4 SUMMIT VIENNA 2015 IS IS

Extended Abstract

Jus Algoritmi: How the NSA Remade Citizenship

John Cheney-Lippold [1]

[1] University of Michigan / 500 S State St, Ann Arbor, MI 48109, United States of America / jchl@umich.edu

Introduction

It was the summer of 2013, and two discrete events were making analogous waves.

First, Italy's Minister for Integration, Cécile Kyenge was pushing for a change in the country's citizenship laws. After a decades-long influx of immigrants from Asia, Africa, and Eastern Europe, the country's demographic identity had become multicultural. In the face of growing neo-nationalist fascist movements in Europe, Kyenge pushed for a redefinition of Italian citizenship. She asked the state to abandon its practice of *jus sanguinis*, or citizenship rights by blood, and to adopt a practice of *jus soli*, or citizenship rights by landed birth.

Second, Edward Snowden fled the United States and leaked to journalists hundreds of thousands of classified documents from the National Security Agency regarding its global surveillance and data mining programs. These materials unearthed the classified specifics of how billions of people's data and personal details were being recorded and processed by an intergovernmental surveillant assemblage.

These two moments are connected by more than time. They are both making radical moves in debates around citizenship, though one is obvious while the other remains furtive. In Italy, this debate is heavily ethnicized and racialized. According to *jus sanguinis*, to be a legitimate part of the Italian body politic is to have Italian blood running in your veins. Italian meant white. Italian meant ethnic-Italian. Italian meant Catholic. By reshaping citizenship standards to practices of *jus soli*, Italian might also come to mean black, brown, Muslim or Hindu. Ultimately, in this case, either essence (*sanguinis*) or presence (*soli*) founds the two main competing theoretical orientations for how citizenship, or "the rights to have rights", is allocated by the contemporary nation-state.

Cue the June 2013 Snowden NSA leaks, when the world learned of ubiquitous state surveillance being conducted on nearly all of the world's Internet and telephony networks. And consider the subsequent concerns around privacy that such a wide-ranging surveillance practice would require. How could the NSA know, in exacting fashion, who was and was not a US citizen? The technicality of this is impossible. But rather than limit the scope of surveillance to acknowledge this impossibility, the NSA had a better idea: why not create a new conception of citizen?

And that's what they did. The NSA decided to implement an algorithmic assessment that assigned foreignness (and by corollary, citizenship) on targeted Internet traffic, enabling legal surveillance if a data subject was deemed to be at least "51 percent foreigner" [1]. This mode of citizenship is what I call *jus algoritmi*. It's a citizenship that, unlike its *soli* and *sanguinis* kin, is not ordained and stable. It's a relationship to citizenship that is temporal and constantly evaluated. One day you might be a "citizen" of the US; another day you might be a "foreigner". *Jus algoritmi* "citizenship" is a categorical assessment based on an interpretation of your data, not your lineage, birth certificate, or passport.

Methods

The key legal justification for all of these programs is importantly echoed in declarative emphasis by an NSA press statement following the first weeks of the Snowden leaks: "NSA's activities are focused and specifically deployed against — and only against — legitimate foreign intelligence targets… [and] all of NSA's analytic tools are limited to only those personnel who require access for their assigned tasks"[2].

While this is not the case for the entirety of the NSA surveillance suite (metadata, as famously argued over the past year, does not cleanly fit into existing US privacy law), the broad stroke by which the NSA attempted to assuage privacy invasion fears seems to always arrive with the mantra of "legitimate foreign targets". The idea of the "legitimate foreigner" is endlessly used as both the justification as well as the deterrent of worry. It defends the programs' mission by always pointing to the dark, brooding "legitimate foreigner". And it seemingly maintains privacy for those of us who know we are not, and could never be conceived to be, foreign.

But there is no direct way to connect a computer's IP address, MAC address, or even profile/email account verifiably to a sense of foreigner or citizenship. The "legitimacy" prefixed upon the foreigner seems to have become more and more pro forma. And so we arrive at the raison d'être for *jus algoritmi*. The NSA determines who is a "citizen" and who is a legitimate "foreigner", in the words of the NSA, by "analysts who use the system from a Web portal at Fort Meade, Md., [to] key in "selectors," or search terms, that are designed to produce at least 51 percent confidence in a target's "foreignness"[3].

The question of "how" this works begins in November 2013, when the New York Times revealed the NSA's SIGINT Strategy for 2012-2016[4]. This document outlines the NSA's plan for the "SIGINT battle space". SIGINT, or signals intelligence, is spy speak for communication between individuals or data. A conversation is SIGINT; so is my email address. In even more layman speak, it's the stuff that surveillance agents hoover up. And in this strategy memo the NSA felt it "must proactively position [itself] to dominate that environment across discovery, access, exploitation, analysis… The SIGINT system and our interaction therein must be as agile and dynamic as the information space we confront".

Continuing even further, the memo states that "for SIGINT to be optimally effective, legal, policy, and process authorities must be as adaptive and dynamic as the technological and operational advances we seek to exploit". Within this five-page document, several goals are declared to dynamically dominate the SIGINT battle space. For the purpose of this talk, two goals in particular should be noted:

"4.2. (U//FOUO) Build compliance into systems and tools to ensure the workforce operates within the law and without worry"

"5.2. (U//FOUO) Build into systems and tools, features that enable and automate end-to-end value-based assessment of SIGINT products and services"[5]

Point 4.2 focuses on in-system compliance, or the automation of NSA operations according to US privacy law. Point 5.2 focuses on automated appraisal of SIGINT products and services, or potentially the use of algorithms to evaluate how SIGINT is assessed. That is, to be "optimally effective" the NSA has to develop "adaptive and dynamic" "legal" processes to gather SIGINT. Across the databases of PRISM, Traffichief, Pinwale, and MARINA, the NSA had moved to what Axel Arnbak calls "automated oversight"[6].

How in the world would you be find out 51 percent confidence in a target's foreignness? We can think about the algorithmic categorizations of gender, race, and class — of how Google takes your search history and determines your gender based on which words you've queried and which sites you've visited. Or how marketing companies can assess your race/ethnicity through wide-ranging surveillance of your browsing and purchasing habits. For these categorizations, queries, sites, and products gender and racialize one's identity. So what "makes foreign" or "makes citizen" a data subject?

While unapparent at first glance, this process can be inferred a small bit in the leaked document "FAA Targeting Procedures A"[7]. Here we get a long list of factors producing a data subject's "foreignness determination", including: communication with an individual "reasonably believed to be associated with a foreign power or foreign territory"; presence on the "'buddy list' or address book" of someone "reasonably believed" to be "associated with a foreign power or foreign territory"; metadata records "acquired by lawful means" that reveals an individual to be "associated with a foreign power or foreign territory"; and even "IP ranges and/or specific electronic identifiers or signatures (e.g., specific types of cryptology or steganography)" that are used by foreign people. More incredulously, in other documentation leaked to Brazilian paper Jornal O Globo, language (like everything but English) used in emails can also be a variable that produces foreignness.

Results and Discussion

There are three key key components that arise from understanding citizenship in this way. One, we are always both citizen and foreigner. A target who is 51% foreigner is neither 100% foreigner nor 0% citizen. We are not either one or the other, but always concurrently both at the same time.

This differs radically from the *soli* and *sanguinis* variants of citizenship allocation because it is not binary. But, much unlike the theoretical citizenship standards of birth or blood, a non-binary description of citizenship might be more apropos. Think about the Arizona law SB1070 in 2010 and the ways that citizenship in the US is heavily racialized as white. A Latino US citizen might not be 100% citizen and 0% foreigner, but something in between.

Two, the valuations between these two are dynamic. As data subjects we constantly produce new information pieces, new inputs (new IP addresses and new social interactions), that are interpreted and valued as determinations for foreignness. Even language, search terms, and web sites visited have the portent to be influential in the output of our foreigner/citizen interpretations. One day the NSA might label me as 49% confidence foreigner; then I take my computer on a trip to Mexico, make a couple friends from Italy, add them onto my Gchat, and return to the US to find that I'm now more foreigner than citizen.

Three, with *jus algoritmi* the index of citizenship moves from a legal, static, enshrined category with historical precedent and juridical standing to a statistical assessment of commonality patterns. At first glance, with no hard-coded center to foreignness or citizen, the "right to have rights" sloshes recklessly across these 51% percentage measure thresholds. But at second glance, reconsider the NSA rubric for evaluating foreignness. Me, a citizen read as foreigner, becomes foreigner. Then, a friend of mine, a citizen read as citizen, talks to me. Eventually, that friend is read as foreigner because I am read as foreigner. It's an incredibly slippery slope precisely because it is impossible to find an anchor holding onto the verifiable truth of one's legal identity.

Conclusion

I am interested in the reframing of the concept of citizenship, a concept that we must think about in both offline, but now especially online, ways. For many of us who are US citizens, our mundane, offline relationships to citizenship is often assumed and unquestioned – it is passive, only activated when we return to the US from a trip overseas.

So what happens when that concept is shaken, removed from its givenness and now put into play? State and corporate surveillance is likely to be one of the most common encounters we have with the rights guaranteed by our US citizenship. Decisions on whether the NSA can spy on us or not is something that is being assessed every time we move, physically — our locations are being tracked by our cell phones, even when they are turned off — log onto our computers, or accept a contact request on GChat. Surveillance like this is surely a reduction of privacy. But it also is a formative redefinition of how citizenship itself can be understood in the future.

References and Notes

1. Gellman, Barton; Poitras, Laura. U.S., British Intelligence Mining Data from Nine U.S. Internet Companies in Broad Secret Program. *The Washington Post* 2013, http://www.washingtonpost.com/investigations/us-intelligence-mining-data-from-nine-us-internet-companies-in-broad-secret-program/2013/06/06/3a0c0da8-cebf-11e2-8845-d970ccb04497_story.html.
2. National Security Agency. 2013 Press Release - Press Statement on 30 July 2013. *NSA*, 2013, https://www.nsa.gov/public_info/press_room/2013/30_july_2013.shtml.
3. Gellman, Barton; Poitras, Laura. U.S., British Intelligence Mining Data from Nine U.S. Internet Companies in Broad Secret Program. *The Washington Post* 2013, http://www.washingtonpost.com/investigations/us-intelligence-mining-data-from-nine-us-internet-companies-in-broad-secret-program/2013/06/06/3a0c0da8-cebf-11e2-8845-d970ccb04497_story.html.
4. National Security Agency. SIGINT Strategy 2012-2016. *New York Times* 2013*,* http://www.nytimes.com/interactive/2013/11/23/us/politics/23nsa-sigint-strategy-document.html.
5. Ibid
6. Arnbak, Axel. NSA Strategy 2012-16: Outsourcing Compliance to Algorithms, and What to Do About It. *Freedom to Tinker* 2013, https://freedom-to-tinker.com/blog/axel/nsa-strategy-2012-16-outsourcing-compliance-to-algorithms-and-what-to-do-about-it.
7. National Security Agency. FAA Targeting Procedures. *ACLU* 2013, https://www.aclu.org/faa-targeting-procedures?redirect=national-security/faa-targeting-procedures.