# Bulk data collection and privacy

## Description



NSA whistleblower Edward Snowden said that when crossing a busy road he instinctively looked away from oncoming traffic for fear of having his image captured on a dashcam. People are more easily recognised face-on than in profile.

That short observation from his book *Permanent Record* delineates some salient themes in the so-called smart city: risk, paranoia and surveillance.

## Nobody is listening

What he and other insiders revealed was that the NSA was able to obtain wholesale communications data from every citizen on the phone network (in the USA and abroad). The initial revelation in *The Guardian* published on 6 June 2013, stated that for telephone communications

> â??the numbers of both parties on a call are handed over, as is location data, call duration, unique identifiers, and the time and duration of all calls.â?•

The NSA was not requesting from the communications service providers the content of conversations, but the meta data.

## Contentless data

We normally think of secret service investigations as directed at key targets, but here the data is collected for everyone on the system whether or not they are suspects. The data is stored in bulk on vast servers ready to be mined as needed. Links can be traced and patterns found. The collection is automated, and no human being need ever see the data â?? unless they are authorised to investigate particular individuals.

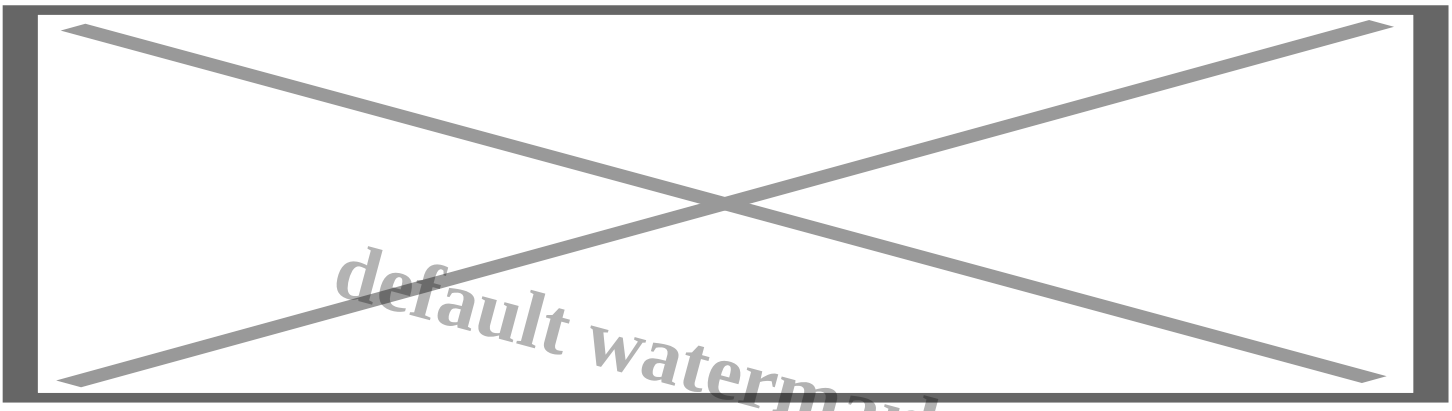
Snowden and others have argued that much about a personâ??s life can be harvested from such meta data, including networks of contacts, lifestyle, activities and competencies. The information field is even more revealing if you include the bulk collection of email metadata, browser histories, debit card

transactions and travel data, especially if linked.

A blog by IT law lecturer Paul Bernal explains the problem of identifying when surveillance actually happens. There are three key moments:

> â??the gathering or collecting of data, the automated analysis of the data (including algorithmic filtering), and then the â??humanâ?? examination of the results of that analysis of filtering.â?•

Does surveillance happen when the bulk data is collected, or when humans inspect the data?




## Surveillance or privacy risk

Bernal explains that the main question is when privacy is put at risk. He writes in terms of *privacy invasion*. I prefer â??riskâ?• as I donâ??t want to presume that thereâ??s such a thing as *complete* privacy, all of the time, in all places and relating to every aspect of a personâ??s life.

According to Bernal, the privacy risk is obvious in the case of installing CCTV cameras in someoneâ??s home connected to some outside agent. The placement of the cameras would strike most people as already providing an extreme risk to personal privacy.

The privacy question arises immediately the moment the surveillance system is set up, when thereâ??s the means for someone (a relative, the landlord, the boss, a law enforcement official) to see what the camera sees, even if they donâ??t ever take the opportunity. Bernal puts the emphasis on *privacy*, about which there are rights and laws, rather than *surveillance*.

By that analogy, setting up a bulk data collection facility initiates the risk to privacy, whether or not surveillance actually happens. Bernal also highlights the intermediate algorithmic processes, which are now assuming greater power and prominence. No human being may ever inspect the data, but will only see the outcome of algorithmic pattern detection. In that case the installation of these algorithms plays a major part in privacy risk as well.

By this reading, urban infrastructures designed to soak up communications data are already complicit in the privacy challenge. Thatâ??s sobering in the context of smart city infrastructures. See post: Reverse Analytics.

## Bibliography

- Bernal, Paul. 2016. Does the UK engage in â??mass surveillanceâ??? *Paul Bernalâ??s Blog: Privacy, Human Rights, Law, The Internet, Politics and more*, 14 January. Available online: https://paulbernal.wordpress.com/2016/01/15/does-the-uk-engage-in-mass-surveillance/ (accessed 27 November 2019).
- Goodwin, Bill. 2017. Mass collection of data on population â??illegalâ??, UK court told. *ComputerWeekly.com*, 5 June. Available online: https://www.computerweekly.com/news/450420162/Mass-collection-of-data-on-population-illegal-UK-court-told (accessed 27 November 2019).
- Greenwald, Glenn. 2013. Edward Snowden: the whistleblower behind the NSA surveillance revelations. *The Guardian*, 11 June. Available online: https://www.theguardian.com/world/2013/jun/09/edward-snowden-nsa-whistleblower-surveillance (accessed 27 November 2019).
- Greenwald, Glenn. 2013. NSA collecting phone records of millions of Verizon customers daily. *The Guardian*, 6 June. Available online: https://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order (accessed 27 November 2019).
- MI5. 2019. Bulk data. Gathering Intelligence. Available online: https://www.mi5.gov.uk/bulk-data (accessed 27 November 2019).
- Moody, Glyn. 2016. Whatâ??s The Difference Between â??Mass Surveillanceâ?? And â??Bulk Collectionâ??? Does It Matter? *Techdirt*, 20 January. Available online: https://www.techdirt.com/articles/20160115/09582933351/whats-difference-between-mass-surveillance-bulk-collection-does-it-matter.shtml (accessed 27 November 2019).
- Poitras, Laura. 2014. *CitizenFour*. HBO Films
- Snowden, Edward. 2019. *Permanent Record*. London: Macmillan
- Stone, Oliver. 2016. *Snowden*. Open Road Films




**Category**

1. Ethics

**Tags**

1. Big Data
2. privacy
3. smart city
4. Snowden
5. surveillance

**Date Created**
November 30, 2019
**Author**
rcoyne99