Title: Against The Black Box
Type: Article
URL: https://ualresearchonline.arts.ac.uk/id/eprint/12426/
Date: 2016
Citation: Smith, Oliver (2016) Against The Black Box. Ubiquity: The Journal of Pervasive Media, 4 (1-2). pp. 39-51. ISSN 2045 6271
Creators: Smith, Oliver

Usage Guidelines

Please refer to usage guidelines at http://ualresearchonline.arts.ac.uk/policies.html or alternatively contact ualresearchonline@arts.ac.uk.

License: None specified

Unless otherwise stated, copyright owned by the author
“If power was previously exerted in the disciplinary practices of design at a built and urban scale, power is shifting into the codes, programs and archives of telecommunications and network technologies.” - Design Act

In our everyday use of technology we step over, round, and through many different systems, services and devices, we connect with some for fractions of a millisecond and others remain with us, in our pockets and about our persons. We put them to work on the information we provide and receive, allowing them to mediate our communications, knowledge and actions, reliant on their inherent capability for data collection, transmission and manipulation. They may seem to be working with us, or for us, but in which direction does the power flow? How can we truly know when it is not possible for us to fully inspect and critique them? They are Black Boxes: we provide an input, they provide an output, but the machinations and manipulations in the interim are hidden from view.

It is necessary to situate this discussion firmly in the context of the revelations made by Edward Snowden: concerning technological data gathering and surveillance programs from the security agencies in the ‘Five Eyes’ program. Revealing, among other things, the widespread collection of internet activity, access to and the ability to control computer networks and the creation or maintenance of backdoors, or security flaws, in consumer electronics, Snowden’s leaks highlighted how little was known about the scale and capabilities of these activities, the potential of many of the consumer devices in daily use to collect and disseminate private information, or be coerced into doing so, and the lack of understanding that even some technology and service providers had of the activities occurring in their own devices and networks. The aim here, then, is to come to an understanding of the structures of power and control that these systems comprise, to lay the foundations for developing existing and creating new methods of resistance and dissent in the face of the Black Box.

**The Black Box**

During World War II, the mathematician Norbert Wiener worked to improve automatic anti-aircraft targeting systems. Taking the exploratory, electrical-engineering approach of black box analysis and applying it to human behaviour, Wiener was able to predict the actions of an aircraft pilot with sufficient accuracy that the guns could fire at the point the aircraft would be when the bullets reached it. Wiener’s breakthrough was that it wasn’t important to know, in advance, anything about the pilot or aircraft; rather, by observing their actions and movements, and analysing patterns within these, enough information could be gathered to make the necessary, precise prediction of what would happen next. In this sense, pilot and machine were taken together as an unknowable ‘Black Box’, as a simplified model that could be described to the rudimentary machines of the time. After the war, Wiener would take this principle forward, becoming a founder of the discipline of ‘Cybernetics’.

---

In his book ‘An Introduction to Cybernetics’, W. Ross Ashby discusses further the ideas of the Black Box, providing us with a bridge between its electrical-engineering origins and Wiener’s appropriation for use on a wider range of more complex systems. Rather than using the Black Box as a tool to gain understanding through simplification, Ashby discusses it as a problem to be tackled. He presents the lack of understanding as problematic, something to be overcome by the experimenter’s use of “certain given resources for acting on it (e.g. prodding it, shining a light on it) and certain given resources for observing its behaviour (e.g. photographing it, recording its temperature).” He gives the example, again military in its origin, of an engineer required to investigate a broken, but top secret, bomb sight:

“Sometimes the problem arose literally, when a secret and sealed bomb-sight became defective and a decision had to be made, without opening the box, whether it was worth returning for repair or whether it should be scrapped.”

With a fundamental distrust of its users, the Black Box, here, becomes a strong manifestation of hierarchical control. The engineer is required to maintain a device, but not trusted to understand it fully, allowed a level of control over its destiny, perhaps, but not able to do this from a position of comprehension: the balance of power is maintained in favour of the Black Box.

For Ashby, the Black Box problem takes one of two forms: ‘The Very Large Box’ and the ‘Incompletely Observable Box’. When Ashby was writing the miniaturisation of general computing devices that has occurred today had not begun, and increases in computational or technological complexity necessarily increased the size of the object housing them. Today, elaborate systems can still be found. Despite their small size, modern computing devices such as smartphones and laptops are highly intricate computers; the phrase Very Large Box, then, can be seen to describe an incredibly complex system and will be used here in that sense. As for the Incompletely Observable Box, early computers were of considerable size, often taking up rooms, or sets of rooms, but it was possible to see their whole, even if it was necessary to climb inside them, and initially at least it was possible to understand their construction by looking at them. The phrase in this context refers to devices, such as the bomb-sight, which were made inaccessible by structures of control, by legislation and secrecy, and in this sense remains relevant to the discussion here, but it is also worth considering for another reason: the internet, the global network joining many computers and devices around the world has produced technology that cannot be easily seen, both because of geographical inaccessibility, with cables running along seabeds for example, and its scale, stretching around the earth and, through satellites, into space.

---

3 ibid.
4 As well as regular phones and tablet computers, thermostats, watches, vacuum cleaners, light bulbs, speakers, thermometers, microwaves, pacemakers, bike locks, signs, bins, washing machines and so on.
The Black Box does not spring into existence, fully formed, as a large, complex entity; it grows through an iterative process of exploration and formalisation of knowledge. Bruno Latour discusses this in terms of the progression of scientific ideas and theories and, in terms closer to the Black Boxes discussed here, Garnet Hertz, in ‘Art After New Media’, shows a technology’s progression towards a “single punctualized object” which is used, not understood and calls it “a requirement of infrastructure and technological development”. These “punctualized” objects are a large part of the tools and appliances we use day to day. Our usage and understanding of them is presented in terms of input and output and, as far as our comprehension goes, they are Black Boxes. With wide network capabilities, trade secrets and and laws preventing their full inspection’, they are Very Large and Incompletely Observable in the fullest sense of the terms.

Before moving on to the core examples of the Black Box, it is important to note that, while this essay will discuss Black Boxes mainly in terms of computational technology, networks and devices they do, of course, occur elsewhere, offline. Equally, there are computational technologies that are not Black Boxes and they are often intentionally and vociferously transparent. One umbrella that such tools fall under is ‘Free Software’, a movement initiated by Richard Stallman upon announcing his intentions to create an alternative to the UNIX operating system that would be free from usage restrictions. It is not possible, due to space restrictions to discuss this in depth here but, for completeness, a brief outline is necessary. Free Software provides its users with a number of freedoms, ranging from access to the source code of the programs they run to the freedom to modify and redistribute them. The word ‘free’ refers not to monetary freeness, but liberty, allowing users to move away from the control exercised by software. This freedom, however, is very localised and can only be exercised within an environment over which the user has complete control to install this software. Furthermore, freedom to modify it means that, while one’s own copy may be privacy enhancing, a similar program elsewhere could be problematic. This is compounded by the fact that many of the devices users interact with in the course of using computational technology are networked and, as we will see in the next section, they present this with varying degrees of legibility to their users, making it difficult to identify the processes and making it impossible to ensure a fully free interaction.

7 https://www.eff.org/issues/drm (accessed 2nd June 2014)
8 https://www.eff.org/issues/cf aa (accessed 2nd June 2014)
11 Free Software is sometimes, for this reason, referred to as Libre Software, or FLOSS (Free, Libre, Open Source Software).
Opaque Interfaces

Digital consumer devices and services do not always offer up legible presentations of what, exactly, they are doing. They offer interaction through graphical user interfaces (GUIs) designed to make it easy and quick for the user to achieve their aims, and the fields of Human Computer Interaction and User Interface/User Experience (UI/UX) Design have focused, in part, on these interfaces, using affordances and metaphors\(^\text{12}\) to achieve smoother, more usable, software tools. This, however, often hides the underlying processes and activities of the devices in use.

The Digital Personal Assistants found on many smartphones, such as Siri for Apple’s iOS\(^\text{15}\), or Google Now\(^\text{15}\) for Android and iOS based systems, are services which allow you to “use your voice to send messages, schedule meetings, make phone calls and more”\(^\text{16}\). These pieces of software, as presented by their interfaces, listen to your voice commands and queries, responding to them by scheduling items in your calendar, setting alarms or writing emails. Beyond this, they are able to search for facts and nearby amenities. The interface for Apple’s initial release of Siri\(^\text{16}\) used speech bubbles, similar to those found in the iOS messages application (fig. 1), to present a conversational view of the user’s interaction with Siri\(^\text{16}\). This, coupled with the fact that the assistant can act on programs within the phone, gives the impression that it inhabits the phone, that any processing or cognition occurs within the user’s device. If this is the case, although the user may not understand the exact way in which it works - it is certainly a Black Box - they surely have access to all the inputs (voice requests, the information on the device) and outputs (organisational software, the screen, the speakers) necessary to conduct an investigation of the Black Box.

This is not the case. It is, in fact, a carefully constructed fabrication. The conversational metaphor, the presentation, and the pacing of the interaction, hide a necessary offloading of processing. The devices on which these assistants run are not powerful enough to undertake the complex analysis required for the illusion of human-machine conversation, they must send the audio to remote, high powered data centers for processing\(^\text{14}\). In this sense, the device and its personal assistant are really just a node in a larger network, passing information to, and receiving it from, remote servers sometimes known as the ‘Cloud’.


\(^{13}\) https://www.apple.com/uk/ios/siri/ (accessed 2nd June 2014)

\(^{14}\) http://www.google.co.uk/landing/now/ (accessed 2nd June 2014)

\(^{15}\) https://www.apple. com/uk/ios/siri/ (accessed 2nd June 2014)

\(^{16}\) Here, Apple’s original version of Siri, released on October 4th 2011, will be used to explore the ways that User Interfaces can obscure the actions of software. For more information, see: http://www.apple.com/uk/pr/library/2011/10/04Apple-Launches-iPhone-4S-iOS-5-iCloud.html (accessed 10th August 2014)

\(^{17}\) Siri is the name of the piece of software, but is also used as a ‘name’ for the ‘assistant’ - the voice the user hears.

While this is, perhaps, fine while virtual assistants are pleasant, seemingly neutral, servants of the most banal scheduling needs, with some entertaining gimmicks thrown in\textsuperscript{19–22}, what happens when they are relied on for more personal, private problems? What does it mean for something understood as a personal service to be part of a large external network, and what happens when these tools reveal themselves as especially partisan gatekeepers of knowledge?

Upon launch, it was discovered by users that “Siri failed to locate nearby abortion clinics. In some cases it suggested pregnancy advice centres as an alternative”\textsuperscript{19–22}, providing a biased, or skewed, filtering of information. Further to this, around seven months after Siri’s launch, asking it the question “What is the best smartphone ever?” provided a result from the third party computational search engine Wolfram Alpha\textsuperscript{19–22} naming a recently released phone by rival company Nokia. A few days later it had ceased to do so, instead providing ‘humorous’ replies such as “Wait... there are other phones?”\textsuperscript{19–22}. The unwillingness to locate abortion clinics was eventually rectified by Apple, who asserted that it was not intentional, but the same capability to increase the impartiality of the tool was used, later, to control access to information on rival devices.

When a change is observed in, or a problem found with a tool (or a multi faceted service, masquerading as a tool) such as this guidance can be taken from Ashby, the engineers, and cyberneticians: the next step is to interrogate the tool as a black box.

Firstly, the inputs need to be monitored. These are knowable, surely: on the simplest level they’re the queries the user speaks and, delving a bit deeper, the settings on the phone and the information in its address book, calendar and so on could be considered an input. Then, the outputs from the system should be measured. These are the qualifying statements and questions the assistant uses to focus the human language it receives into machine readable queries, and the results the assistant returns from these.

Though the inputs are adjusted and the outputs are compared scientifically and systematically, it’s still likely that the results will contain a strange bias. As a result, expanding the area of investigation can lead only to the realisation that this device cannot possibly contain so much information about such a range of topics, including those that are time and location sensitive, it must be querying information services elsewhere (such as the aforementioned Wolfram Alpha). It is, therefore, not just a Black Box but, rather, a Very Large Black Box. Leaving aside their own status as Black Boxes, the search engines and other sources the assistant queries can be questioned in the same way it does\textsuperscript{20–22}, but still different results will be received. For example, while Wolfram Alpha informs the user that the best phone to buy would be a Nokia, Siri is insistent.

The conclusion to be drawn, then, is that, despite being presented as a self contained, localised system, this piece of software is, in fact, a node in a distributed, networked system containing both

\textsuperscript{19} http://shithatsirisays. tumblr.com/ (accessed 3rd June 2014)
\textsuperscript{20} http://www.bbc. co.uk/news/ technology-15982466 (accessed 3rd June 2014)
\textsuperscript{21} http://www.bbc. co.uk/news/ technology-18071342 (accessed 10th August 2014)
\textsuperscript{22} ibid.
\textsuperscript{23} This requires a fair amount of assumption - the nested levels of Black Box involved here start to bring in a fair amount of opacity
publicly accessible services and inaccessible proprietary services - the hidden systems within Apple’s data center. It is certainly a Black Box, but now revealed as an Incompletely Observable Black Box. Aspects of the service are intentionally hidden for two main reasons. Firstly, the purposes of user experience, allowing the user to feel they are able to query their device directly which gives an illusion of efficiency, sleekness and privacy. Secondly, for the purposes of business, both overtly, ensuring the presentation of only Apple’s services, and covertly, hiding the systems involved, the precise algorithms and data behind such services being a valuable trade secret.

Although one of the examples given above is ostensibly playful and tongue in cheek, they both show that the power lies with the software as a gatekeeper of knowledge. While it is possible for the user to know how a piece of local software functions by observing it, much like a traditional electronic device, when much of the functionality is external it is possible, and almost guaranteed, that it will change, constantly shifting the power relationship between user and device.

**Leaky Connections**

While seemingly local devices may hide their true, networked, nature behind user interfaces, what of the internet, the World Wide Web, surely there’s no artifice there? It’s part of a web browser’s presentation, via the address bar, to show that a website is separate from the user, the content changes without their control and when there’s no connection, there’s no access. However, while the user may be aware of their connection to a network, it’s not necessarily clear what, or how many, parts of that network they are connected to. It’s normal to visit a web address (for example, http://www.bbc.co.uk/news/) which returns specific content and one might logically assume that signifies a connection only to the single place that content is held, in this case on the BBC’s servers.

It is possible to test this assumption. Many modern web browsers, such as Google Chrome provide tools for web developers to profile the sites they are building, to check loading times and other aspects of their performance, and these tools can be used on any publicly accessible site. Visiting news.yahoo.com24 and viewing the network connections made, it can be seen that, in total, there were 100 requests made in the 10 seconds it took the site to load. Of these, 57 were from domains identifiably related to Yahoo, such as ads.yahoo.com, or l.yimg.com, with the other 43 coming from other domains with variable levels of identifiability: googleads.g.doubleclick.net explains itself, to a certain extent, while something like s1.2mdn.net offers little in the way of clues25. This can be seen on many sites across the web, bringing in data related to advertising, page content, user tracking, page analytics and so on.

---

24 http://news.yahoo. com was the most visited news site according to Alexa’s top 500 sites list (http://www.alexa.com/topsites/category/Top/News) at the time of visiting (28th August 2014). Except for http://reddit.com, which I’m even more unsure about classifying as a news site than I am Yahoo’s.

25 2mdn.net is registered to Mark Monitor, a ‘brand protection’ provider and a part of Thomson Reuters. But visiting the page returns nothing.
Some of these further requests came from the page initially loaded - for example adding the images that were part of the layout, or scripts to add functionality – but others were loaded in by other scripts. For example, inspecting one resource from http://ams1.ib.adnxs.com shows that it is a piece of javascript which loads in a further piece of javascript from http://cdn.adnxs.com which itself loads content from, or sends data to http://ib.adnxs.com. Visiting the last URL provides nothing human readable, so it’s logical to assume that it’s providing information to a server.

Some browsers do provide an indication of the connections a site makes, often as a small tab in the bottom left of the window which displays the current resource being loaded, but this not often particularly visible and ordinarily operates in real time which, given the speed of networks (UK average, 17.8Mbits\(^\text{27}\)), and the small size of the resources loaded means that it’s unlikely that users will see all of them, or even necessarily perceive that there are any at all.

Coupled with the inability to see the processes occurring on external servers, beyond the URLs accessed and the files they return, the nebulous nature of visiting a web pages shows the internet to be an Incompletely Observable Black Box. It is possible to get a sense of the requests and actions undertaken by a web browser but, as with the third URL linked above, the investigation can be quickly brought to a halt by a further Black Box whose inputs can be seen, but whose outputs are hidden.

These outputs can, therefore, lead anywhere and it’s very difficult, if not impossible to know where this might be, at least just by investigating the technology available to us. The documents released in early 2013 by NSA whistleblower Edward Snowden make it clear that it’s often impossible to know what’s being done with our digital information, the extent to which this data is collected, redirected, stored and analysed means we have little to no idea exactly what’s occurring, when and to whom.

They do, however, go some way towards revealing some of what might be happening to our data. The documents reveal ‘Optic Nerve’, a program from the British agency Government Communication Headquarters (GCHQ), which collected images from the webcam feeds of Yahoo webcam users, storing them for analysis. In this case Yahoo, the provider of the service, “denied any prior knowledge of the program, accusing the agencies of ‘a whole new level of violation of our users’ privacy’”\(^\text{26}\), but the problems are similar whether the service provider is aware, or not.

Networks, especially the internet, necessarily require passing of information to multiple parties, if it is not possible to know what happens at each step, with each party, then it is not possible to know, with certainty, every resting place for the data sent, and what has been, or is being, done with it. By trusting conversations and images to a service that functions as a Black Box, users are, in turn, opening their data to a larger network of collection, sending it to unknowable numbers other Black

\(^{26}\) HTML, the primary language used to markup and build websites, provides only structural and textual content within its files, further resources such as images or videos are stored as separate files on the same, or different servers and referenced from the HTML.


Boxes. Optic Nerve is just the program that has been made public. It is not possible, without further leaks, to know what, or how many, other similar programs there may be.

Talkative Objects

Lacking a full understanding of the machinations behind an optional digital assistant on high-end smartphones, and having assets loaded into web pages behind the scenes could be seen as optional problems. After all, if one doesn’t want to talk to a smartphone, search engines are still available, and if the internet seems too opaque then there are still libraries and public information centers, at least for the time being. The problem is only one part, of some parts, of two specific devices: the smartphone and the personal computer.

However, as part of the trends towards wearable technology and an ‘Internet of Things’, the number of devices on this network is growing in number and in variety of forms. Things that are traditionally singular, isolated objects are increasingly becoming computationally able, sensor rich, networked devices. These devices, formerly mute and static, are becoming talkative objects, with access, often very close, to information about things such as health, finances, relationships and with access to the wider network, they gain the ability to broadcast this to interested parties.

In his talk ‘The Coming War on General Computation’ Cory Doctorow highlights the changing relationship we have with computers:

“As a member of the Walkman generation, I have made peace with the fact that I will require a hearing aid long before I die, and of course, it won’t be a hearing aid, it will be a computer I put in my body. So when I get into a car – a computer I put my body into - with my hearing aid - a computer I put inside my body - I want to know that these technologies are not designed to keep secrets from me, and to prevent me from terminating processes on them that work against my interests.”

These computational, wearable and drivable devices are not just computers in terms of having processors, and running code, they are also likely to be networked devices too. In their paper ‘Security and Privacy in Implantable Medical Devices and Body Area Networks’ Rushanan et al. describe medical devices with the capability for “wireless data transfer (or wireless medical telemetry) for monitoring and configuration without sacrificing patient mobility or requiring surgical procedures to physically access the devices”, revealing that these devices are implanted in more than 25 million patients in the United States. The paper goes on to describe the difficulties involved in analysing the security of these devices, with “proprietary protocols” making it difficult

for traditional tools to develop methods appropriate for analysing them. That the devices are seen as a Black Box even by the security researchers attempting to improve them makes it incredibly difficult to imagine gaining the assurances of control and knowledge Doctorow wants, especially at the level of the end users of these devices.

It may be unusual to consider a pacemaker, an insulin pump, or a hearing aid as a networked device, but there will be moments in their use where it becomes clear that they are. When a doctor wirelessly looks up the pacemaker’s activation data to check that it’s working correctly, for example. There are other devices, however, that may never make clear their talkative nature either because they are built to hide it, or because they are formerly mute objects and have been retrofitted to become vocal.

Germany’s Der Spiegel, one of the recipients of Snowden’s leaked documents published a number of pages from what they termed “The NSA’s Spy Catalog”\(^3\), a series of specifications, use cases and datasheets for some of the intelligence service’s surveillance equipment. As part of this, they reveal devices that can be implanted in computers, routers and other pieces of network infrastructure, or placed within the signal chain to allow access to the information within them. For example, the ‘RAGEMASTER’ (fig. 2) is a 6mm long piece of circuitry, whose operation is outlined as follows:

> “The RAGEMASTER taps the red video line between the video card within the desktop unit and the computer monitor, typically an LCD. When the RAGEMASTER is illuminated by a radar unit, the illuminated signal is modulated with the red video information. This information is re-radiated, where it is picked up at the radar, demodulated, and passed onto the processing unit, such as a LFS-2 and an external monitor, NIGHTWATCH, GOTHAM, or (in the future) VIEWPLATE. The processor recreates the horizontal and vertical sync of the targeted monitor, thus allowing TAO personnel to see what is displayed on the target monitor.” \(^3\)

Essentially, this tool silently copies the video signal being sent to a computer screen, while still allowing the original signal through, and can be queried by an external device to allow the intercepted imagery to be seen on a display external to the system, room and, potentially, building it is installed in.

In his keynote ‘Art As Evidence’\(^3\) at transmediale 2014, Jacob Appelbaum, a security researcher and journalist, and one of the first people to gain access to the Snowden files, discusses some of these devices and their implications. He describes the act of ‘interdiction’ undertaken by the NSA calling it “the process whereby all of those objects previously mentioned, gain a little extra attribute”. Using an example of a keyboard ordered online by a fellow developer on the TOR

---

\(^3\) [http://www.spiegel.de/international/world/a-941262.html](http://www.spiegel.de/international/world/a-941262.html) 30th December 2013 (accessed 30th August 2014)

\(^3\) [ibid.](http://www.spiegel.de/international/world/a-941262.html)

\(^3\) Jacob Appelbaum: *Art As Evidence*. Transmediale 2014 Keynote. [http://www.youtube.com/watch?v=ndx0eox0Lkg](http://www.youtube.com/watch?v=ndx0eox0Lkg) (accessed 5th June 2014)
he outlines the way that interdiction is used to intercept the package and, potentially, augment it with these tools for listening and broadcasting. As he says, it is not necessarily possible to be sure that interdiction, or other interference has taken place, and I would argue that this continues even if the devices are opened up and inspected. It is mentioned many times in the catalogue pages that the devices are often made from off the shelf components to ensure they cannot be traced back to the NSA.

Intercepting but not interfering with existing communications and piggybacking on existing networks allow these surveillance Black Boxes to take advantage of the existing Black Boxed layers within the technology they surveil. By trusting our data and communications to one or more sets of Black Boxes, we open it up to many more. As Appelbaum goes on to say “When the physical world and the internet come together, there are these convergence points in which everyday objects are actually weaponised and turned against us”.

**Conclusion**

The process of identifying objects, tools, as Black Boxes and, thereby, as potentially talkative, opaque or leaky actors rather than purely as sleek servants of our desires is an important one. It is the first step towards countering their control and coming to an understanding of the realities of a digital, connected world. As shown above, the Very Large and Incompletely Observable nature of the Black Box can make this identification tricky, but we have seen three core outcomes of the Black Box and, through identifying these, we can begin to hone in on further examples. Firstly, by understanding that these objects which might, at first glance, seem to be self contained and knowable, are actually opaque interfaces to a nebulous, concealed network, we can begin to direct our investigations. From this awareness that a connected device could be acting, through its leaky connections, in multiple, undisclosed ways with the information we provide it helps us to reconsider that provision, thereby reducing its knowledge and potential for control. These two considerations allow us to regard the connected devices we use as talkative objects, systems that report our actions and intentions to remote actors, either directly and intentionally or through offering themselves as easily hijacked listening points. This delineation and description of the Black Box is the first step in being able to act upon and observe it, and thereby counter it. We’re now able to scrutinise it as Ashby’s engineer might probe a mechanical object, looking for flaws or gaps, glitches and failings, windows to understanding hidden deep inside the Black Box.

---

35 TOR is a distributed network that helps provide anonymity for its users. More about TOR can be found here https://www.torproject.org/
Fig. 1.
Side by side Siri and messages screenshots showing the similarity of the interface metaphors.
Fig. 2.
The RAGEMASTER datasheet. Showing the device, and outlining some of its specifications.