

Can Google make us smarter?

eLearning
Symposium 2019

Professor Leslie Carr
Web & Internet Science
Web Science Institute
bit.ly/LeslieCarr

A black and white portrait of Professor Leslie Carr, a middle-aged man with a grey beard and glasses, wearing a patterned shirt and a jacket. He is looking directly at the camera. The background is slightly blurred, showing what appears to be a conference or meeting setting with other people and screens.

Professor Leslie Carr

Web & Internet Science,
Web Science Institute

**What is the impact
of the Web on
Society?**

**What is the
impact of Society
on the Web?**

bit.ly/LeslieCarr

WAIS

Web and Internet Science

WAIS researches socio-technical systems at scale using a combination of data science, computer science and social science methods. We study the interaction between the computational and human constituents of very large scale information systems.

wais.ecs.soton.ac.uk

WSI

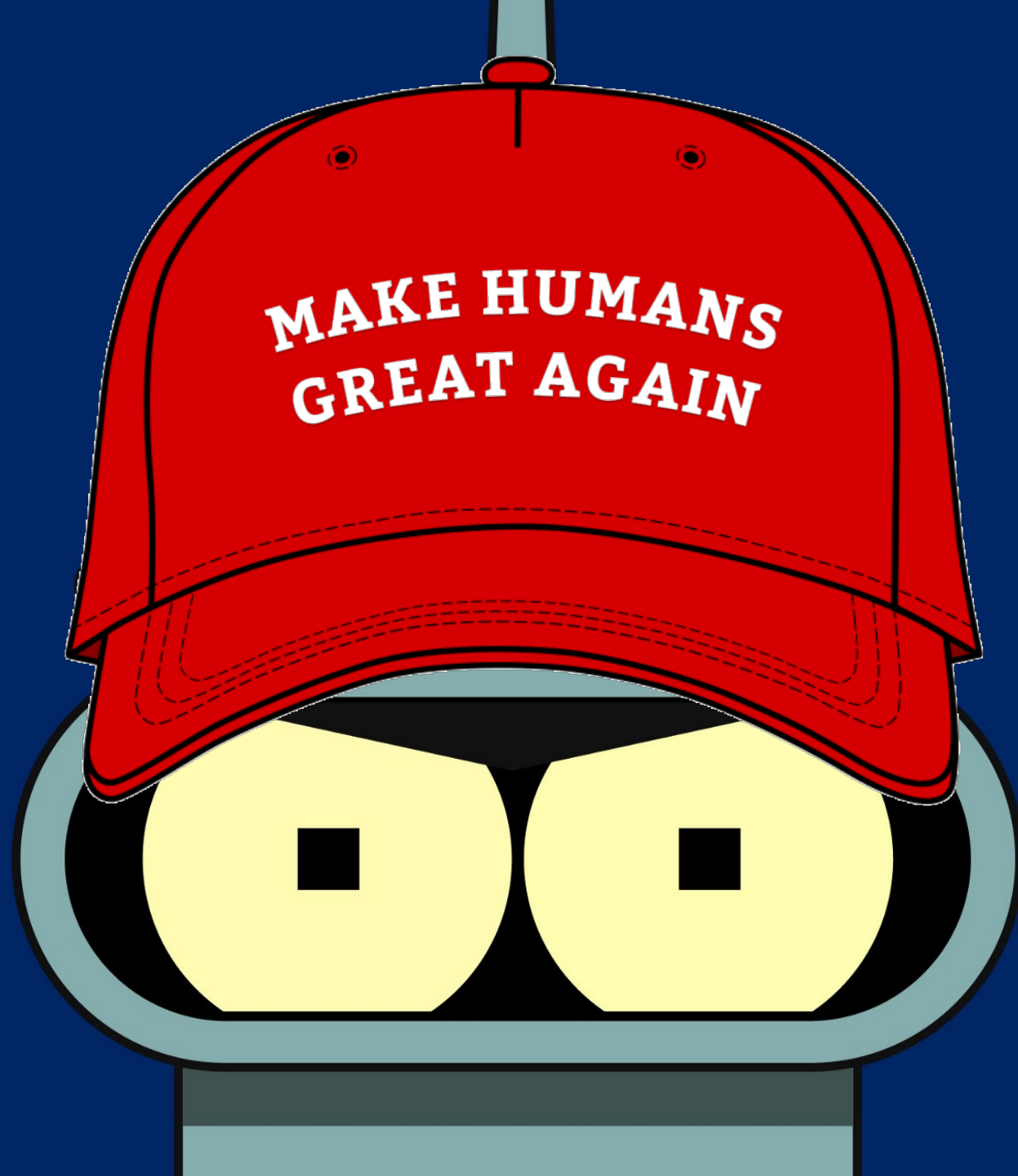
Web Science Institute

WSI coordinates the University of Southampton's globally recognised expertise on the development and social impact of Web technologies, offering analysis, tools, data and advice to government, business and civil society.

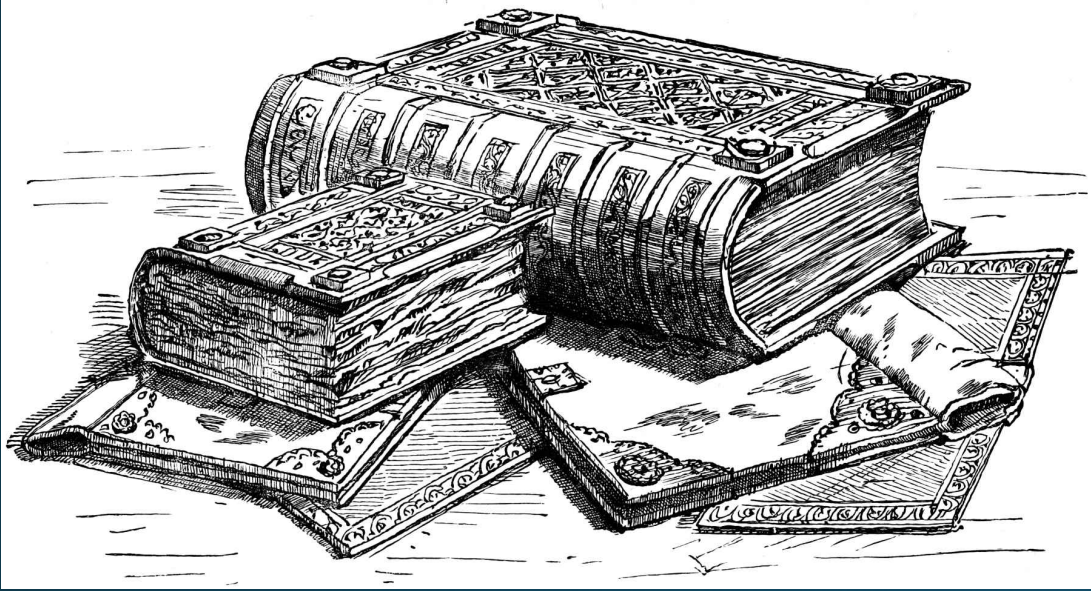
www.southampton.ac.uk/wsi

Can Google make us smarter?

- Artificial Intelligence has a very high profile over the last few years, as huge quantities of data from the Web have given computers new kinds of capability. But what about Human Intelligence? Google is building smart cars that understand the roads, but can Google make smart humans that understand the world?
- In this talk I will look at some of the ways that we have tried to make computers help us to be more intelligent, from Indexing the Internet, to Webs of Semantics, to Data Storytelling.



Internet & Web



- The Domesday book (1086).
A list of everyone, what they do and what they have. For tax purposes.



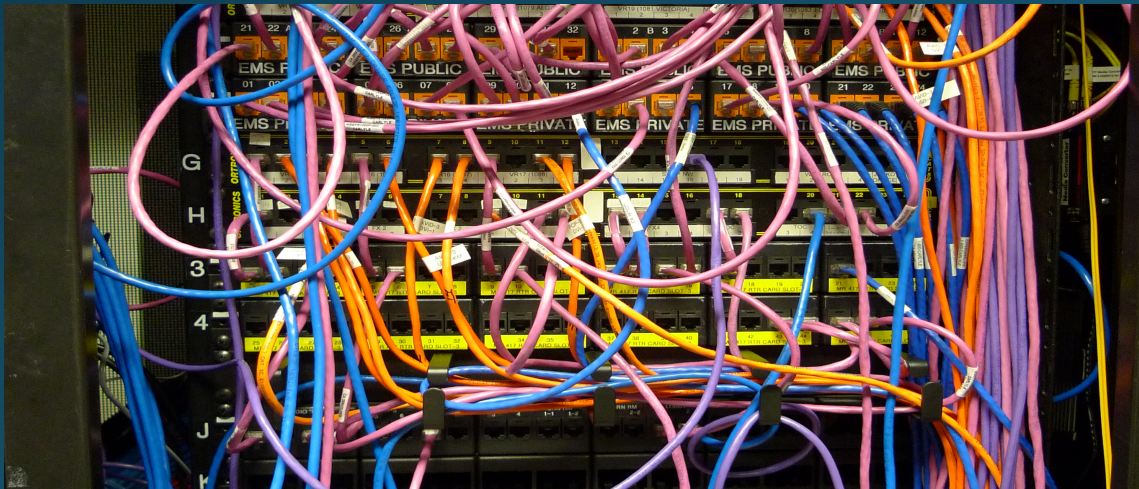
- The Internet (2019).
A list of everyone, what they do and what they have. For advertising.

Historic Attempts at Webs

Sponsor	System	Scope	Real	Date	Important Properties
Finance / Press	Reuters	Professional, centralised		1850	News & stock information (originally carrier pigeon and subsequently telegraph)
Private Institution	Mundaneum	Public, centralised		1920	Based on indexing technology (the library card)
Military	Memex	Scholarly, individual, centralised	X	1945	Aimed at Scientists and Technologists in WWII
Media	Xanadu	Public, decentralised	X	1960	Focused on DRM, reuse and writing for “creatives”
Media	CEEFAX	Public, national, centralised		1970	Broadcast, linked, not participatory
Government	Minitel	Public, national, centralised		1980	Commercial services and information
Academy (CS & HEP)	FTP / Archie / Anarchie	Public, decentralised		1985	Download resources (papers, reports) to hard drives and print them on LaserWriters.
Commerce	Hypercard, HyperTIES	Private, centralised		1988	Personal applications, sometimes tied to multimedia resources on CDROMs / video disks
Academy (HEP)	WWW	Public, global, decentralised		1990	Universal naming, linking, interoperability, participative. No writing & indexing.
Academy (CS)	Microcosm	Private, centralised		1990	Sophisticated linking and openness for personal information stores
Academy (CS)	HyperG	Public, centralised		1990	Extension of Web for with support for writing, indexing and consistency management.
Commerce	AOL, CompuServ	Public, centralised		1990	Dialup walled garden access to email, forums, chat rooms and information resources

The Internet

- A network of computers and routers and cables and clients and servers.
- In fact, lots of networks all networked together. An inter-net.
- A worldwide network of computers sending and receiving messages



Vint Cerf, "father of" the Internet

World Wide Web

- The “inventor of” the Web
 - URL, HTTP, HTML
 - Browsing
- A network of interconnected documents built on a network of interconnected servers by a network of interconnected people.



Tim Berners-Lee,
“inventor of” the Web



How New Knowledge Tech Grows



Siri, Alexa

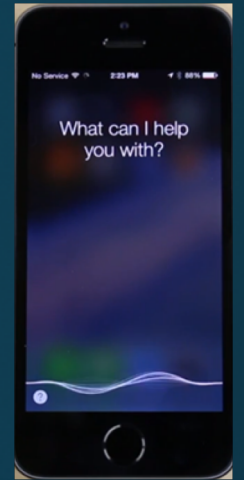
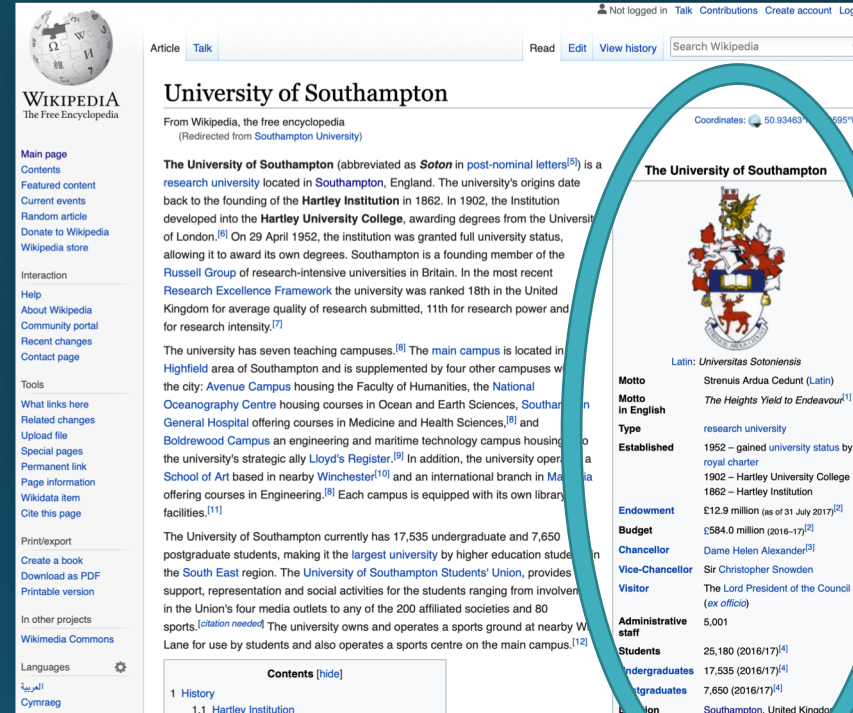
DBpedia

Wikipedia

Wiki

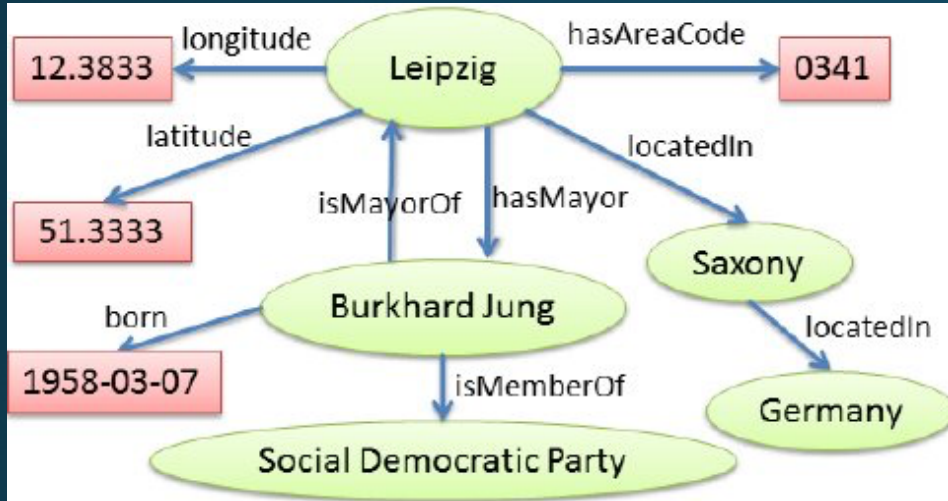
Web

Internet



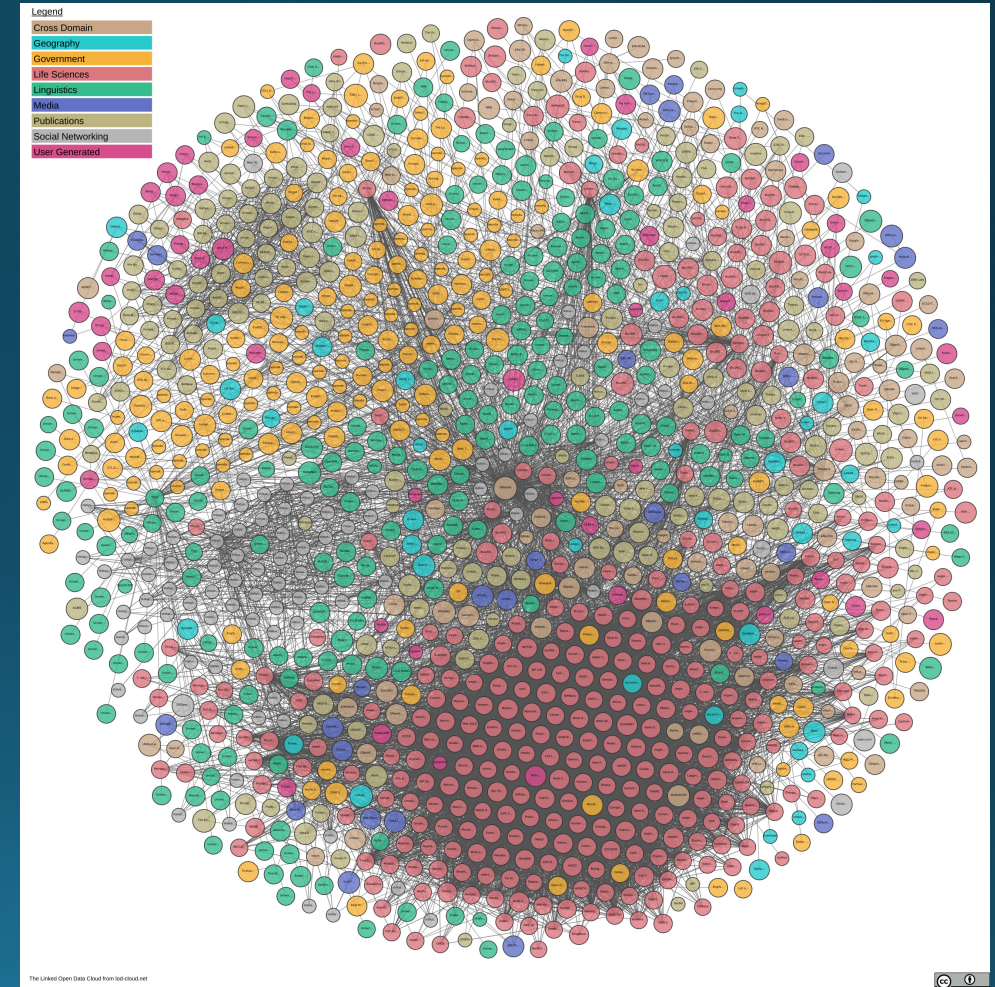
- **Vertical generativity** means allows you to pile platforms on platforms on platforms to achieve ever more sophisticated applications
 - But it requires 'open' licenses across generations of platforms.

Semantic Web



```

<bibo:Journal rdf:about="http://aims.fao.org/serials/c_249cf355">
  <bibo:ISSN>0028-4793</bibo:ISSN>
  <dc:title>New England journal of medicine</dc:title>
  <dc:alternative>The New England journal of medicine</dc:alternative>
  <ags:publisherPlace rdf:resource="http://aims.fao.org/aos/geopolitical.owl#United_States"/>
  <ags:publisherName>Massachusetts Medical Society.</ags:publisherName>
  <dc:language>eng</dc:language>
  <dc:issued>1928</dc:issued>
  <foaf:homepage>http://www.nejm.org/content/index.asp</foaf:homepage>
  <dc:accrualPeriodicity>weekly</dc:accrualPeriodicity>
  <dc:subject rdf:resource="http://aims.fao.org/aos/agrovoc/c_10394"/>
  <dc:subject rdf:resource="http://dewey.info/class/610/about"/>
  <bibo:coden>NEJMAG</bibo:coden>
</bibo:Journal>
  
```

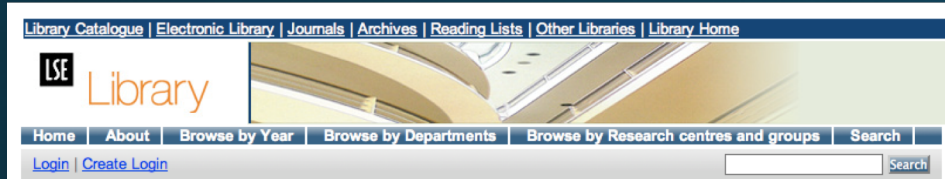


Open Information

- Open Access to Research Outputs
 - Supported Research Funders all over the world
- Open Research Data
 - Demanded by UK government, private funders & UK Research Councils
- Open Educational Resources
 - From closed VLEs to sharing resources and MOOCs
- Open Government Data
 - Supported by G8



Open Access Research: EPrints



Welcome to LSE Research Online

Welcome to LSE Research Online. A collection of research produced by LSE academics: articles, working papers, book chapters, conference papers and more.

Atom RSS 1.0 RSS 2.0

About

To deposit your research in LSE Research Online, you can email it to lseres@lse.ac.uk or you can deposit items directly.

Latest Additions

View items added to the repository.

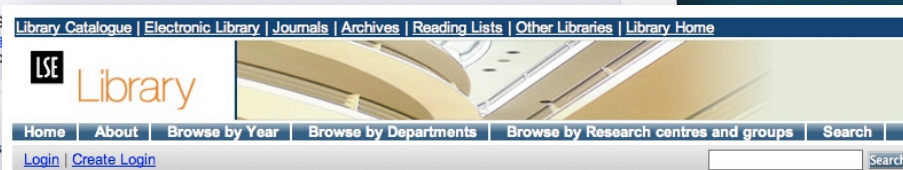
Search Repository

Search the repository using a search engine.

Browse Repository

Browse the items in the repository.

LSE Research Online supports Open Access.



Supporting people with AIDS and their carers in rural South Africa: Possibilities and challenges

Campbell, Catherine and Nair, Yugi and Maimane, Spongile and Sibiya, Zweni
carers in rural South Africa: Possibilities and challenges. *Health & place*

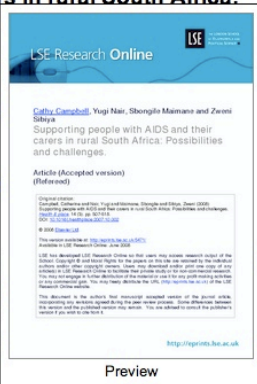


PDF - Requires a PDF viewer such as [GSview](#), [Xpdf](#) or [Adobe Reader](#).
215Kb

Official URL: <http://www.sciencedirect.com/science>

Abstract

Under-served rural areas—home to over half of people in sub-Saharan Africa—have a case study of the existence and quality of support networks available to people in African rural area. Drawing on 45 interviews and 13 focus groups, we identify key agencies in civil society, the public and the private sectors. The most effective agencies are faith-based organisations, traditional healers and local leaders, a private sector agencies. We discuss ways in which existing and latent network



Preview

Item Type: Article

Additional Information: © 2008 Elsevier Ltd

Rights: <http://www.lse.ac.uk/library/rights/LSERO.htm>

Uncontrolled Keywords: HIV/AIDS; Rural areas; Carers; Home-based care; Primary care; Support networks; Facilitation

Library of Congress subject classification: [R Medicine > RA Public aspects of medicine > RA0421 Public health. Hygiene. Preventive Medicine](#)
[H Social Sciences > HV Social pathology. Social and public welfare](#)

Institutional repository for the London School of Economics

<http://eprints.lse.ac.uk/>

One of 145 UK Institutional Research repositories



Open Govt Data data.gov.uk (8,300 data sets)



data.gov.uk^{BETA}
Opening up government

1 Data Apps Consultation Forum Blogs Ideas Linked Data Resources About

FCO Travel Advice

Posted by Foreign and Commonwealth Office on 17/02/2011 | 2 comments | ★★★★★

Foreign & Commonwealth Office (FCO) travel advice for you to re-use in your own applications, widgets and content. Designed by Directgov Innovate for developers or content aggregators to build feeds of travel advice, get embassy contact information and FCO travel news through a RESTful, content-negotiated interface. Any development platform with an HTTP stack or any widget generator that takes an ATOM feed will be able to use the content.

► Overview

▼ Resources (download)

Download  FCO Travel Advice

► Details

► Contact information

Comments (2) [Add new comment](#)

 **How ?**
Tue, 08/03/2011 - 12:24 — Anonymous
How would I integrate this into my travel insurance site ?
Thanks in advance
[reply](#)

 **Take a look at what how I**
Tue, 15/03/2011 - 20:47 — michaelg
Take a look at what how I have used the data here
<http://www.ebusinessconsultant.co.uk/foreign-travel-advice/>

Share this   

Related apps

Travel Tools Dashboa...
by Andy Wrigley | ★★★★★
A Travel Dashboard providing a huge range of tools for independent and business...

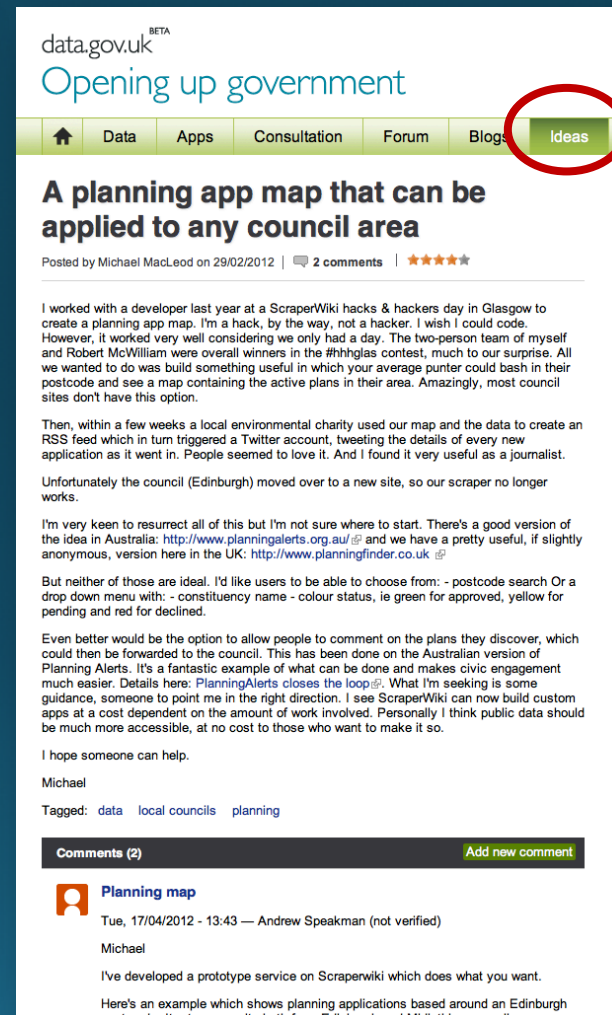


Travel Information b...
by Andrew Wrigley | ★★★★★
A widget for travel websites, providing the latest Travel Advice by Country from drop down...



Have you used our data to create an application?
Let us know...
[Share your app](#)

Have you got a good idea for the data?
We can help you find a developer...



data.gov.uk^{BETA}
Opening up government

Home Data Apps Consultation Forum Blogs Ideas

A planning app map that can be applied to any council area

Posted by Michael MacLeod on 29/02/2012 | 2 comments | ★★★★★

I worked with a developer last year at a ScraperWiki hacks & hackers day in Glasgow to create a planning app map. I'm a hack, by the way, not a hacker. I wish I could code. However, it worked very well considering we only had a day. The two-person team of myself and Robert McWilliam were overall winners in the #hhglas contest, much to our surprise. All we wanted to do was build something useful in which your average punter could bash in their postcode and see a map containing the active plans in their area. Amazingly, most council sites don't have this option.

Then, within a few weeks a local environmental charity used our map and the data to create an RSS feed which in turn triggered a Twitter account, tweeting the details of every new application as it went in. People seemed to love it. And I found it very useful as a journalist.

Unfortunately the council (Edinburgh) moved over to a new site, so our scraper no longer works.

I'm very keen to resurrect all of this but I'm not sure where to start. There's a good version of the idea in Australia: <http://www.planningalerts.org.au/> and we have a pretty useful, if slightly anonymous, version here in the UK: <http://www.planningfinder.co.uk>

But neither of those are ideal. I'd like users to be able to choose from: - postcode search Or a drop down menu with: - constituency name - colour status, ie green for approved, yellow for pending and red for declined.


Even better would be the option to allow people to comment on the plans they discover, which could then be forwarded to the council. This has been done on the Australian version of Planning Alerts. It's a fantastic example of what can be done and makes civic engagement much easier. Details here: [PlanningAlerts closes the loop](#). What I'm seeking is some guidance, someone to point me in the right direction. I see ScraperWiki can now build custom apps at a cost dependent on the amount of work involved. Personally I think public data should be much more accessible, at no cost to those who want to make it so.

I hope someone can help.

Michael

Tagged: [data](#) [local councils](#) [planning](#)

Comments (2) [Add new comment](#)

 **Planning map**
Tue, 17/04/2012 - 13:43 — Andrew Speakman (not verified)
Michael
I've developed a prototype service on Scraperwiki which does what you want.
Here's an example which shows planning applications based around an Edinburgh postcode. It returns results both from Edinburgh and Midlothian councils.

Bespoke apps produced by independent developers for government data

Shared ideas: Michael built a 'planning proposals' web app. Some turned it into an RSS feed. Someone else turned that into a Twitter feed. Someone else used scraperwiki...

Why Open Knowledge is Vital

Increasing cognitive capabilities

Evaluation

Did Germany make the right choice in relocating their post-unification capital from Bonn?

Synthesis

Create a set of guidelines for choosing alternative national capitals in a post-climate change Europe.

Analysis

What were the overriding factors in the choice of sites for the European parliament and commission?

Application

Choose the most appropriate capital for the Isle of Wight

Comprehension

Why is Paris the capital of France?

Knowledge

What is the capital of France?

Bloom's Educational Taxonomy demonstrates why closed systems are bad – human understanding is built on top of successively better knowledge platforms.



The Turing Test

- What's the difference between a computer and a person?

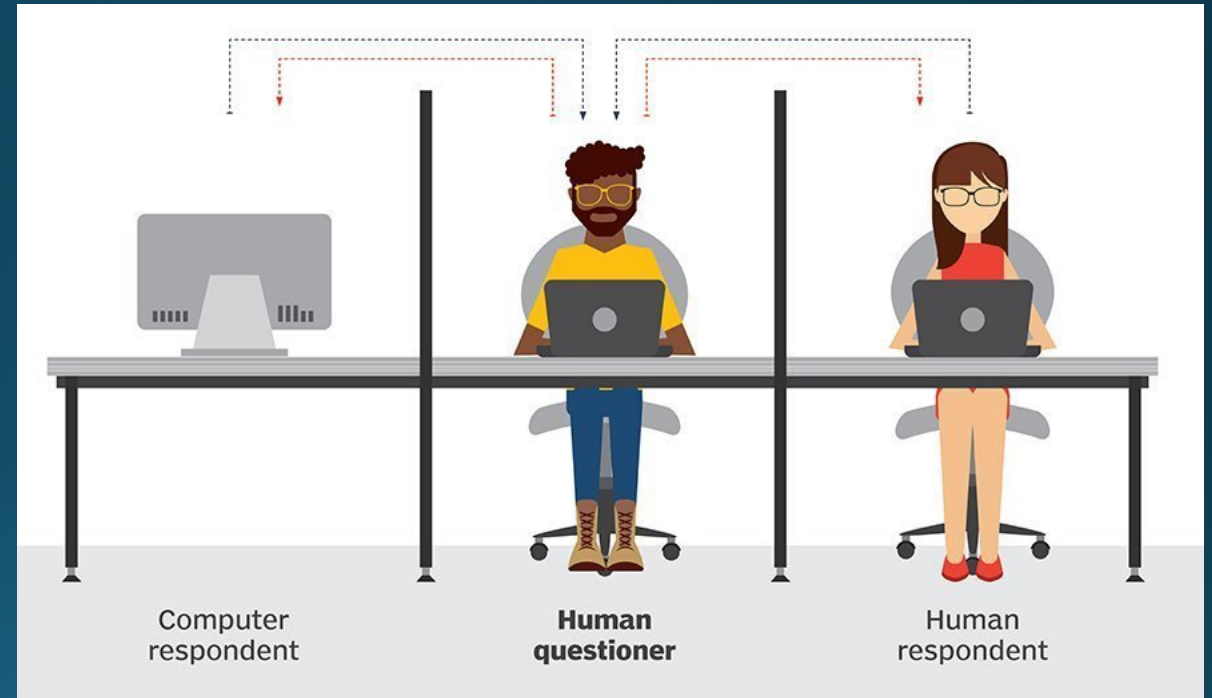
A. M. Turing (1950) *Computing Machinery and Intelligence*. *Mind* 49: 433-460.

COMPUTING MACHINERY AND INTELLIGENCE

By A. M. Turing

1. The Imitation Game

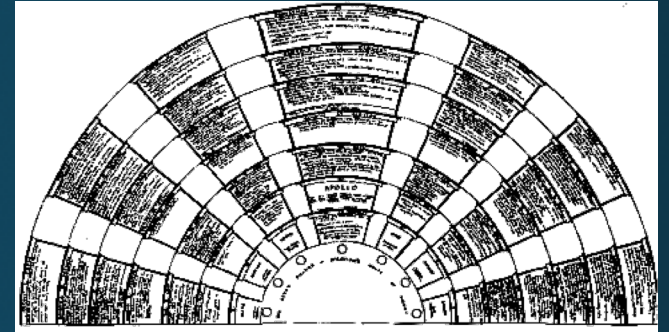
I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.



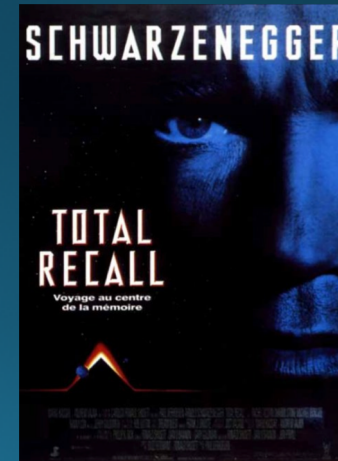
- Can you tell the difference?

Personal Knowledge, Recall and Memory

- The "memory theatre" was an aspect of a science of the imagination which was practiced from Classical times up to the Renaissance.
- Writing, books, CDs, Web extend our cognitive storage capacity
 - at some reduction in access time
- Can you get away with not internalising something at all?
 - If your environment has sufficient smarts
 - Google



Guillio Camillo's Memory Theatre 1550



Philip K Dick
We Can Remember it For You Wholesale

Who Wants to Cheat at **MILLIONAIRE**



- How well can you perform if one of your WWTBAM lifelines is 30-seconds on Google?

- How well can you perform in an Internet Open Book Exam?


Sports Entertainment Business Features Opinion Tech

< freep / tech >

Heather Newman: The right search words can prove to be a lifeline

April 25, 2001

Steve Perry of Charlottesville, Va., had won \$500,000 on the Sunday broadcast of "Who Wants to be a Millionaire" and had all three of his lifelines left.



His question, for \$1 million: What was Carol Brady's maiden name on "The Brady Bunch"?

He asked the audience. He eliminated two wrong answers (which wiped out the audience's top two choices). Then he made a fateful decision: He phoned a friend who was supposed to look up the answer on the Internet.

"C-a-r-o-l B-r-a-d-y m-a-i-d-e-n n-a-m-e," he spelled out for his buddy. The friend burned some time by asking him to respell a couple words, but then there was silence.

"Oh, man," the friend said as the 30-second timer ticked.

"I'm checking," he said.

Then the clock ran out.

UAI 2003 LAM ET AL. 337

1 Billion Pages = 1 Million Dollars? Mining the Web to Play "Who Wants to be a Millionaire?"

Shyong (Tony) K. Lam
Computer Science Dept.
University of Minnesota
Minneapolis, MN 55455
lam@cs.umn.edu

David M. Pennock
Overture Services, Inc.
74 N. Pasadena Ave., 3rd floor
Pasadena, CA 91101
david.pennock@overture.com

Dan Cosley
Computer Science Dept.
University of Minnesota
Minneapolis, MN 55455
cosley@cs.umn.edu

Steve Lawrence
NEC Laboratories America
Princeton, NJ 08540
lawrence@necmail.com

Abstract

We exploit the redundancy and volume of information on the web to build a computerized player for the ABC TV game show "Who Wants to Be a Millionaire". The player consists of a question-answering module and a decision-making module. The question-answering module utilizes question transformation techniques, natural language parsing, multiple information retrieval algorithms, and multiple search engines; results are combined in the spirit of ensemble learning using an adaptive weighting scheme. Empirically, the system correctly answers about 75% of questions from the *Millionaire* CD-ROM, 3rd edition—general-interest trivia questions often about popular culture and common knowledge. The decision-making module chooses from allowable actions in the game in order to maximize expected risk-adjusted winnings, where the estimated probability of answering correctly is a function of past performance and confidence in correctly answering the current question. When given a six question head start (i.e., when starting from the \$2,000 level), we find that the system performs about as well on average as humans starting at the beginning. Our system demonstrates the potential of simple but well-chosen techniques for mining answers from unstructured information such as the web.

[5] are formal enough to be solvable in principle, though are far from trivial to master in practice due to exponential size search spaces. In chess, checkers, and backgammon, current machine players rival their best human competitors. Recently, attention has turned to less structured game environments, like crossword puzzles [16], video games [30], and soccer [29], where game states, actions, or both are not easily enumerable, making a pure search formulation unnatural or impractical.

"Who Wants to be a Millionaire?" is a trivia game where actions are enumerable, though competence depends on the ability to answer general-interest questions—often requiring common sense or knowledge of popular culture—and to make decisions based on confidence, expected reward, and risk attitude. True human-level competence at *Millionaire* will likely require excellence in natural language processing and common sense reasoning. We present a first-order system that exploits the breadth and redundancy of information available on the World Wide Web to answer questions and estimate confidence, and utilizes a decision-theoretic subsystem to choose actions to maximize expected risk-adjusted payoffs.

2 RELATED WORK

2.1 QUESTION ANSWERING

A large body of research exists on question answering. For example, see the Question-Answering Track [32] of the Text Retrieval Evaluation Conference (TREC). Systems in this track compete against each other to retrieve short (50 or 250 byte long) answers to a set of test questions.

Question-answering systems typically decompose the problem into two main steps: retrieving documents that may contain answers, and extracting answers from these documents. For the first part of the task, retrieving a set of promising documents from a collection, the systems in the TREC QA track submitted the original questions to various information retrieval systems [32].

A number of systems aim to extract answers from documents. For example, Abney et al. [1] describe a system in

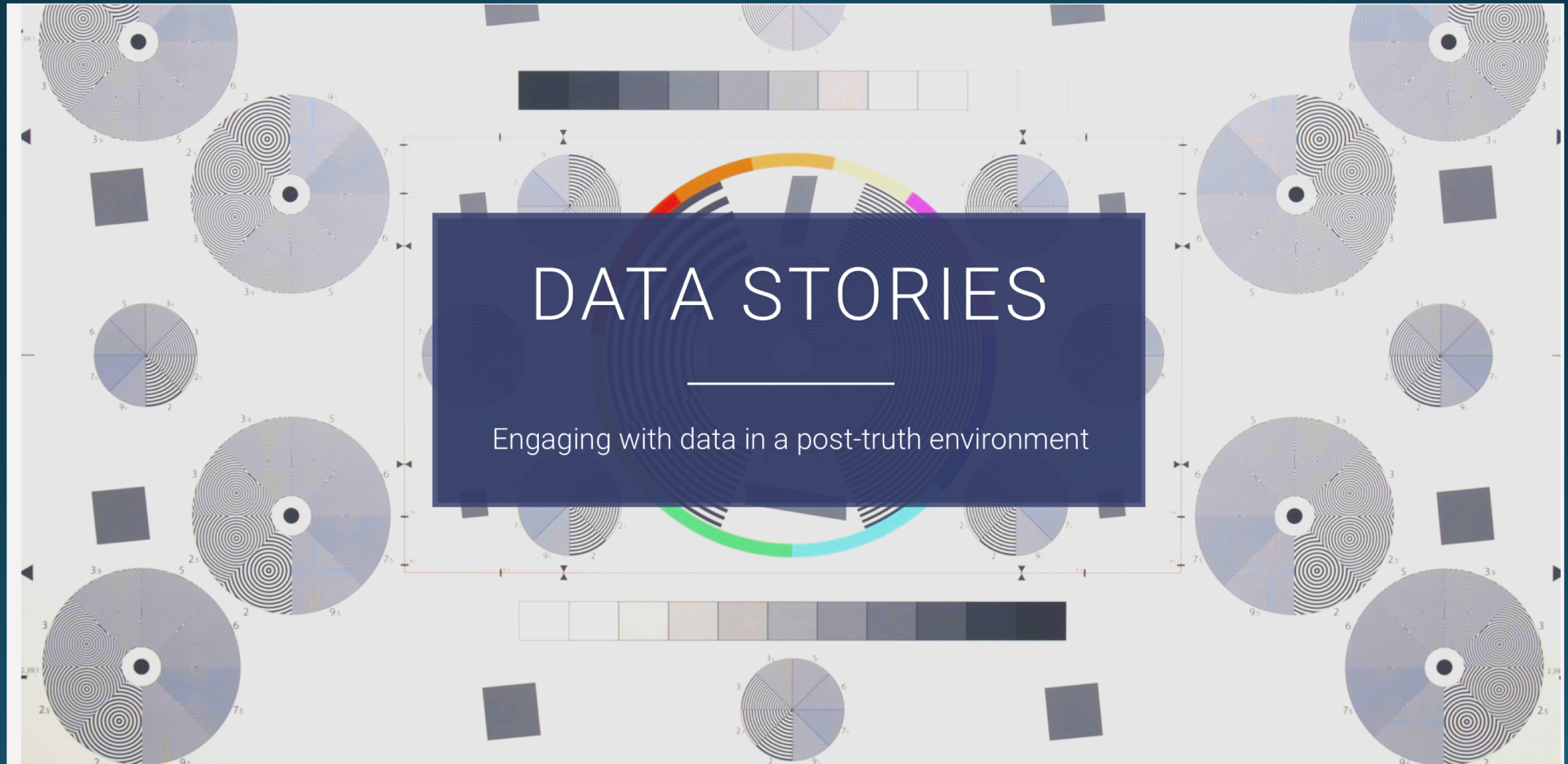
1 INTRODUCTION

Machine competence in games has long served as a benchmark for progress in artificial intelligence (AI). While we seem hardly close to building systems capable of passing a full-blown Turing Test, machine excellence in a growing number of games signals incremental progress. Games such as chess [13], checkers [27], Othello [7, 18], and Go

This work conducted at NEC Laboratories America, Princeton, NJ.

See http://www.angelfire.com/ky3/nellylunaticsfan/atgs_wwtbam/perry.html

How is Knowledge Being Used?



“One of the interpretations of the EU referendum result and the rise of Donald Trump in the US is that we are now living in a **post-truth** society - a world in which anecdotes shared on **social media** and invented numbers thrown on the sides of buses are more trusted and influential than official **statistics**, extensive **research**, and proven **expertise**. In this world, scientists, **statisticians**, **analysts**, and **journalists** must find new ways to bring hard, factual data to **citizens.**”

“**Data must entertain as well as inform, excite as well as educate.** It must be built with social media sharing in mind, and become part of our everyday activities and digital interactions with others.”



Data Stories looks at **frameworks** and **technology** to bring data closer to people through **art, games,** and **storytelling**

It examines the impact that varying levels of **localisation, topicalisation, participation,** and **shareability** have on the engagement of the public with factual evidence.

It delivers **tools** and **guidance** for **communities** and **civic groups** to achieve wider participation and support for their initiatives; and empower **artists, designers, statisticians, analysts,** and **journalists** to communicate through data in inspiring, informative ways.

CURRENT DATA SHARING PRACTICES ON TWITTER

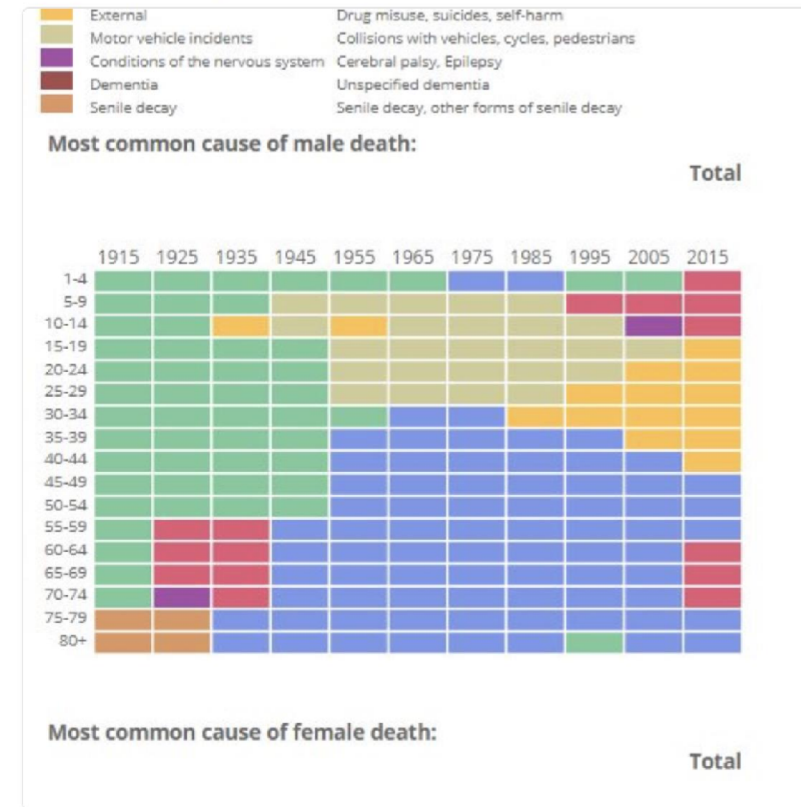
- What evidence can we see of data sharing activities?
 - What form is data being shared in?
 - How are the various stages of the data science pipeline represented?
 - Does anyone share raw data?
 - Do narratives explicitly reference the data that they are built on?
- How common is data sharing
 - Who is it done by?
 - How do they do it?
- What kind of data is (not) being shared?
- Who makes use of the data for what purposes?

OFFICIAL DATA

- 6 week Twitter study of ons.gov.uk
- 1186 original tweets made by 898 people, with 4906 subsequent retweets
- 15 most active tweeters, half work for the ONS or are official accounts of the ONS
- Most retweeted tweet (503 times) is by a BBC journalist mentioning an ONS data visualisation
- One of the 64 separate tweets about this ONS data release

Fascinating @ONS graphic on what's killed most people.

[visual.ons.gov.uk/causes-of-death...](https://visual.ons.gov.uk/causes-of-death)



12:21 PM - 18 Sep 2017 from City of London, London

504 Retweets 476 Likes



22

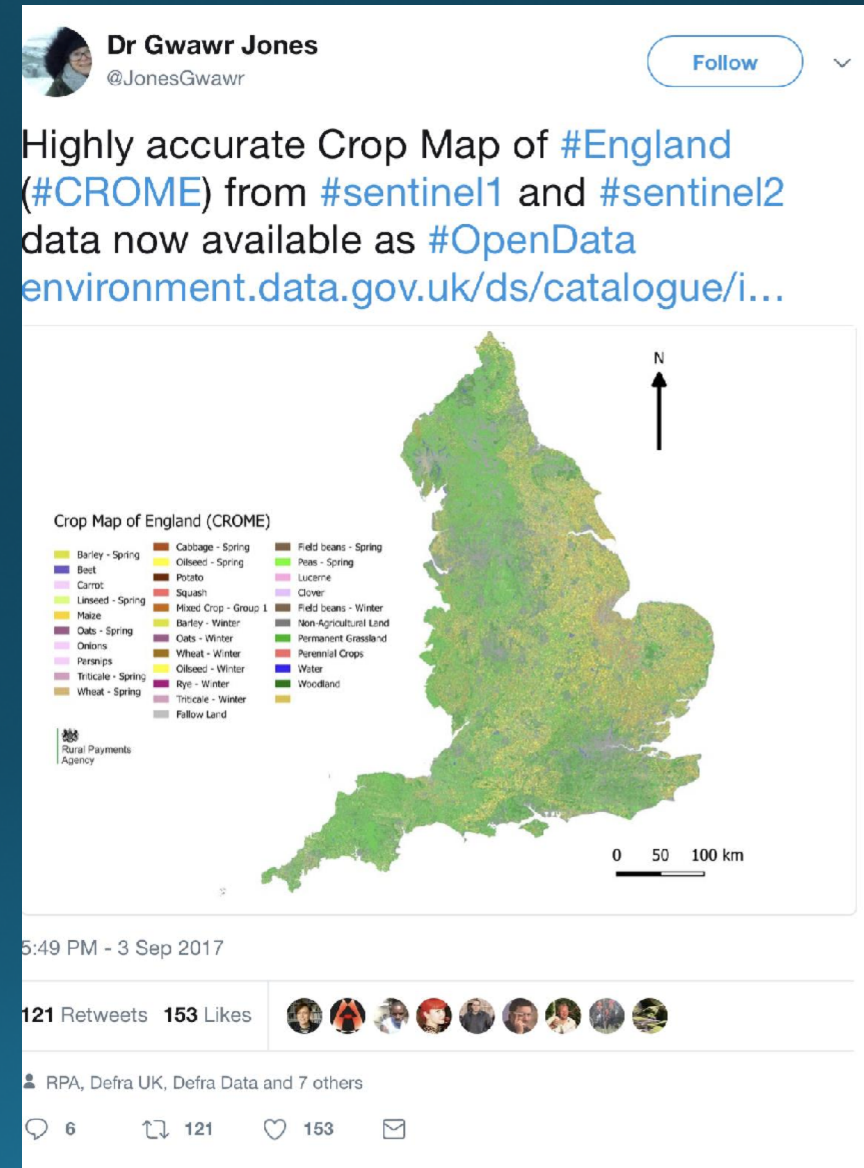
504

476



OPEN DATA

- Six week Twitter study of data.gov.uk
- 113 original tweets made by 87 different accounts, with 258 subsequent retweets
- No bias towards organisational affiliation is present in the set of active retweeters
- The single most retweeted tweet (121 times) is by a Joint Nature Conservation Committee earth observation specialist. Mentions a crop map visualisation from environment.data.gov.uk



Most Twitter-popular UK data journalists

mccandelish	David McCandless	London-based writer, designer, author and journalist. Follow @infobeautiful for my infographic / data visualisation work
Erik_Kaplan	Erik Kaplan	Investigative #Journalist on #Business #Data #Environment #Intelligence #Politics #Science #Technology & #Trade for #Newspapers & #AlternativeMedia.
martinstabe	Martin Stabe	Data journalist based in London. Head of interactive news @FinancialTimes. @ftdata #ddj PGP: 82AC A19A 2708 FE96 2C79 F636 5927 CBBF E81E F667
carlapassino	carlapassino	Freelance journalist focusing on the intersection of politics, housing, food, farming and the environment across Europe. Keen on data, tea and coffee. Mum
SophieWarnes	Sophie Warnes	Data Journalist at the Stats Palace. Sends https://t.co/bJONL7KLcm every Sunday. Loves cats, maps; hates pandas.
DataMinerUK	DataMinerUK	Data Journalist at @TheTimes. Interested in all things #ddj and #opendata. Opinions are my own. Get in contact if you have #data that needs digging.
jgro_the	Jack Grove	Times Higher Education reporter on employment, careers, PhDs, ECRs & Europe. CIPR Data Journalist of the Year 2016.
miketaylorsport	Mike Taylor	Broadcaster, live radio sport commentator (football/cricket/athletics/motorsport etc). Data journalist.
rosenbaum6	Martin Rosenbaum	BBC journalist on politics & current affairs, radio documentary producer, specialist in freedom of information & data. ESRC Council member.
ian_a_jones	Ian Jones	Data and graphics journalist at the Press Association .
abookbinder	Alex Bookbinder	Journalist, researcher, itinerant Canadian. Ex- @OpenSociety, @FrontierMM, @xchange_org, others. Migration, media, conflict, data, SE Asia. Views mine.
carldavidturner	carl	Film-maker, journalist, data nerd and cyclist. I'm not always serious. All views are my own, RTs not endorsements.
NadiaElghamry	Nadia Elghamry	Head of Digital @estatesgazette and Data Editor. Journalist. Data junkie. Social media lover. Foodie. Baker. Feeder. Netballer.
michaelcross	Michael Cross	Journalist in Muswell Hill, London (and Broadstairs). News editor Law Society Gazette ; free data fan. Views personal, retweet not an endorsement, obviously.
analopezct	Ana Lopez	Journalist, globetrotter, mad about data and science. I need mountains and meditation in my life. EN/SP/FR https://t.co/AzlyhKqng7
nicholasdunbar	Nicholas Dunbar	Author of 'The Devil's Derivatives' and 'Inventing Money'. Financial analyst, data visualiser, former journalist, founder of Risky Finance
danwainwright	Daniel Wainwright	Data Journalist #ddj at BBC News Online . These are obviously my views and no-one else's (who else would want them?).
clairemilleruk	Claire Miller	Senior Data Journalist at Trinity Mirror , maths geek, feminist. Book - Getting Started with Data Journalism: https://t.co/rXw8qtoTR7
alekswis	Aleksandra Wisniewska	Interactive data journalist @FinancialTimes aleks.wisniewska@ft.com
Ashley_J_Kirk	Ashley Kirk	Data journalist @Telegraph. Visiting #ddj lecturer @CityJournalism
PorcelinaD	Pamela Duncan	Data journalist with @GuardianData. Previously Irish Times
marce Ellison	Marc Ellison	Photo/data/graphic journalist, worked for: 60 Minutes, BBC, Al Jazeera, Huff Post, Guardian, CBC, Globe and Mail, Toronto Star, + Vice.

Spreadsheets

- XLSX
 - 0 spreadsheets
- Google sheets
 - 1475 original tweets from 1067 unique accounts with 6923 retweets
- Most retweeted spreadsheet (1188 times)
 - schedule for the timings of INKIGAYO broadcasts (a famous Korean livestreamed pop music program with live voting) in various cities around the world.
 - The tweet is sent by an account promoting BTS, a recent high profile K-pop band (the first to win a Billboard newcomers award in the US) and gives detailed song broadcast timings.

BTS Schedule: Converted Times for A.R.M.Y.S

File Edit View Insert Format Data Tools Add-ons Help

100% View only

Credit @bstwings_views | Share URL: bit.ly/btsxarmytimes

EVENT : INKIGAYO

TIME (KR): OCT 8, 12:10 PM KST

TIME (UTC): OCT 8, 3:10 AM KST

**Daylight Savings*

If there are any changes or corrections to be made, send an email to btsviews@gmail.com or tweet [@bstwings_views](https://twitter.com/bstwings_views)

City	Event Date/Time	Timezone
Seoul	Oct 08, 12:10 PM	KST
Abu Dhabi	Oct 08, 07:10 AM	GST
Algiers	Oct 08, 04:10 AM	CET
Amsterdam	Oct 08, 05:10 AM	CEST*
Ankara	Oct 08, 06:10 AM	TRT
Athens	Oct 08, 06:10 AM	EEST*
Auckland	Oct 08, 04:10 PM	NZDT*
Baghdad	Oct 08, 06:10 AM	AST
Baku	Oct 08, 07:10 AM	AZT
Bangkok	Oct 08, 10:10 AM	ICT
Beijing	Oct 08, 11:10 AM	CST
Beirut	Oct 08, 06:10 AM	EEST*
Belize City	Oct 07, 09:10 PM	CST
Berlin	Oct 08, 05:10 AM	CEST*
Bern	Oct 08, 05:10 AM	CEST*
Bogotá	Oct 07, 10:10 PM	COT
Bridgetown	Oct 07, 11:10 PM	AST
Brussels	Oct 08, 05:10 AM	CEST*
Changsha	Oct 08, 06:10 AM	EEST*
Chengdu	Oct 08, 12:10 AM	ART
Chongqing	Oct 08, 05:10 AM	EET
Colombo	Oct 08, 05:10 AM	SAST
Dhaka	Oct 07, 10:10 PM	CDT*
Dubai	Oct 08, 05:10 AM	CEST*

Spreadsheet Categories & Use

- Visual inspection of 100 highly retweeted sheets.
 - sports statistics (including gambling analysis)
 - computer games statistics
 - catalogues of resources/assets (including artist's videos or a series of TV episodes)
 - selling goods / artwork / services for a trader or fan group
 - co-ordinating donations / volunteers, political information
 - co-ordinating political activity
 - music voting
 - buying on behalf of an artist
 - monitoring cryptocurrency offerings.

Report of simple list	10%
Report of rich data (columns, visualisation)	40%
Complex report with analysis (sports /financial)	10%
Promoting action (voting / purchasing)	15%
Co-ordinating crowd action	20%
Other (timetable / schedule)	5%

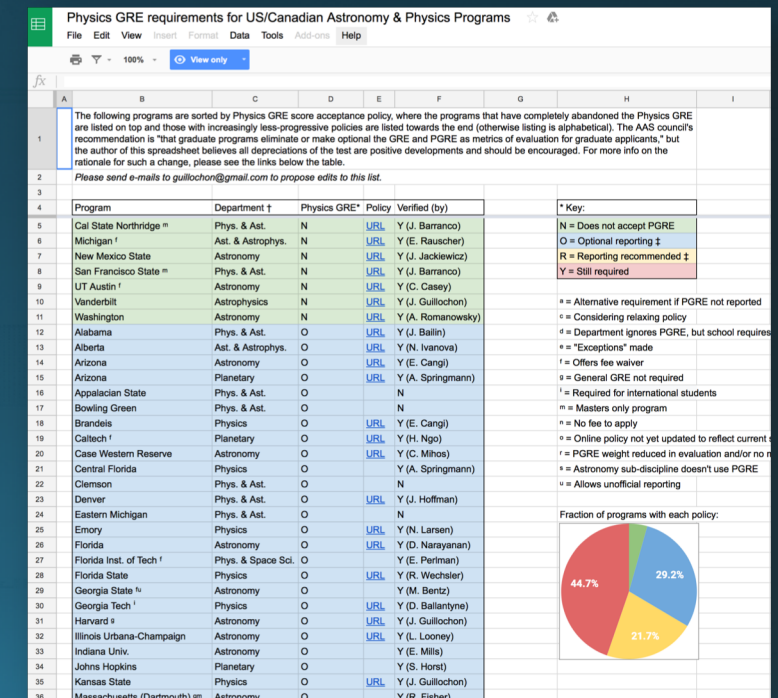
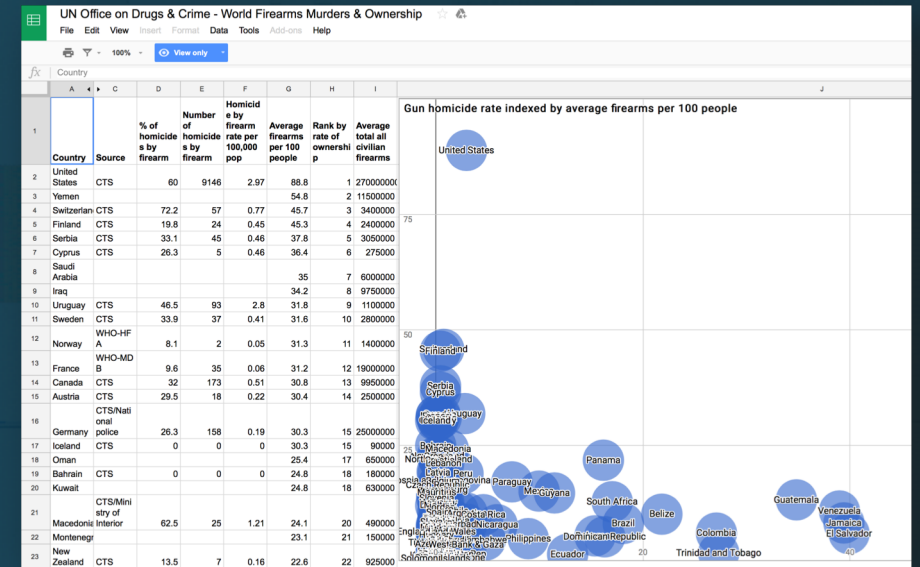
Use of Spreadsheet Charts

- 5% (29) of sheets contained charts
- 4 charts intended to promote subsequent use and discussion
 - A survey of fanfic writing community from attendees at a NYC festival
 - A maths teacher who takes part in Maths Teaching discussion groups, tweeted a Google form for to record preferences for banana ripeness. Data analysis and maths lessons ensue. The experience is blogged by the spreadsheet author.
 - A study on the citation of Registered Reports in Cognitive Neuroscience. This may be an example of pre- or alt-publication of bibliometric research.
 - Historic weather data collected by a local citizen offered to a “sports weather” journalist.

Games (trading, playing, curation)	7
Politics (monitoring, organising, arguing)	6
Surveys (attitudes or phenomena)	4
Financial investment analysis	3
Personal list of assets / achievements	2
TV / radio (voting/ratings)	2
Trading (orders)	1
Miscellaneous data collection <ul style="list-style-type: none">- Historic weather data- Boeing 787 production data (hobbyist)- Google Analytics audit of Udemy- Academic citation analysis	4

Use of Spreadsheet Charts

- 2 charts intended to support an argument or discussion.
 - A spreadsheet is used to plot UN data on firearms in a bubble chart in a discussion thread between pro- & anti- NRA positions. The author is a senior technologist in Microsoft.
 - A spreadsheet of the use of the Physics GRE in N American University Physics admission processes. The tweet is sent by a delegate at the Conference for Undergraduate Underrepresented Minorities in Physics, not the spreadsheet author.



Data Arguments

then prepare to have your mind blown with data

Daniel B @csuwildcat · Oct 3
Vegas shooting was awful, but shouldn't be used to take citizen gun rights that save 300-500k ppl/yr from injury or death. (Fed crime stats)

Jean-Paul Cane @cane_jean · Oct 3
I'm ok if we decide as a country that we're willing to accept costs of ownership, but let's not pretend that more guns mean safer society.

Daniel B @csuwildcat · Oct 3
Then prepare to have your mind blown with data:
To the contrary, data shows "absolutely zero" cause/correlation to more guns = more crime:



Daniel B @csuwildcat · Oct 3
If your inferred thesis was correct, you'd "at minimum" expect the data on those charts to show the inverse.

Jean-Paul Cane @cane_jean · Oct 3
Americans are 25x more likely to die from gun violence than ppl living in similar countries. That correlates directly w/ gun ownership.

Daniel B @csuwildcat · Oct 3
No, it doesn't - data clearly shows the inverse. FBI/DOJ state 80% of all gun deaths are directly linked to the drug trade. End the Drug War

Daniel B @csuwildcat · Oct 3
I'm honestly surprised you made a claim that more guns is correlated w/ more deaths immediately after I posted the data showing the opposite

Jean-Paul Cane @cane_jean · Oct 3
What does this graph tell you?



Image: Gun Deaths vs. Gun Ownership ← Inductio...
Found on Google from mark.reid.name
images.google.com

Daniel B @csuwildcat · Oct 3
Replies to @cane_jean
Shows me the author cherry picks datasets (OECD). Here's the full set, which I calc'd myself; very different result:



2:04 PM - 3 Oct 2017

Daniel B @csuwildcat · Oct 3
Replies to @csuwildcat @cane_jean
The distribution of a more diverse dataset reveals the most important point: gun effects are culturally/socially driven. (for US: Drug War)

Adam Rackis @AdamRackis · Oct 3
economist.com/news/leaders/2 ...

its madmen, but Min was armed only with a knife, so none of his victims died.

If America is ever to confront its ob...
with guns, th...
rate is four times higher than Brit...
times higher than Germ...
an anti-American bigot prepared...
Americans are four times mo...
murderous than Britons, could p...
pretend that no connection exists between...
those figures and the fact that 300m guns are...
"out there" in the United States, more than...
one for every adult.

Daniel B @csuwildcat · Oct 3
Adam, the data is crystal clear - your assertion/claim that more gun prevalence generates more gun crime and gun deaths is abjectly false:



bigot

idiot

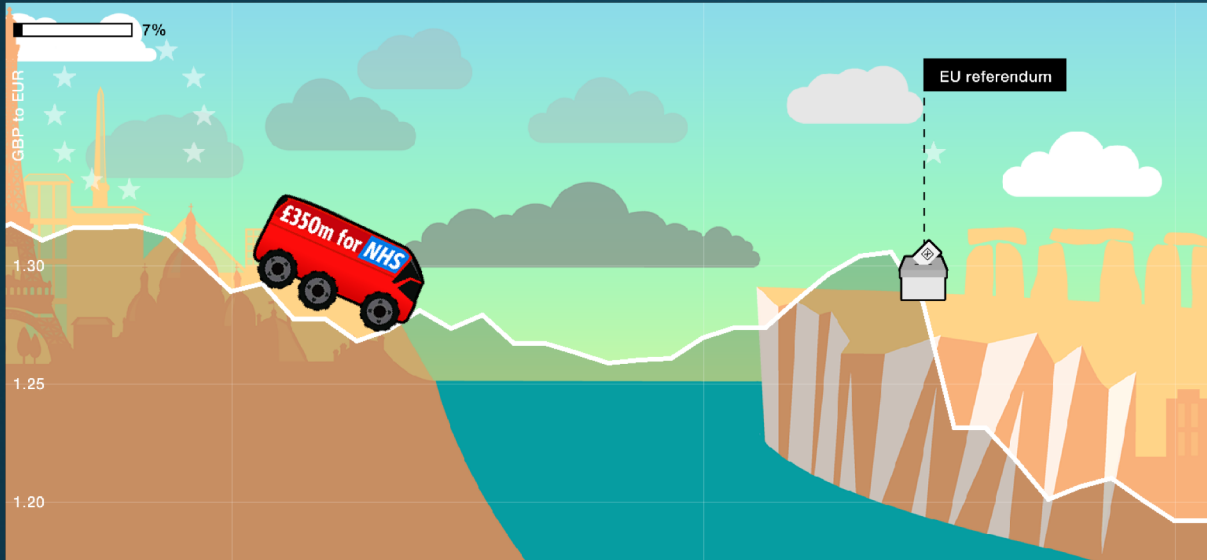
Daniel B @csuwildcat
Head of Decentralized Identity @ MSFT. Fought for the Web @ Mozilla. Self-taught coder. Voluntaryist libertarian. Reminder for idiots: this is a personal acct
Los Gatos, CA
backalleycoder.com
Joined September 2008
5 Followers you know
1,915 Photos and videos

BTC_JPC @cane_jean
Bitcoin/crypto enthusiast, geek, surfer, rock climber, interested human
South Carolina, USA
Joined October 2014
30 Photos and videos

BTC_JPC @cane_jean · 9h
Not since ETH was \$7 last year have I observed a greater divergence and fundamentals in a crypto asset. We may go lower first, but I'm n... against the honey badger. #bitcoin

RAMIN NASIBOV @RaminNasibov · 17h
Cuban floors! (by Sebastian Erras)

Data Games



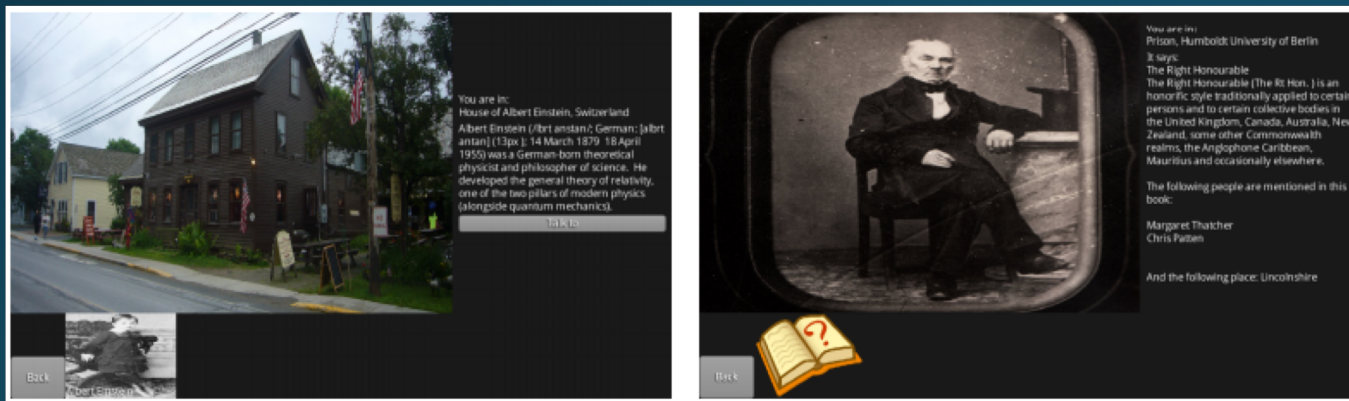
Brexit Bus

- Player controls a (familiar) red bus
- “Road” is actually a graph of the value of the Pound (£)
- Player can explore how certain events affected the value

<https://advisa.se/en/research/brexit-bus/>

Generative Adventures

- Using *OpenStreetMap* and *Wikipedia*, generate the locations, NPCs, and (crucially) the plot of a mystery/adventure game
- Player follows a trail of clues and interacts with real-world historical figures



<http://julian.togelius.com/Barros2016Playing.pdf>

Brexit Bus

- Player controls a (familiar) red bus
- “Road” is actually a graph of the value of the Pound (£)
- Player can explore how certain events affected the value

Concluding Thoughts

- I've gone on too long, haven't I?
- Educating humans or computer systems is seriously difficult
- Helping humans use their skills on the VAST amount of readily available knowledge is really important.
- We have only just started.