You Are Being Tracked, Evaluated and Sold: an analysis of digital inequalities

Bev Skeggs, Professor
Director, Atlantic Fellows Programme
International Inequalities Institute, LSE
September 2017
ESRC funded (2013-2016, ES/KO10786), with support from Sociology and Computer Science, Goldsmiths, University of London

https://values.doc.gold.ac.uk/
https://values.doc.gold.ac.uk/firstfindings/

Bev Skeggs and Simon Yuill
Trading faster than the speed of light

• 120 milliseconds is about one-third of the time it takes a human to blink an eye
• Bids are made by advertisers for access to your data 50 billion times a day, or about half a million times per second.
• 100,000 requests from advertisers per second (Any longer than that, and it threatened to delay the Facebook page load).
• Light travels fast (299,792,458 meters per second), but not fast enough.
• In 2014 Facebook had 52,000 unique attribute data signals on one profile. It buys other data to develop this profile.

XGQnGPa1MCmThgb9wN4vL0UpgBUUtWg.rg.FTN.0.AWUxZtUF

From Antiono Garcia Martinez (Facebook Product Manager 2016) “Chaos Monkeys”
Advertising use of programmatic marketing is the digital industry equivalent of financial flash trading (see Michael Lewis’s “Flash Boys”)
Trading and tracking

- As you finish this sentence if you have anything open on your laptop the chances are you will have been tracked and traded, especially if you are a high net worth individual with an influential group of friends, often in less than a millisecond (a thousandth of a second 1/1,000), likely over 100 times. In 2016 estimated 60k per day.
- Uni of Cal study of Alexa (web traffic data analytics) found 5,000 trackers on the top 100 websites. Google is the top tracker.
- Nb: potential future net worth also makes a difference so students are often highly rated (family/friends/education dependent)
- It depends on opening webpages that may or may not be of interest to advertisers. Every time you open a new webpage you are being assessed for trading potential (phones more protected than desk/laptops)
- Even if you are not high net worth you will have been traded...... in a different arena, for different purposes, different products, and...
- NB: not just about trading ....tracking is also used for state surveillance

Facebook continually uses machine learning algorithms (with over 100,000 signals) to experiment with your data to enable it to match to advertisers. You provide the signals every time you use the internet eg device used, speed of connection, browsers, emails, webpages, banking, language in messaging, videos watched, reactions to, click bait, location data, your networks, plus historic data, plus data bought from brokers –

Or from governments in the case of the recent Indian government Aadhar scandal –), data from government agencies (eg NHS to Google)

- You can turn your phone off. Turning off the Facebook platform makes no difference.
Rhythms of interaction: “lifeness”

Research question: what happens when intimate relations eg friendship are monetized?

In 2013 asked here at the LSE (for the BJS annual lecture) at the very beginning of the project: “Values beyond value: Is anything beyond capital?”

Began as a study of social interactions, ended as an investigation into new forms of capitalism

https://values.doc.gold.ac.uk/interactions/
What we did:

1. A random digital survey of people’s use and attitudes to use with 154 participants
2. A Facebook App that provided data about participants’ activity from the Facebook API, including all user-generated posts in their newsfeed with which they had interacted (with their permission)
3. A plug-in for Firefox browsers, called Admonitor, that gathered data we were not able to collect directly from Facebook itself (this included advertising shown on Facebook, the order of posts as they appear at the top of users' Facebook timelines, and advertising included on other sites via systems such as GoogleAds). (with participant’s permission)

It took 6 months to write the software as it was frequently interrupted by Facebook’s changing code.

1. Developed heuristic device to analyse 200 million ad words
2. A live website provided daily visualizations of the data as it was collected – participants could see tracking and interactions
3. Generation of databooks to enable the reading across different data
4. Post-data collection interviews with 15 participants to review their data with them

We were registered as Facebook developers and of our participants 33 managed to register as Facebook testers. See https://values.doc.gold.ac.uk/interactions/

All software is open source and designed by Dr Simon Yuill available at: https://github.com/valuesandvalue

We tell the story of the difficulties in “The Methodology of a Multi-model Project Examining how Facebook Infrastructures Social Relations” (in Information, Communication and Society, 2015). Open access.
Key to interaction flows:

NB: the grey matter, behind the Facebook platform
Grey matter = tracking on browser
Manipulating the news feed for attention

https://values.doc.gold.ac.uk/firstfindings/
Loose networks made it more difficult for FB to profile.

Tight and influential networks enable FB to build profile from friends and target more effectively.
FB tracking of general browser use – collating all sites data, keyword matching-enable profile building

Sites tracked by Facebook

Information embedded in URL’s of online adverts and trackers

=data extracted when ad words correlate with tracker source

https://values.doc.gold.ac.uk/
The trail of Facebook as it tracks your activity on other websites. The most visited at the top. The circles are scaled to a size that represents their relative size to the number of ads that have shown a match (if live you can see how many times you have been targeted by a particular ad company on your Facebook platform.)
Facebook processing power (Manovich (2016) 100 Billion Data Rows per Second)

• i300 million photo uploads per day on Facebook as 968 million people log on daily.
• In 2014 Facebook processed 600 TB of new data per day.
• Scuba, Facebook's fast slice-and-dice data store, stores thousands of tables in about 100 terabytes in memory, ingesting millions of new rows per second whilst deleting just as many. Throughput peaks around 100 queries per second, scanning 100 billion rows per second, with most response times under 1 second.
• Over 2 billion active users monthly (disputed - but still significant). If we subtract China and Russia (own systems)- that’s only about 35b left!
You do **NOT** have to be on the Facebook platform in order to be tracked.

You just have to be on an internet page with a link to Facebook = nearly all.

Check the bottom of the page you are on for the f symbol  eg:
Failure to regulation a global oligopoly

• In June 2015 a national government (Belgium) tries to block Facebook tracking non-Facebook users without their knowledge or permission.

• It employs 5 US computer science departments to collect evidence of tracking. Until this point (June 2015) Facebook has denied that it is tracking non-users of its platform (nb our research 2013-4).

• They found that Facebook tracks the web browsing of everyone who visits a page on which its logo is inserted. It tracks the computers of users without their consent and tracks those who do not have a FB account or have explicitly opted out of FB.

• When a user visits a third-party site that carries one of Facebook’s social plug-ins, it detects and sends the tracking cookies back to Facebook.

• **They also found that the opt-out mechanism on FB actually enables tracking**

**However, on June 29th 2016 Facebook Wins Appeal against Belgian Court Case Over Storing Non-User Data**

The ruling was based on the fact that Belgian courts do not have international jurisdiction over Facebook Ireland, where the data concerning Europe is processed.
Why so much tracking and trading?

- To make money, pay shareholders, invest, growth.
- The majority of Facebook’s revenue 97% comes from advertising.
- Facebook market capitalisation (intraday): $270.37 billion (16.10.15), $312.3 (04.03.16), $321.97 (20.04.16), $369.04 (03.10.16), $451.36 (20.04.16), $495.28 (20.09.17). Source: NASDAQ.
- To keep ahead of the market against competition, hence have to control competition eg Instagram, what’s app.
- To be able to sell itself as having the “best matching” of consumer to advertiser.
- Eg USP =

Two billion people use Facebook every month. 1 of every 5 minutes that people in the US spend on mobile is on Facebook or Instagram. 500 million Instagrammers use the app each month.
Digital capital: Profit without production: Costas Lapavitsas

Digital companies run platform services (eg Facebook and Google). They do not make products to be sold. Microsoft and Apple have products.

They rent you space on their platforms paid for by the right to your personal data.

They employ a tiny number of people = 17k (labour to market cap and capital costs –small fixed capital).

They make money from extracting revenue from non-productive areas of the economy, from the circulation of capital eg from traditional advertising, from consumption, competing for the same market.

They extend the circuits for the circulation eg Facebook ready eg via messenger to enter financial markets via fintech and de-regulation of bank data ......
Attempt to monopolize data through generating an indispensable global utility

- Although they follow the “traditional logics of capital”: consolidation of interests via monopolization of different fields, diversification globally
- Consolidation through buying other companies eg (next slide)
- Global diversification: eg games in developing countries, “free basics” in 49 countries, mobile banking in China, P2P, drones, 3D, dumb phones, telecommunications, Facebook for SIM in Argentina, Chile, Columbia.
- The value of a company’s assets is not so much a function of the products and services it produces as of the ability to control risk via image, branding, development, IPs, trademark, copyright, trade secret laws, patents and licensing: the overall control of knowledge.
- FB creates value through proprietary rights (Bohm and Land 2012; Harvey 2005; Ekman 2012; Perelman 1998) which along with the above comes from access to and rights over to your data – your agree to the right for FB to use your data
- Facebook creates monopoly making a ubiquitous and unavoidable utility – operates as if electricity – necessary for network connection (eg spotify).
Buy the Competition: consolidation, monopolization, diversification
Facebook business says about its advertiser auctions

• “Advertising on Facebook makes it easy to find the right people to capture their attention”

• The advert auction determines which adverts should be shown to which people. Using information that you provide in the advert creation process, the auction shows your advert to the people (algorithmically matched) who are most likely to be interested in it.

• When showing ads, we try to balance two things:
  • Creating value for advertisers by helping them reach and get results from people in their target audiences
  • Providing positive, relevant experiences for people using the Facebook family of apps and services

• “Our goal is to match the right ad to the right person at the right time”.

• It is selling access to your “eyeballs” and charges according to how valuable your eyeballs are! (eg do you have the purchasing power, are you likely to buy, what else have you bought recently, what are your spending patterns, how much debt, etc....)
Facebook’s monopoly utility advertising capacity

• Facebook’s scale for real-time advertising cannot be matched remember above – faster than the speed of light and getting faster).
• In the tradition of monopolies it can reduce revenue through high demand because it can access the largest audience, unmatched by any other company.
• Elkin (2016) argues this is unlikely to change as no one else has the capacity to compete.
• Gehl (2014) demonstrates how the software architecture of Facebook shaped global advertising standards Arola’s (2010) shows how the IAB standards were in part determined and “template-driven” by Facebook.
• It can experiment with your data: “the Ads team would slice off tranches of the FB user base in rich ads markets and dose them with different versions of the ads system to measure the effect of a new feature, as you would test subjects in a clinical drug trial” (Martinez, product manager Facebook 2016).

Now controls
More than 50,000 algorithms

75 billion transactions per day
6x the trade daily of NASDAQ

14 global offices

Tracks with API’s, cookies, web-beacons, Chango a keyword tracker

Known as “programmatic advertising/marketing”

Restructuring the advertising industry as demands for “creatives” diminished by the speed and power of algorithmic auctions

We found they also sell access to other data auctioneers to trade your data eg:

All checking your browser use via machine learning algorithms
We also found data brokers extracting data from the FB platform: FB also buys data from data brokers to enhance the profiles it develops

- We had 1,968 examples of Acxiom, 23 from Experian. In 2015 Acxiom brags that it has, on average, 1,500 pieces of information on more than 200 million Americans.
- Oracle collects 3 billion profiles from 15 million websites and 700 million social messages daily. One of its specialties is the tracking of TV viewing habits.
- **Competition between is intense**
  - All based on real-time scoring when all signals you send (eg web browsing, detected through tracking) is sold as a “hit” (match) or packaged to develop a profile of your “worth”
  - There is a great deal of competition by stealth to collect information on how best to profile you. Remember this comes from a range of sources: government data (eg social security, national insurance, NHS Google sifted in UK, criminal stats, all purchases, credit cards, ATMs, all browser use, owning a cat! plus geo-demographic, everywhere you go, when you look at digital adverts on the street.
  - Every day life has been colonised by data extractors and traders, all in the interests of advertising/directing consumption.
  - US Federal Trades Commission reported (2014) it is impossible to find out how you are being packaged and sold. You can’t even find out how these data brokers collected all the data they have on you. You may never know.
  - Or you may find out when you realise something has gone badly wrong (50% inaccuracy: World Privacy Forum Report 2014). Experian’s credit arm had 15 million customers’ information, including social security numbers, breached in 2015.
- **Beyond regulation.**
Data brokers plus Facebook – trading between

• FB has “social data” eg speech patterns, head movement, face recognition, friendship networks, rating of influence, affective responses, which gives them an advantage in profiling. But only IF combined with all the other demographic data eg from Experian (traditional DBs) GPS, health, welfare, consumption data, some which it buys off other data brokers to extend and build a person’s profile.

• All data controlled through intellectual property rights.

• Facebook will soon have access to all financial data in Europe: as a result of a European Directive PSD2 (Jan 2018), it will be able to access the API’s of traditional banks and experiment with that data without regulation.

• FB operates as a metadata economic ecosystem (all data flows through eg other platforms like Spotify) Haynes and Nguyen 2014)
Why significant? differences in tracking and trading: “Worth and waste” (marketing classifications)

But we know from the 2007-8 financial crisis that nothing is wasted. What we also know is that debt is one of the largest financial markets. Those who are not sold commodity objects are often sold debt........
Generating digital inequality: not just debt but also a different surveillance

- Those already subject to conditions of impoverishment will experience more surveillance with higher stakes, harsher consequences and lack of resources to seek redress (Gandy 2009; Gangadharan 2012; 2015).
- Madden et al. (2017) identify networked privacy harms in which users are held liable for their own behaviour and the actions of those in their networks.
- They will be subject to excessive state surveillance: eg Job applications more likely to be a source of surveillance, and more likely to be subject to “threat scoring” – a person’s propensity to be involved in a crime
- Performative commercial categorization by data brokers as eg “rural and barely making it”, “fragile families” as in, the categorizations will have future effects.
- Credit scoring algorithms considered to be a “trade secret” (but 1 in 5 have errors); state profiles also often inaccurate, but very difficulty to change profiles eg re incarceration (Pasquale 2014)
- Almost impossible to bring a legal case (difficulty of having the technological power to collect the evidence eg to challenge Steve Bannon’s $1.3 billion super computers)
- No algorithmic accountability
- The already impoverished are targeted for predatory financial products, payday loans, online classes or debt relief services (Federal Trade Commission Report on Big Data 2016). Techniques such as “lead generation” rely on targeting the vulnerable, what Jackie Wang (2017) calls the new racial capitalism which begins with parasitic governance and predatory lending that extends credit only to dispossess later. Predatory lending has a decidedly spatial character and exists in many forms, including subprime mortgage loans, student loans for sham for-profit colleges, car loans, rent-to-own scams, payday loans, and bail bond loans.
- Privacy is an economic positional good.

- Paradoxically high net worth individuals also subject to hyper surveillance but in order to target for consumption not to target for debt, policing and future impoverishment.

- Very different digital tracking with very different consequences.
- It is what Jodi Dean calls “communicative capitalism” or Shoshana Zuboff calls “surveillance capitalism”, or Jackie Wang (2017)
Tracked the arc of a US person’s life – primary school tests, education funding, teacher scores, college recruitment eg sub-prime university targeting eg Trump uni

Job applications, CV scanning, evaluate worker performance, tracking social media

Banking, credit, spending habits

All search histories

All health data

Predictive policing

Feed off each other over a lifetime “a death spiral of modeling” hence WMD

But also inherited.....

Opaque, unregulated, incontestable: “these models are constructed not just from data but from the choices we make about which data to pay attention to and which to leave out. These choices are not just about logistic, profits and efficiency. They are fundamentally moral” p218
AADHAAR: The Indian Identity Card System (or the Silicon Valley biometric gold rush)

- The largest biometric database in the world (but also in Kenya, Bangladesh...)
- “sold” to population as a means to provide access to welfare, public services, subsidies and rehabilitation provision; “empowerment” of the poor.
- But linked to cash machines (Aadhaar Pay), 2016 bill allowed private companies to access all data (but not individuals)
- Basic issues – labourers without fingerprints! (36% failure)
- Enabled the withdrawal of welfare provision
- In 2016 130m users hacked, scores of incorrect data recorded

- Wikileaks (26 August 2017) revealed that CIA (plus French and English defence companies) have access to the whole system via the private biometric data company Crossmatch that helped devise AADHAAR (disputed!): see Novara media for full debate as Indian court rules on data privacy.

- Globally a whole industry “cyber insurance” industry has been established as a private solution to state use of digital companies eg recent UK NHS ransomware scare (Morozov 2017) ; isn’t it supposed to be the other way round? (an ever expanding realm of industrial global scale privatization)
State tracking and trading:

- They have also become a service for the state, offering up data on a scale that most national states cannot generate, as we saw with PRISM leaked docs; FB, Google handing over vast amounts of data to NSA in 2009.

- Facebook founder Jim Breyer has a relationship to the venture capital wing of the CIA In-Q-Tel with....

Trump’s tech advisor and another Facebook founder Peter Thiel, founder of PayPal (deregulation of financial transfer) and Palantir, an intelligence gathering company working with the CIA and US military. A trans humanist Ayn Rand far right libertarian, Thiel secretly funded legal case against Gawker. Also targeted activists, reporters, labor unions and political organizations using hacking tools to break into computers

- The Connectivity Lab at Facebook is working on new aerospace and communication technologies with experts from Ascenta, NASA’s Jet Propulsion Laboratory, NASA’s Ames Research Center, and the National Optical Astronomy Observatory (see internet.org).

- Ippolita (2015) details the involvement of key Facebook members in libertarian politics in the USA, including TheVanguard.Org, a neoconservative pressure group set up to attack Moveon.org. These are global organisations.
A particular sort of state
Machine learning and algorithms are not innocent:

Facebook’s Secret Censorship
Rules Protect White Men From Hate Speech But Not Black Children

A trove of internal documents sheds light on the algorithms that Facebook’s censors use to differentiate between hate speech and legitimate political expression.

by Julia Angwin, ProPublica, and Hannes Grassegger, special to ProPublica, June 28, 5 a.m. EDT
Algorithmic advertising as political tool; a platform for manipulation

• Facebook admitted in July 2017 that it found that more than 5,000 ads, costing more than $150,000, had been placed on its network between June 2015 and May 2017 from "inauthentic accounts" and Pages, likely from Russia. (source: https://www.wired.com/story/facebook-ties-more-than-5000-political-ads-to-bogus-russian-accounts/)

• Advertising not just about products but political “fake news”
https://www.wired.com/2017/02/veles-macedonia-fake-news/
To recap; back to the research

• Our findings detail how Facebook’s desire for future capital accumulation means:
  • They are constant tracking you via Facebook and all other sites which have an embedded
  • They constantly collate all your browser data to match ad words and add to profile
  • Through their trading they shape your network over time, how you interact – with whom, when, and how
  • It is your attention they want
  • They are constantly expanding capacity to experiment with your data
Their USP is their ability to do all the above
Inequality?

• By stealth
• Classifications unknown
• No accountability
• Debt silo/sub prime market creation not known until legal challenge reveals it (eg Trump university)
• We do not know how our data traces are being traded and valued
• We are the fodder, the resource from which value is extracted in order to sell us stuff: a constant circle of expropriation.
• Little protection
• Litigious companies
But not just the making and legitimation of structural inequality; it is political, we become aligned to the desires and rhythms of capital

- We found Facebook makes us do things.
- It shapes our habits, such as the FOMO (fear of missing out), the compulsion to constantly check, to broadcast, to lurk and the many activities we identified through our research,
- To perform ourselves in social media form: extracting profit from the performance of subjectivity.
- It makes us pay attention to certain things over others (significance of sensation)
- We even walk in anti-social ways!
- This often means smart phones have become similar to prosthetics, rarely leaving our hands, converting our time use for the interests of consumption.

Our intimate relations are re-shaped (eg friendships) – but hard to identify

- Power works through us connecting devices to bodies, people to people and advertisers to us.
- This oozing of capital interests for potential future accumulation into our daily life and habits has become normalised – our research participants were “resigned”: they described their use of FB as “a contract with the devil”, “a necessary evil”,
- This is ideology without ideas, an ideology of expressive action, ideology of form not content, an ideology made from connective convenience
Ideological? But not directly – power working through us

• By stealth
• By demanding attention
• By filling our time
• By shaping type and form of communication
• Results in acquiescence to power by:
  • bodily habit
  • convenience
• By confusing forms (advertising as a political platform)
• By resignation and normalisation
• It works through form (as a utility), classification and content

• *Facebook only have values when it will increase their value*
What to do?

• Many proposals to regulate and nationalise

• New general data Protection Regulation EU regulation (GDPR) from May 2018 may temporarily deter data blockers because the regulation insists that companies have the consent of those they are using. Trackers should be made visible to users.

• Facebook and Google unlikely to be effected by the GDPR as they have “consent” when you sign up and agree terms.

Nationalisation a good idea but by what sort of government? – certainly not the one that passed into law in the UK in 2016 the investigatory Powers Act? An act that gave the police the most sweeping surveillance powers in the western world? (unmatched even by the US) that legalises state hacking of all mobile and internet communications, one “more suited to a dictatorship than a democracy” (Open Rights Group)

• Not the Indian government which sells its data globally

• The UK government that already uses the services of private companies for its own surveillance of citizens (as revealed by Snowdon and Wikileaks), but also to protect the systems such as the NHS which have been privatised.

• In the UK we are currently in a massive land grab for billionaires via our data – not just tracked and traded but politically manipulated (we are living with the Brexit consequences). See anything written by the brilliant Carole Cadwalladr

• The state needs regulation as much as the private companies.
But how?

- Regulate and nationalise
- Address privacy not just as a human right but as economic positional good too
- We need to be able to rethink what do we want technology to do, not do to us: turn technology to social (ist) uses.
- At a time when stealth is spreading into different spaces and becoming ever more stealthy as the internet of things intrudes into our homes eg Alexa and Echo, for instance, are listening out for us at all times, in our own homes, for “pre-emptive capture”. Don’t let them in!
- Make the machine learning work harder – used to think obfuscation, going random – getting harder to do
- Minimize information given
- Take back time, break habits, tear off the prosthetic
- Avoid all online shopping, ATMs, use public computers. Use cash!
- Learn to hack/code
- (then change the device identifier on your phone)
- Go dumb
- Turn them off!

- EDUCATE AND ACTIVATE
What Facebook Knows About You

by Julia Angwin, Terry Parris Jr. and Surya Mattu, ProPublica
September 28, 2016
Blockers; confusers

• Via private companies and ad blockers eg Abine DNT+ (http://www.businessinsider.com/this-is-how-facebook-is-tracking-your-internet-activity-2012-9?IR=T)

• Free; Track me not (Computer Science, NYU https://cs.nyu.edu/trackmenot/). Works by creating confusion.
Alternatives to Facebook

- https://crabgrass.riseup.net/
- https://diasporafoundation.org/
- https://ello.co/manifesto (launched 2014)

- https://www.facebook.com/OffiziellAnonymousPage
- Involved in establishing WorldSocial which hosts ZSocial, UTNESocial, DollarsandSenseSocial, SyrizaSocial, GreenSocial (the U.S. Green Party). “There will be no advertising or sale of user information. If a government gets any WorldSocial information, it will be because they stole it”.
• https://www.eff.org/deeplinks/2013/02/howto-opt-out-databrokers-showing-your-targeted-advertisements-facebook

• https://lifehacker.com/5994380/how-facebook-uses-your-data-to-target-ads-even-offline

• https://lifehacker.com/5843969/facebook-is-tracking-your-every-move-on-the-web-heres-how-to-stop-it

TTC workshops and tools:

ONONYMOUS.ORG [2015-]

2015

Ononymous.org brings together materials around online protection.

The website, launched in August 2015, features projects by ourselves as well as by others working in the field, including Front Line Defenders, EFF, Open Data City, The Tor Project, The Centre for Investigative Journalism and Access Now.
Tactical Technology Collective

ABOUT US

Founded in 2003, Tactical Tech is a non-profit that has been working worldwide to demystify and promote technology in the context of activism for over a decade.

Working at the intersection of tech, activism and politics, Tactical Tech reaches more than three million people worldwide annually through events, training, online resources and exhibitions. We are an international group of technologists, activists, designers and practitioners based in Berlin, who work with citizens, journalists and activists to raise awareness about personal data, privacy and digital security.

HISTORY

https://tacticaltech.org/projects/glass-room-london#