



eCitizen.TV

Legal and Policy Research Blog of the **MIT Media Lab Human Dynamics Group**

[eCitizenConnect](#) [Personal Media Network](#)

[Blog](#)

[MIT Research](#)

[About](#)

SUNDAY, JANUARY 12, 2014

Big Data and Identity

The trend toward wide adoption of big data and data driven services and systems is showing early signs it will bring profound implications not only for how online identity works and in some sense, what *identity actually is*. The problems and prospects arising as a result of this trend are significant. This is a good moment to start describing, discussing - and deliberating about - how the big data age is expected and intended to unfold.

Examples

Below are some very simple illustrations of how identity is expressed and operates with some of the Google+ re-posts and comments that have been coming in related to my previous post on CivicID (yes, it's a very **meta** approach to selection of examples!):

Andy Oram
 Shared publicly · Jan 10, 2014
 2 circles

Multiple identities in the public sphere through **#CivicID** (links to MIT videos) <http://bit.ly/1iqoCds> (via **+Dazza Greenwood @Nstic**)

Relationship of Core Identity Infrastructure and CivicID
[bit.ly](http://bit.ly/1iqoCds)

To demonstrate how comment threads evolve, here is an example from another post I made yesterday discussing a crowd-sourced wiki site to draft legislation in a new open, participatory model:

ABOUT ME



+ Dazza Greenwood



Follow

359

I'm the Dazza Greenwood who is

www.DazzaGreenwood.com

[View my complete profile](#)

TOPICS

[eCitizenConnect](#)

[Identity](#)

[LegalScience](#)

[PersonalData](#)

SUBSCRIBE TODAY



Posts



Comments



BLOG ARCHIVE

▼ 2014 (24)

► April (3)

► March (5)

► February (8)

▼ January (8)

[ClearButton: A Personal Data Notice Registry](#)

[Computational Legal Science](#)

[The State of the Network](#)

[Fresh Light on the Personal Data Social Compact](#)

[Big Data and Identity](#)

[Wikifying the Law](#)

[Relationship of Core Identity Infrastructure and C...](#)

[MIT Meda Lab & Janrain Demo of CivicID](#)

► 2013 (10)

GOOGLE+ FOLLOWERS

Google+

Matthew Baggott
Shared publicly · Jan 9, 2014

Dazza Greenwood originally shared:
Wikifying the Law
I've been participating in a new project of a state assemblyman
[Read more](#)

Wikifying the Law
ecitizen.tv

+3 3

Notice how my identity as the blog author is leveraged via the Google Plus account and how the identity of the people who comment or re-post are logically grouped together by the common thread of communication and uniquely linked by reference to the URL of this page... the makings of big data, automated models of user-behavior and metadata based attributes create extended layers of associations with other people or groups, and ties to ideas or opinions, etc. The expectations that do or should apply to these facets of identity are unclear and raise a wide range of public policy, legal and basic social questions.

What Rules Do (or Should) Apply?

A quick look at the most apparently relevant Google user agreement (the Google "**Embedded Content Policy**") which raises as many questions as it provides answers perhaps partly due to the newness of these types of services and the emergent nature of how these and other social network type services will interact and evolve. One clause of the embed policy states:

"Do not use, sell, or transmit to others any data obtained from the embedded Google+ content, including data related to the user's interaction with embedded Google+ content. For the avoidance of doubt, this prohibition includes, but is not limited to, any use of pixels, cookies, or other methods to recognize users' clicks on embedded Google+ content, the data of which is then disclosed, sold, or otherwise shared with other parties."

One wonders if the act of embedding the content on a blog such as this Google Blogger powered site is the "use" and "transmission" to others (of) data obtained from the embedded Google+ content" because the embed process results in "including data related to the user's interaction with embedded Google+ content" and therefore is apparently inconsistent with the words. Yet, it would seem the act of using Google embed features is designed and intended to facilitate behavior like including the embedded content on a Google powered social blog post.

From Social Media to Social Compact

Evidently, the relevant boundaries and broadly understood sets of responsibilities, rights and relationships will settle remains to be determined. While Google+ comment make nice examples, the ownership, control and responsibilities governing social network comment systems may not seem to be especially high stake. But the big data aspects of identity are starting to color a range of higher profile and deeper value aspects of life as well.

The FTC is starting to pay closer attention to so-called big data companies and their practices with respect to privacy and consumer protection. Upcoming seminars hosted by the FTC this spring will shine a light on new trends in Big Data and their impact on consumer privacy. Consider the newly formed industry initiative called "**About the Data**" which is intended to provide some access to the types of personal data collected by participating companies. This data drives automated systems that can impact or drive anything from selection of the specific advertisement you might see on a screen based on your overall and current behaviors at a particular moment to bigger impact contexts such as how insurance, employment or financial decisions will be made about you. This industry effort signals that existing companies depending on access to a flow of personal and

Dazza Greenwood
Add to circles

359 have me in circles [View all](#)

GOOGLE+ BADGE

Dazza Greenwood
google.com/+DazzaGreenwood

Lately, my identity is DazzaGreenwood.com (names, identifiers and contexts all rolled ...)

[Follow](#)

359 followers

identity data recognize the need to address the issues and perhaps seek to frame some of the public dialog of the imminent future. But what would a complete set of relevant issues look like and how can a coherent framework for dialog and debate be structured to ensure the viewpoints of individuals, industry, policy makers and others are adequately engaged and balanced?

The types of positioning and framing of issues both by regulators and industry signal early movement to address a large scale change carrying with it a big set of issues. The relevant topics, when looked at carefully, implicate fundamental underpinnings of a broader political and economic expectations about ownership, property rights, civil liberties and basic expectations about the role and relationships people can expect in the emerging era of big data societies and the digital economies. And - as expected at times of major transition - identity is one of the topics in the middle of the issue map.

Decision Point: Time for a New Deal on Data?

Today, we are witnessing the transformation of identity itself into a constituent dimension or component of big data of loosely connected automated systems. Everyday interactions and behaviors are increasingly occurring by and with computational models that constitute core aspects of individual identity. These statistical methods and expressions result in the way people appear and in a sense, who they are in the reality of a networked world. The model-based identities operate within liquid big data environments that flow as dynamically as white water river. Mathematic, probabilistic algorithms are literally how identity works in these contexts. And in this way, big data identities are emerging as new types of entities that largely exist outside of the direct understanding of the people they identify, describe and perhaps define.

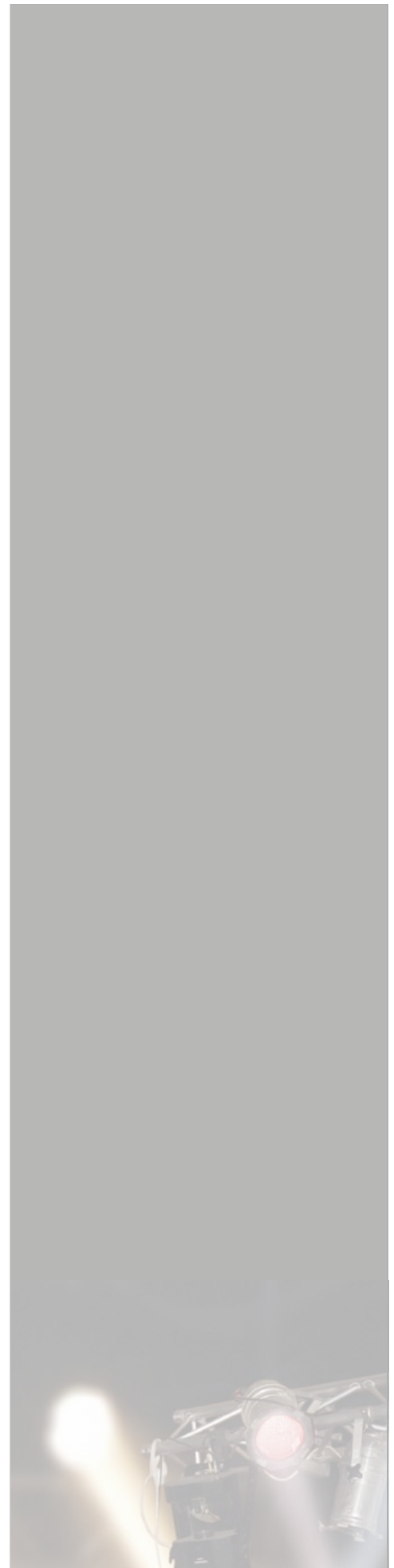
The trend toward greater transparency and individual knowledge, consent and control with respect to their personal data will be the most relevant key policy and business issue in this context. Identity is about as personal as personal data gets. A basic question going forward is: In an age of big data, how do people know who they are, and what aspects of their identity belong to them? The New Deal on Data, advocated by MIT Professor [Sandy Pentland](#) hold the essential elements of a workable solution. I am privileged to be part of Sandy's research team and have had the opportunity to learn about and gain experience applying the New Deal on Data concepts in business, legal and technical prototypes and other projects.

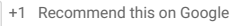
I'm optimistic that the New Deal on Data approach can be refined, adapted and adopted to form the basis of a broader renewed agreement - a kind of re-stated social compact - upon which economic, political, cultural and other expectations can be safely premised. A forum like the NSTIC multi-stakeholder private sector led "[Identity Ecosystem Steering Group](#)" and the [World Economic Forum](#), among others, are great examples of the types of places where the issues and options can be hashed out from several key vantage points. Eventually, business conventions, legal frameworks and technical standards will need to reflect and implement a broader general agreement on ownership, control and the new institutions of a big data age.

How identity is treated within this larger economic and public debate will be the pivot upon which we launch into the coming era. As my beacon light of high level perspective and foundational good sense [Dan Geer](#) noted at the [MIT KIT launch](#) last year, this is probably the last generation with a real say in what the fundamental premises of the renewed or replaced social compact will be. The period of decision is upon us. The big data age is happening and it has implications for the basic deal on which we rest expectations about who we are and what we can expect from our lot in life. Pivot one way, and we can successfully reinterpret but maintain the basic values and principles in the next era. Pivot another way, and the opportunities for profoundly different types of society and social structures may result. Hopes for utopian and risks of dystopian futures about, as do opportunities to further realize the historic dreams of freedom, prosperity and peace. At the individual, organizational and societal levels, engaging to deliberate and decide a direction is a choice, and - if only be default - going with the flow is also a choice.

As [Dan Geer admonished](#), let us choose wisely.


Posted by [Dazza Greenwood](#) at 10:40 AM







1 comment


Google+



Add a comment as Dazza Greenwood

Top comments



Dazza Greenwood

1 year ago · Shared publicly

This is Your Identity... On Big Data

The trend toward wide adoption of big data and data driven services and systems is showing early signs it will bring profound implications not only for how online identity works and in some sense, what identity actually is . Examples Below are some illustra...

+1

+1 · Reply

Older Post

EMAIL UPDATES

Submit

