

Data adsorptents, data emitters and databases in politics

by

Amelia Andersdotter

This essay intends to give an introduction to the politics of data protection and the challenges that face users of new technologies, politicians and society. Our everyday lives, our commerce and interaction with public institutions are now built around information. This is increasingly becoming the primary source for power and influence. In particular, those with more information have acquired more power and more influence. This text describes the proliferation of technologies surrounding us, a few possibilities facilitated by those same technologies and some aspects of empowering everyone interacting with them with respect to data control. I've chosen to describe this new reality through three core concepts:

Data adsorptent¹ is the term I intend to use to describe the emerging myriad of technologies with a primary purpose of gathering and processing data. Computers are obviously an example, and so are mobile phones. But software and applications installed on these devices can also act as data adsorptents – in the case of smartphones, internet usage, phone calls, text messages and online applications increase the data collectivity even further. In Sweden we have small data adsorptents installed in every house

¹ Adsorption is a term used in chemistry to describe when molecules from a gas or a liquid (high energy molecules) attach themselves to a solid surface (low energy molecules). The resulting compound is usually used as a catalyst or tool for other processes. It describes the behaviour of technologies collecting, processing and storing data fairly well.

connected to the electricity grid since July 1, 2009², designed to supply electricity providers with enough information about the consumption of each household to individually optimize power transmissions through the grid. Traffic cameras collect data about cars passing by. Data adsorptents collect, briefly put, data from their surroundings and usually give it a meaning in a context³.

A data emitter is, similarly, someone who broadcasts data to their surroundings. It can happen consciously or unconsciously, and the broadcast data either persists with the data absorptent afterwards or is of a more transient nature. Imagine for instance the act of using a stove, and in that way broadcasting information about your cooking habits to the world, or having a fixed, non-secret phone number and in that way broadcasting your location to anyone you call.

Gathered data is often catalogued in databases. Sometimes they are very large, like the Google registers of performed searches matched with IP-addresses. Sometimes they are smaller, like the member register of a local chess club. When databases are electronic, they can often be cross-linked and we get something called profiling. If you happen to cross-link the electricity provider data on power consumption with the records of sold appliances to that household you could create a fairly accurate ecological footprint of the household. Mapping an entire city gives a fantastic image of ecological footprint variations over blocks and city regions.

A library is sort of a database. It's a catalogue of information emitted by somebody (an author, for instance) and later collected by somebody (a librarian perhaps) and organised into a database (or bookshelf). A data emitter whose data ends up in a library usually has a good understanding of what data they've broadcast

² http://www.lundsenergi.se/Kundtjanst/Fjarravlasning/Fragor_och_svar_fjarravlasning

³ This could, for instance, bring about a future augmented reality.
http://en.wikipedia.org/wiki/Augmented_reality

and how it's presented. They have the possibility to anticipate the effects, with readers quoting them in surprising situations being the exceptions.

In the information society, it is not as straightforward. Neither emission, collection, or subsequent use of the data is entirely straight forward. Consider for instance the well-known copyright controversy, where data is released in the form of an artistic work, enters into peer-to-peer network databases and ends up getting spread to and used by people all over the world in unanticipated ways. In the digitised environment it is, just like in the library, they who have the most knowledge about the databases and their content who are relatively more powerful. Librarians can help visitors find books, they usually have access to the record of books the visitor has previously borrowed. In the ubiquitous information technology environment the librarians are Google, Apple, E.ON or Albert Heijn (a Dutch supermarket chain). These new librarians have also limited access to their systems of collecting data and how they organise it: hardware, firmware and software are closed to users. Whoever has the most precise and abundant data catalogued about the most objects will have the upper hand on the objects the data concerns, and the current infrastructure is making it very difficult for users to ensure that companies are living up to their promises, or devising ways of escaping their data collection.

Europe has a tradition of public libraries. The libraries have been open to the public, membership has been more or less free, and membership mostly implies full access. The present information infrastructure is nowhere near similar. On a European level we are discussing net neutrality, tiered pricing and full restriction of access to connectivity for users acquiring data from the network. In emission control, the critique of Apple's application distribution model for their smartphones is a good example: For a long while, it was impossible to install and use non-Apple verified applications on

iPhones⁴. Apple reserved the right to scrutinize all applications prior to approving their inclusion in the official Apple store. In this way, the user was completely dependent on Apple, and they controlled completely what you could and could not do with your phone. Apple still enforces this policy to some extent, and jailbreaking still renders your warranty void.

The American lawyer Lawrence Lessig has written “code is law” and that they who control the code control the digital environment. The Apple store illustrates that. But the information society we're currently building is not limited to the Apple store or the online social networks of Lessig⁵. On the contrary, if your windshield is wound up and down by electronics, these will have all the possibility in the world of storing data about your windshield usage without you ever knowing or being able to hinder it. It is trivial to collect data about when and for how long the engine has been started and what the power output was or tracking and storing information about the path the car has travelled. A data emitter can in many cases be happy about this or at least not mind. The data collection can possibly provide useful services, like road-finding or alerting the car owner of problems with the car. But by driving we are also effectively powerless to cease our data emissions even in those situations we do not wish to receive extra functionality – the necessary information about the data collecting units is just not accessible enough.

If you're a parliamentarian discovering these issues, you will probably sit down immediately to search for legal solutions to the problem of data processing. Smart meters on the electric grid should not be obligatory, they can be installed at the last transformation point of the grid prior to the households. Anonymous OV-chipcards⁶ on public transports can not cost more money than

⁴ http://en.wikipedia.org/wiki/Jailbreak_%28iPhone_OS%29

⁵ <http://pdf.codev2.cc/Lessig-Codev2.pdf>

⁶ An OV-chipcard is a commutation card implemented in many public transport systems around Europe.

personalised one: privacy is an unconditional right. Profiling of citizens by means of cross-linking databases between companies or even within companies should be further limited.

In the regulatory debate, an analysis of the imbalance caused by some parties having the ability to acquire and analyse lots of data about other parties, even with consent, is not very present: the EU Data Protection Supervisor Peter Hustinx, has underlined at several occasions the importance of users having access to and control over data already adsorped: to know about and consent to what data is being collected about you, and the right to demand that the party adsorping that data ceases to make use it. Access to and the ability to alter the information infrastructure for private use is rarely discussed at an EU level, and especially not in the context of data emission control or access to such control.

We need further discussion on a political level about direct control over data adsorpting technologies. People moving in an ICT-networked society reasonably should have the ability to control and customize their data emissions on their own, and not be limited to post-emission access and removal demands⁷. A persistent informed consent approach risks causing more inconvenience for people moving around in society than it enhances privacy protection, especially considering the absorptent will have to interact with an emitter to acquire the consent. A solution must instead include creating a proper system of technological accountability, where the ubiquitous technologies can be researched, investigated and understood by anyone who is exposed to them. It will be a challenge for legislators to find ways of granting people that control. Data emitters will also have a challenge in finding their desired level

⁷ It's often not trivial to remove data safely either. A typical scenario is that the data is not deleted and just marked as being able to being over-written, which is not sufficient data to be unrecoverable. See Gutmann, 1996:
http://www.cs.auckland.ac.nz/~pgut001/pubs/secure_del.html

of emission and adjust them according to time, place or data adsorptent.

Most people are probably unaware of when and how their data is being collected, and it's becoming increasingly difficult to keep track of all the situations where data adsorptents are introduced for optimization purposes. Technologically it's at a level beyond most people's skills, but for preventing misuse of data adsorptents it is crucial that people have the ability to develop the necessary skills for controlling data adsorptents.

We could do this by opening up technical specifications, architectures and interfaces. This would allow for a user-driven development of parallel technologies hindering or limiting data emitter output. It would create a greater stress on technology providing or data gathering enterprises to fulfill their obligations with respect to data protection legislation, since they would be scrutinized by as many people as have the desire to do so. The openness of the technology removes the problem of interoperability, since the choice for interoperation rests on the parallel developer. There is a potential for emitters to manage their emitted information more conveniently, for instance as is the case when one user has loans with many banks and wants to control them via one and the same interface⁸. The ability to uphold user's rights and freedoms in society would be moved towards the users, instead of resting with public institutions.

Some attempts to create technical solutions to manage data emissions on the user side are RFID Guardian⁹ and De Privacy Coach¹⁰. They detect data adsorptents in their environment and

⁸ This is essentially not possible today. Very closed structures in the banking and financing industries are currently being reviewed by the European Commission. The security by obscurity also makes it next to impossible for new market entrants to get onto the market.
http://ec.europa.eu/internal_market/finances/index_en.htm

⁹ <http://www.rfidguardian.org>

¹⁰ http://www.difr.nl/?page_id=10

allow users to set privacy levels for each collector in their ambience. To create such devices, standards in technology is vital. Interoperability – when software or hardware can co-operate with each other – plays an important role in being able to make such privacy guarding devices with universal applicability.

But if we imagine that privacy guards become a commercial product, it's easy to see how standards and interoperability does not make the power imbalance go away. The standards and interoperability are provided by the data adsorptents, and agreed between data adsorptents. Public institutions are ready to take on the role of independent privacy evaluator of data adsorptent agreements and the post-emission data control. The data emitters themselves are disadvantaged: the right to evaluate and keep check of technology ends up with the commercial collectors or public institutions. Actual user ability to create their own emission control mechanisms and privacy quality checks of the data processing systems is small. The prohibition to reverse engineer patented new technologies (including for research purposes) enforced in most European states further aggravates the imbalance.

The ideological conflict between free or closed information infrastructure is old. We have the idea that developer autonomy with regards to deciding the level of openness in their hardware or software must be preserved on the free market, but we are also struggling with how people in a networked, digitalised society must be able to maintain their rights with respect to states, companies and other people.

Society needs to allow for all types of voluntary emission, and all types of voluntary reception. It also needs to create abilities for restriction of emission and reception. It seems, at least to me, that we have very few ways of achieving such a structure without making our infrastructure open and free.