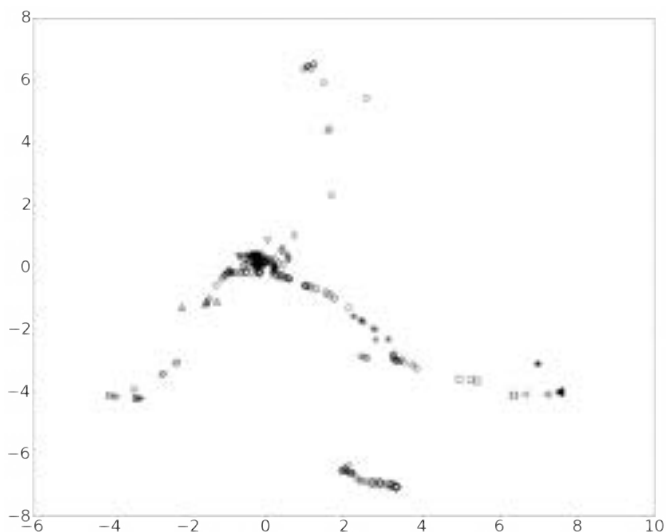
Big Data & Society
July–December 2015: 1–10
© The Author(s) 2015
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/2053951715616649
bds.sagepub.com

SAGE

The screenshot shows two parts of the research output. The top part is a mobile app interface for 'MobileMiner' with buttons for 'Start', 'Stop', and 'My'. The bottom part is a web browser view of the CKAN dataset page for 'socket'. The page includes a search bar, a 'Manage' button, and a 'Download' button. A text box states: 'The game "TheLine" is making rather frequent use of an internet connection.' Below this, the source is listed as 'MobileMiner App Data'. A table displays 21686 records with columns: _id, uid, process, protocol, ip, port, opened, and close. The table shows a list of network connections, many of which are to 'com.onetouchgame.TheLine'.

_id	uid	process	protocol	ip	port	opened	close
21685	1802305...	com.onetouchgame.TheLine	tcp	7.1.1.7	3002	2014-07-09T13:25:06	2
21686	1802305...	com.google.android.apps.plus	tcp	5.4.12.10	443	2014-07-09T13:21:33	2
21684	1802305...	com.onetouchgame.TheLine	tcp	7.1.1.7	3002	2014-07-09T13:20:03	2
21682	1802305...	com.google.android.apps.plus	tcp	5.4.12.10	443	2014-07-09T13:11:48	2
21683	1802305...	com.onetouchgame.TheLine	tcp	7.1.1.7	3002	2014-07-09T12:41:47	2
21681	1802305...	com.onetouchgame.TheLine	tcp	7.1.1.7	3002	2014-07-09T12:11:46	2
21679	1802305...	com.google.android.apps.plus	tcp	5.4.12.10	443	2014-07-09T12:09:45	2
21680	1802305...	com.onetouchgame.TheLine	tcp	7.1.1.7	3002	2014-07-09T12:09:42	2
21678	1802305...	com.onetouchgame.TheLine	tcp	7.1.1.7	3002	2014-07-09T11:56:59	2
21676	1802305...	com.onetouchgame.TheLine	tcp	7.1.1.7	3002	2014-07-09T11:52:17	2
21677	1802305...	com.google.android.apps.plus	tcp	5.4.12.10	443	2014-07-09T11:46:35	2
21675	1802305...	com.onetouchgame.TheLine	tcp	7.1.1.7	3002	2014-07-09T11:36:44	2
21673	1802305...	com.google.android.gm	tcp	6.2.12.10	443	2014-07-09T11:02:37	2
21672	1802305...	com.google.android.apps.plus	tcp	5.4.12.10	443	2014-07-09T10:46:23	2

Computer Social Networks



https://www.buzzfeed.com/craigsilverman/fake-news-real-ads?utm_term=.qxj0NA96WN#.tnl4YW73zY



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)**



Engaging the digital human

New platforms of engagement

Exporting Techno-Cultural workshop

- **Critical data examination for the rest**
- **Engage new audiences beyond the classroom and the academic workshop?**
- **How can the techno-cultural method inform the development of new collaborative spaces to facilitate innovative humanities research that can engage the general public, augmenting knowledge exchanges between experts and non-experts?**



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.

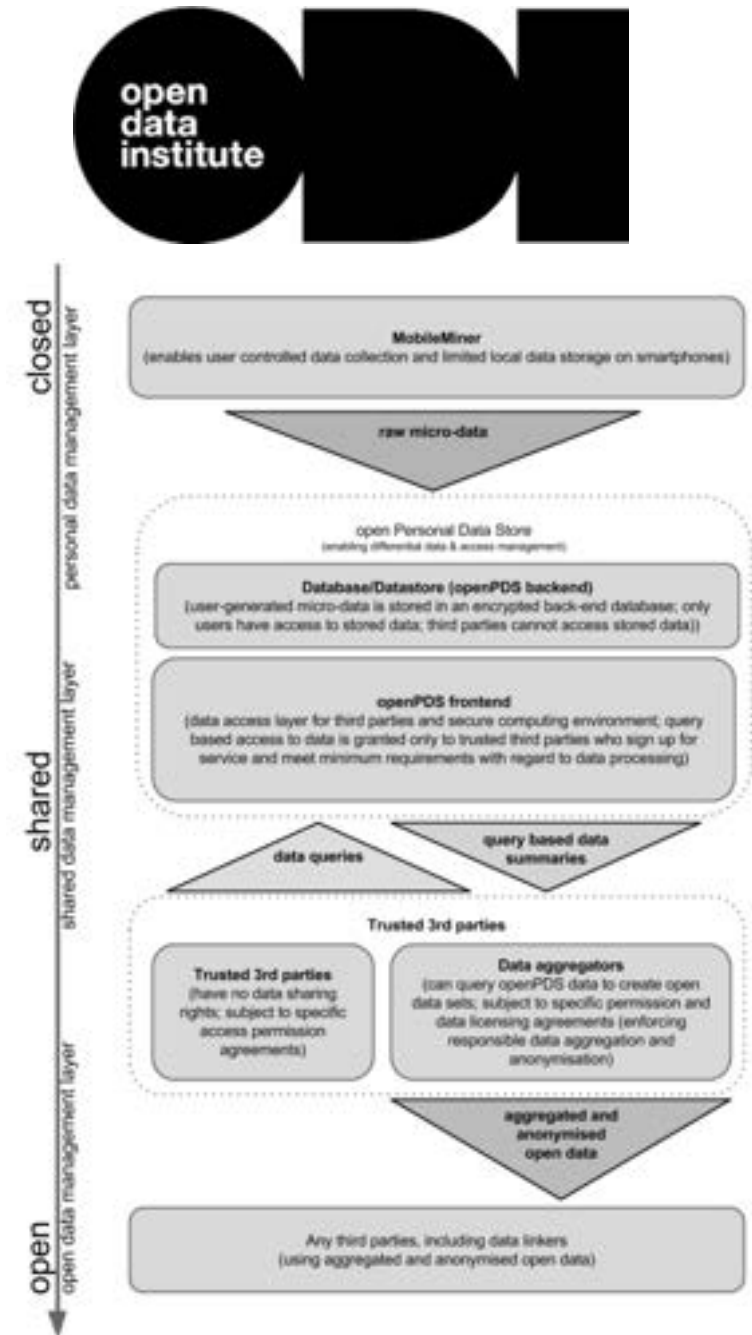
I. INTRODUCTION

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-

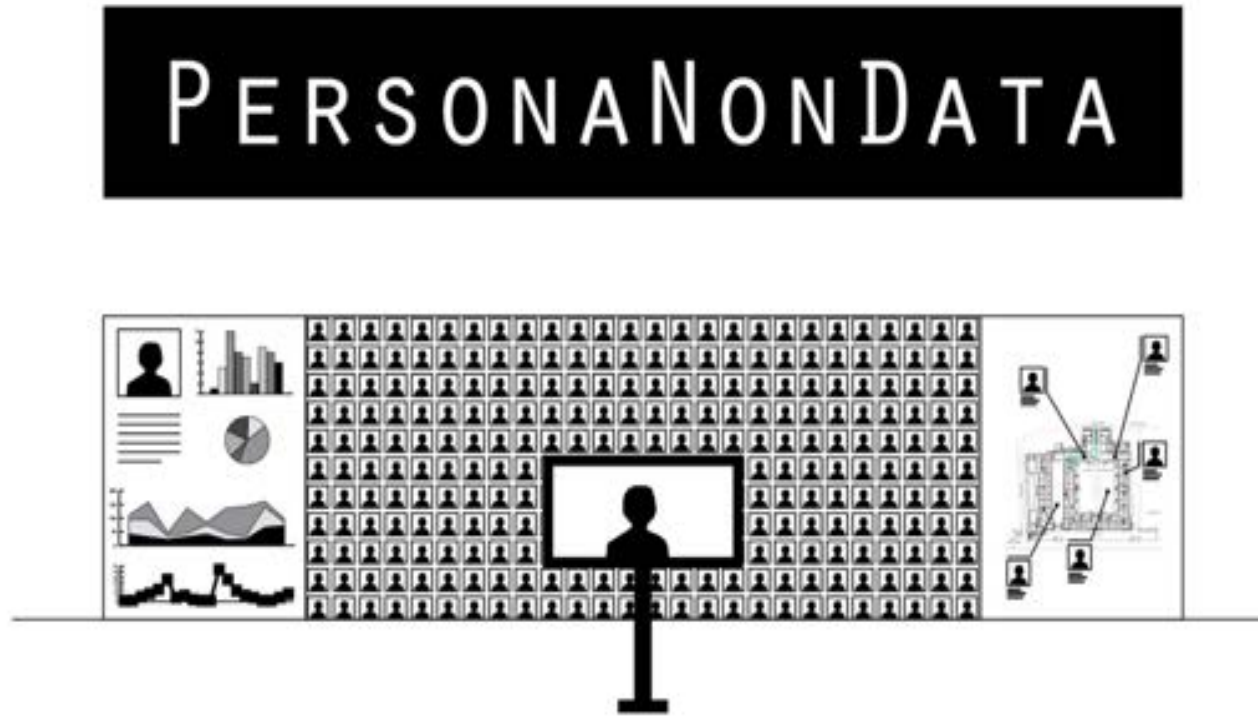
anonymisation technologies for publishing cultural data. This approach will cultivate open data cultures, for example, by presenting the potential surplus from integrating it with other linked data resources such as the concept ecosystem of DBpedia [1].

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.



Big Bang Data



Salvatore Iaconesi & Oriana Persico

January-March 2017

Somerset House Gallery

<http://artisopensource.net/persona-non-data/>



Internet of Cultural Things



Arts & Humanities
Research Council



Data Translation Zone: KCL & Tactical Tech Community Partnership



The Glass Room

Tactical Tech Collective

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

Thank you

References for all the grants

SoBigData Research Infrastructure (2015-09 to 2019-08)

H2020: [654024](#)

URL: <https://grants.uberresearch.com/501100000780/197882/SoBigData-Research-Infrastructure>

Our Data Ourselves (2013-09 to 2015-06)

AHRC: [AH/L007770/1](#)

URL: <https://grants.uberresearch.com/501100000267/828AB831-87F5-44DC-84B0-5F22769821BE/Our-Data-Ourselves>

Empowering Data Citizens (2014-08 to 2015-10)

AHRC: [AH/M002551/1](#)

URL:

<https://grants.uberresearch.com/501100000267/1C1C8D07-D739-4DB9-B973-C9D6C57163CF/Empowering-Data-Citizens>

An Internet of Cultural Things: Creative Explorations of Data in Cultural Institutions (2015-09 to 2016-09)

AHRC: [AH/M010015/1](#)

URL:

<https://grants.uberresearch.com/501100000267/EC35E997-E3D0-4CE8-BEED-DAEE7D722687/An-Internet-of-Cultural-Things-Creative-Explorations-of-Data-in-Cultural-Institutions>

References for all the publications

i) Hacking the social life of Big Data

Big Data & Society | 2015-12-01 | journal-article

- DOI: [10.1177/2053951715616649](https://doi.org/10.1177/2053951715616649)

ii) Simondon on Datafication

Digital Culture and Society | 2016-12-01 | journal-article

- DOI: [DOI 10.14361/dcs-2016-0206](https://doi.org/10.14361/dcs-2016-0206)

iii) Mining mobile youth cultures

Proceedings - 2014 IEEE International Conference on Big Data, IEEE Big Data 2014 | conference-paper

- DOI: [10.1109/BigData.2014.7004447](https://doi.org/10.1109/BigData.2014.7004447)

iv) Towards a mobile social data commons

Proceedings 2015 IEEE International Conference on Big Data, IEEE Big Data 2015 | conference-paper

- DOI: [10.1109/BigData.2015.7363932](https://doi.org/10.1109/BigData.2015.7363932)