ELSI of big data in health care

SNSF Professor Effy Vayena
Health Ethics and Policy Lab
Epidemiology, Biostatistics and Prevention Institute
University of Zurich

Volume, Velocity, Variety, Veracity

- **Ubiquitous** (from dating, to hiring, to voting, to policing, to searching for terrorists, etc.)
- **Everyone** is involved and affected (not necessarily by choice)
- **Ever-growing**, ever-expanding (metrics, IP addresses, space)
- **Hyped** (means different things to different stakeholders)

Big DATA in Medicine & Health Care

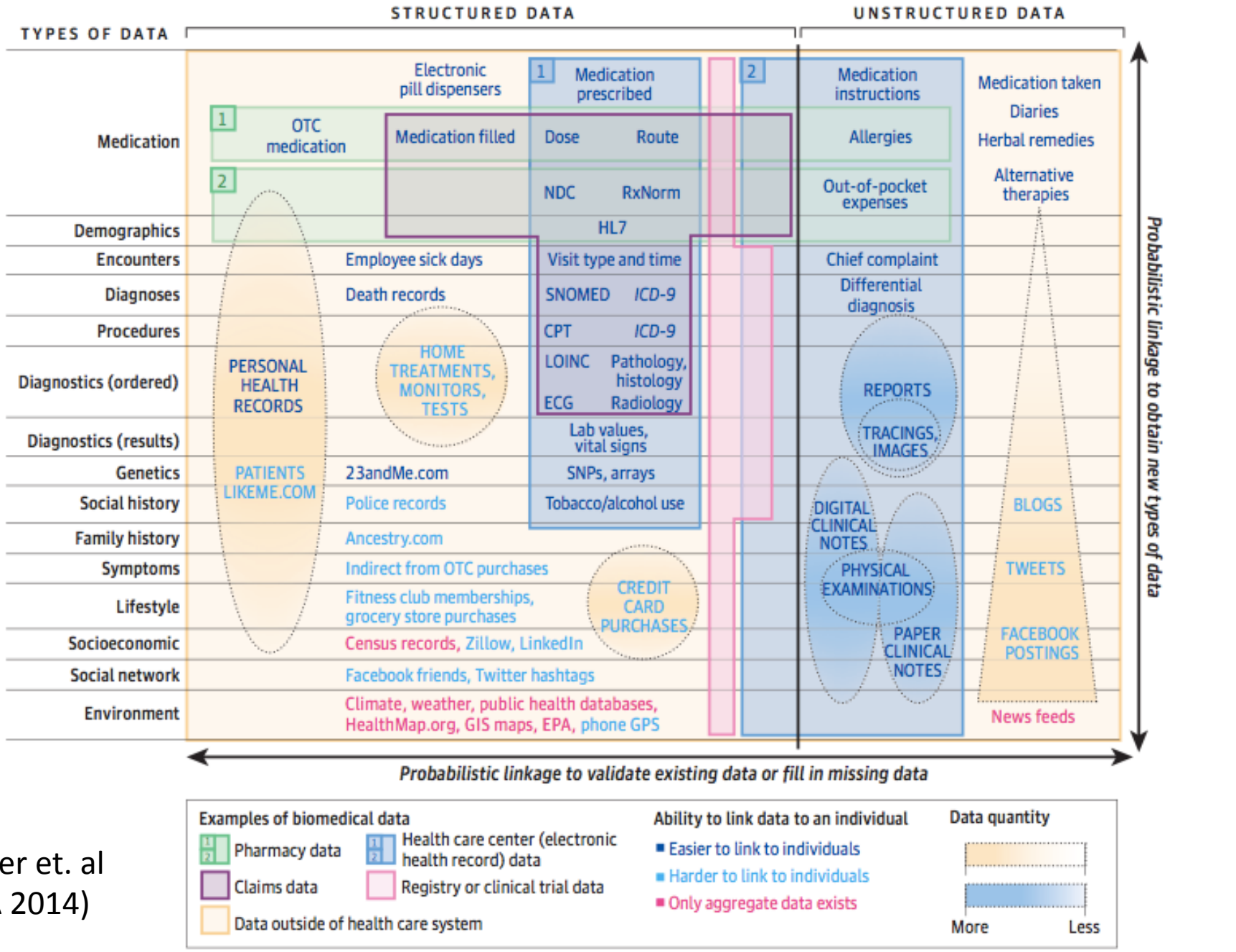
- EHR
- mhealth
- Body-adapted wearable electronics
- Personal genomics
- Quantified Self

- Loyalty cards/purchases
- Phone data
- Web searches
- ….

"The patient is an enormous repository of information that needs to be harvested as a partnership not only in clinical care but in discovery as well. It is the only way we will **define wellness and its progression to disease,** rather than **traditional medicine that defines disease and its progression to death**. "

Dennis Ausiello , Nature Biotech, 2015

Figure. The Tapestry of Potentially High-Value Information Sources That May be Linked to an Individual for Use in Health Care



(Weber et. al JAMA 2014)

Figure 1 Timeline of insomnia-related tweets from representative individuals. Density distributions (probability density functions) are shown for seven individual users over a two-year period. Density on the *y* axis highlights periods of relative activity for each user. A representative tweet from each user is shown as an example.

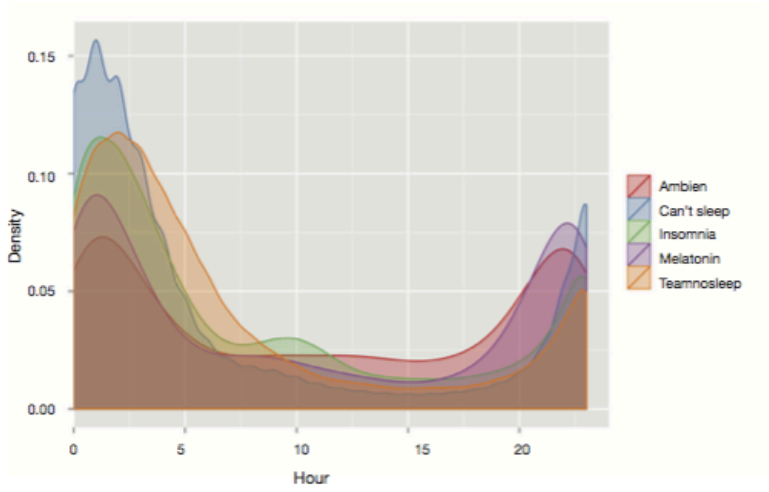


Figure 2 Time distribution of insomnia-related tweets. Density curves (probability density functions) for multiple insomnia terms are shown, illustrating what time of day (using a 24-hour clock) tweets were posted. Density on the *y* axis highlights times of relative activity for each keyword. Tweets were from our own database of geo-tagged tweets, with 1,315,236 tweets shown here across a two-year period. Note that the time of each tweet was converted into the user's local time.

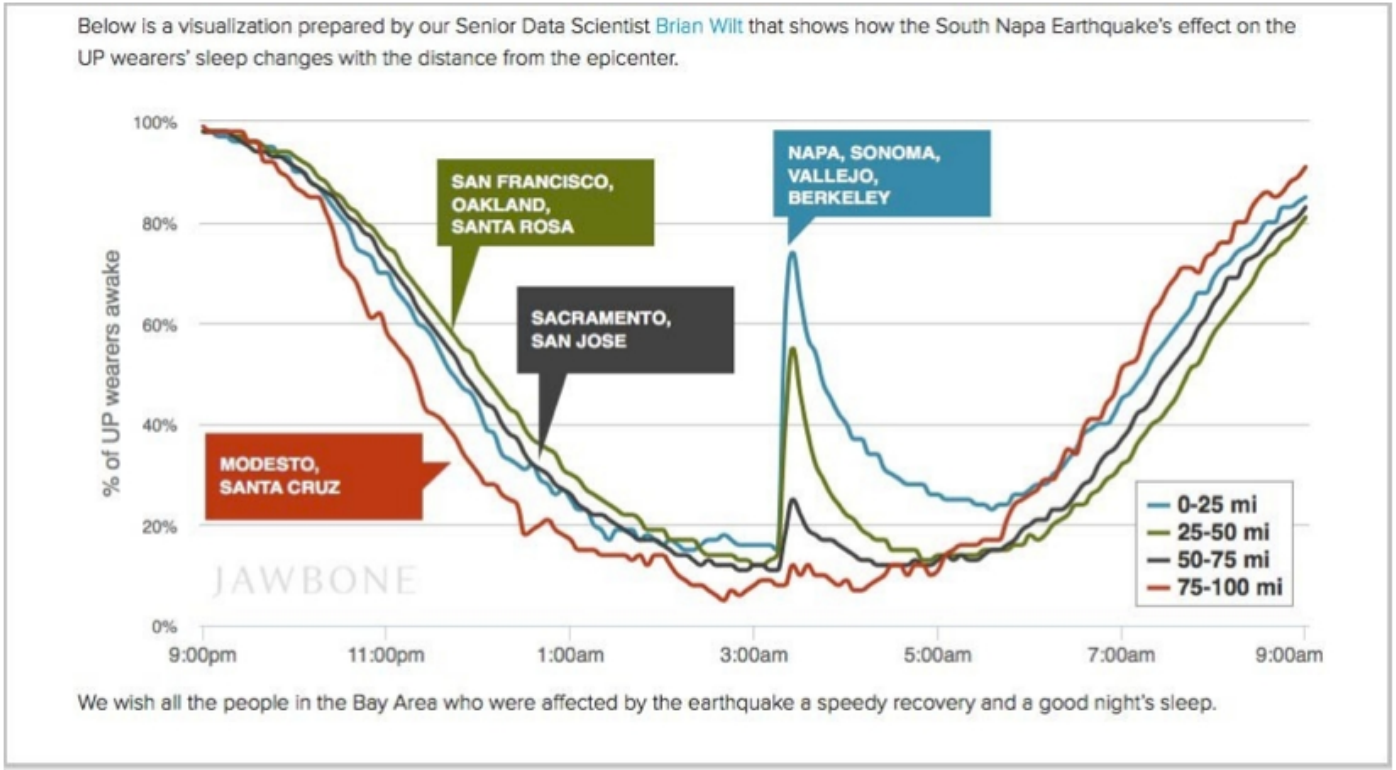
The digital phenotype

Sachin H Jain, Brian W Powers, Jared B Hawkins & John S Brownstein

In the coming years, patient phenotypes captured to enhance health and wellness will extend to human interactions with digital technology.

Nature Biotechnology, 33:5 2015

> *I wear a Jawbone Up fitness tracker, and I was surprised when I saw its blog posts that show research about sleep, like this one on circadian rhythms and another about the earthquake in Napa. I was one of those 55 percent people in Oakland who woke up during the quake! Did I sign up to be in a sleep study when I bought my Jawbone Up?*
>
> *— Brandon, Oakland, California*



Awake in Oakland after the Aug. 24 earthquake in Northern California.

http://america.aljazeera.com/articles/2014/10/29/sleep-study.html

1. Privacy

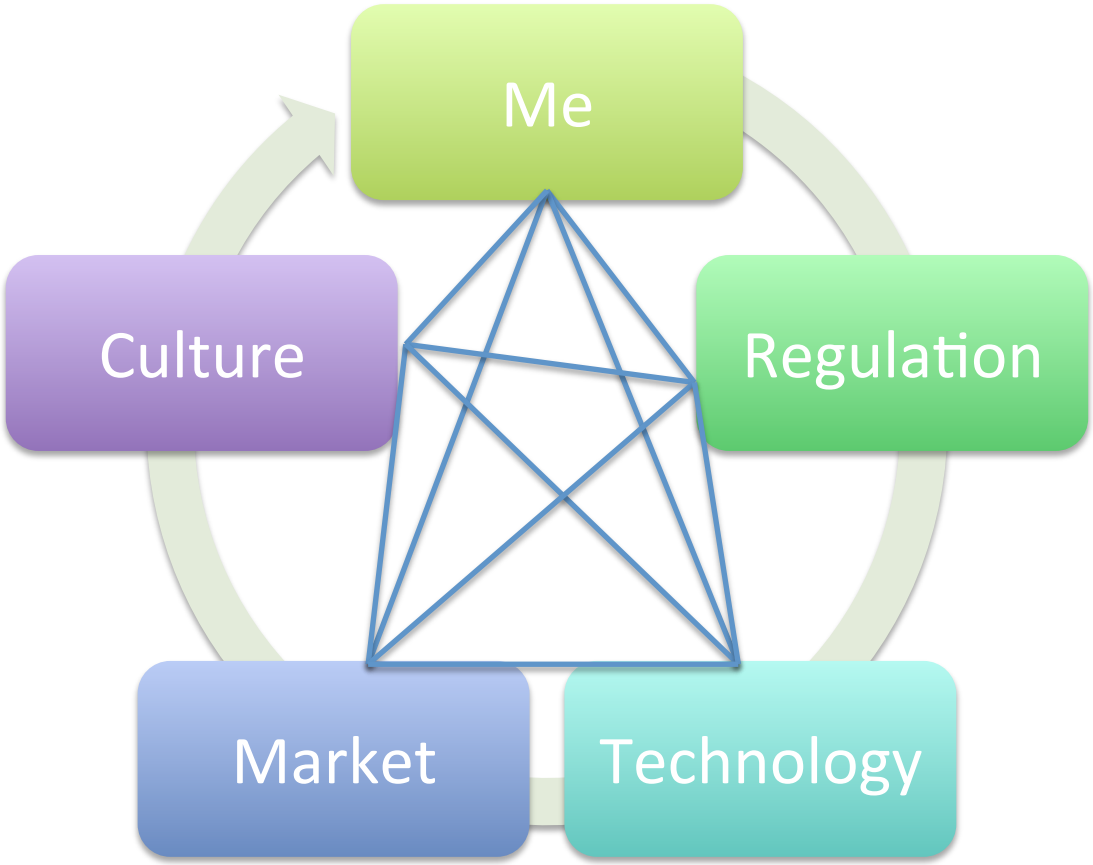
“There’s privacy issues. We've got to figure out how do we make sure that if I donate my data to this big pool that it's not going to be **misused**, that it's not going to be **commercialized** in some way that **I don't know about**.

And so we've got to set up a series of structures that make me confident that if I'm making that contribution to science that I'm not going to end up getting a bunch of spam targeting people who have a particular disease I may have.”

President Obama, **Precision Medicine Initiative**

February 2016

. Privacy



Collection

Storing

Processing

Sharing

Dissemination

.Privacy

- "…big biomedical data are scattered across institutions and intentionally isolated to protect patient privacy." (Weber et. al JAMA 2014)

- World Economic Forum: obstacles to data commons- privacy and security.

. Public and private

- "Analog" system of protection
  - Opting out
  - Anonymization
  - Reliance on notice and consent

2. Consent

- Necessary but not sufficient
  - Unknown uses, immense possibilities
  - Digital trail (authorization, not consent)

  - Consent innovation
    - General
    - Granular
    - Dynamic
    - Portable

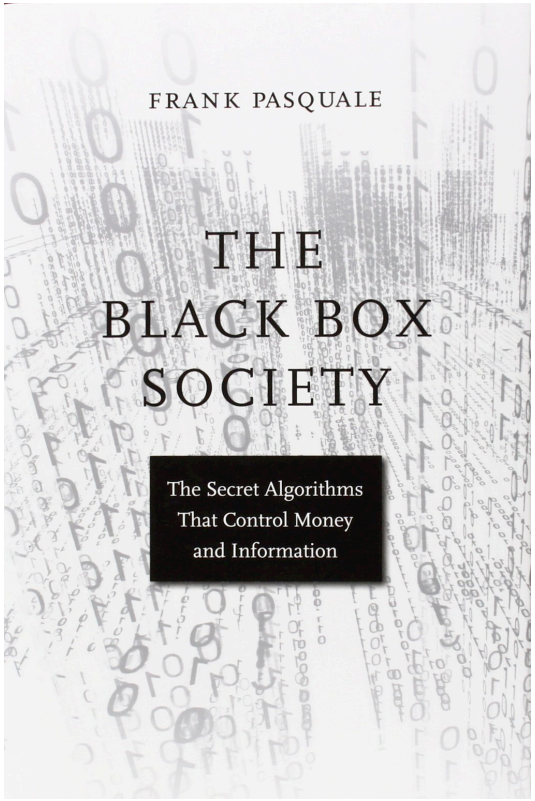
3. Data uses

- Focus on ethical **use**
  - Ethics review
  - How to review big data proposals?
  - Just standard review?
  - Risk-benefit assessment
  - Benefit sharing (IP, patents)

4. Governance of data initiatives

Principles of governance

- Transparency
- **Accountability**
- Participation
- Deliberation
- Fairness



5. Public involvement

- Digital literacy
- Empowerment (personal data space/vaults/stores)
- Opportunities to act
- Opportunities to govern

high quality care for all,
future generations

NHS England

Our work | News | Events | Publications | Resources | Statistics | Contact us

Home › Our work › Technology, systems and data › The care.data programme – collecting information for the health of the nation

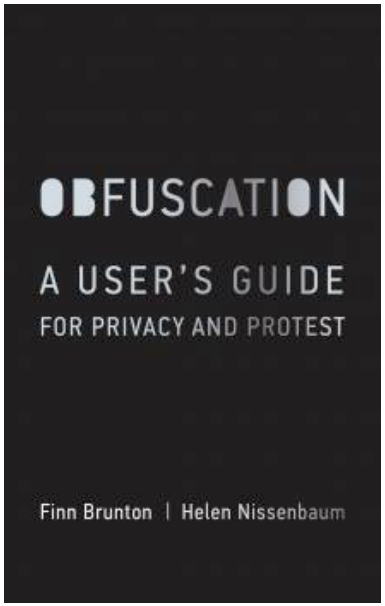
search the site

Advanced search

NHS choices Visit NHS Choices for patient information

The care.data programme – collecting information for the health of the nation

What is care.data?



Expanding data ecosystem in health

Shifts:

① Data taxonomies (biomedical data?)

② New methods (black box)

③ Public and private distinction

④ New actors (data companies)

⑤ New roles (CEO-statesman)

Vayena E.&Gasser U. "Strictly biomedical? Sketching the ethics of the Big Data ecosystem in biomedicine
in B. Mittelstedt & L. Floridi (eds.) The Ethics of Big Biomedical Data (Springer 2016)

Big data regulatory ecosystem

European Data Protection Supervisor, Opinion, September 2015