



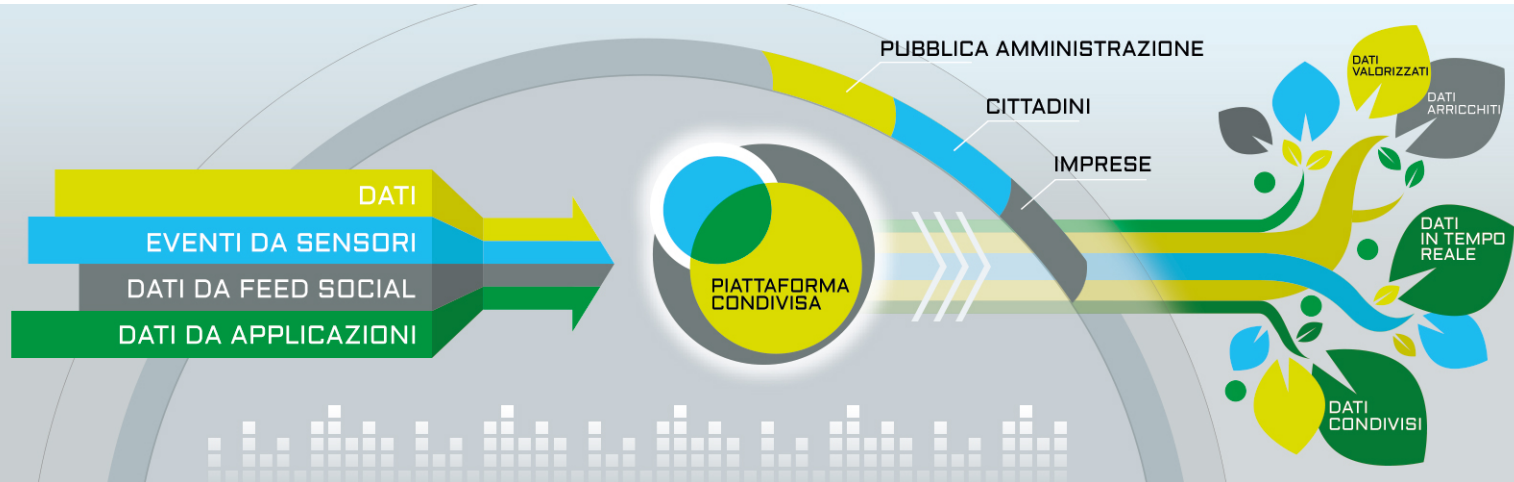
Nexa Center for Internet & Society

Politecnico di Torino

Studying the Internet, exploring its potential & experimenting new ideas

**Open data e aspetti economici
dei dati personali**

<http://www.smartdatanet.it/>

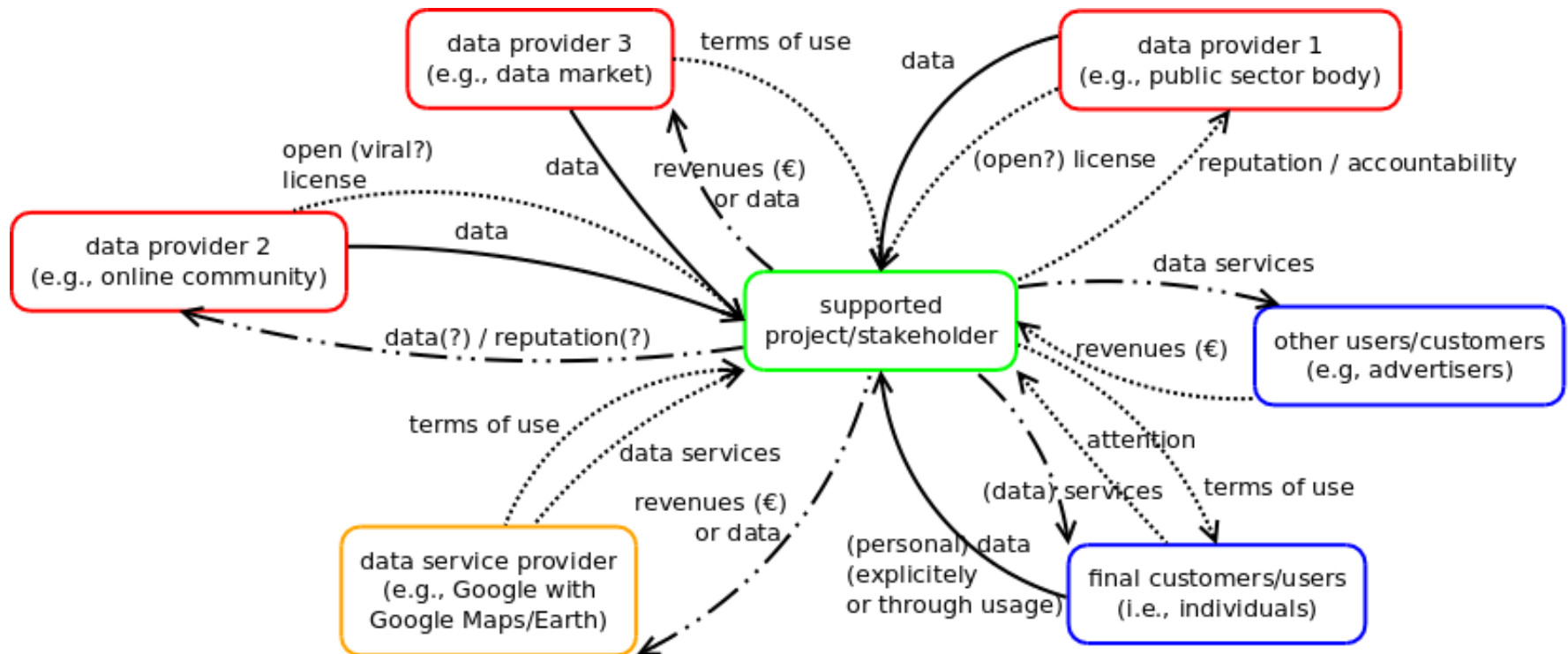


Fonte: CSI Piemonte

Bilanciare diritti e opportunità

- European Partnership for Big Data Value
- “Chilling effects”
 - Incertezza giuridica (IPRs, privacy, aspetti contrattuali)
- Smart disclosure e data portability come elementi abilitanti
- Idea di “data value networks”

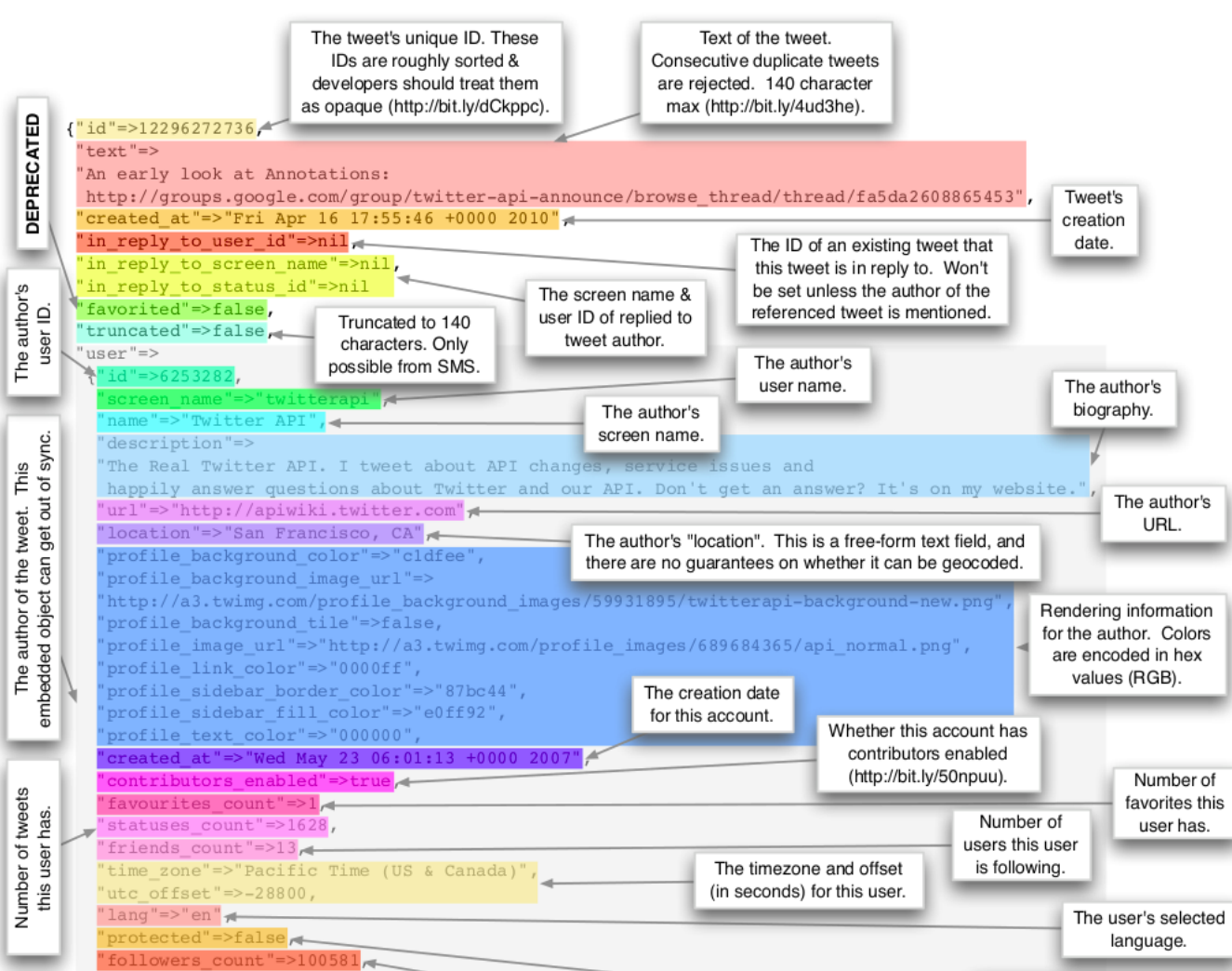
Archetipo di “data value network”



Big data (?)

- “Why big data evangelists need to be reprogrammed”
- "Too often, big data enables the practice of apophenia: Seeing patterns where none actually exist, simply because enormous quantities of data can offer connections that radiate in all directions. In one notable example, Leinweber (2007) demonstrated that data mining techniques could show a strong but spurious correlation between the changes in the S&P 500 stock index and butter production in Bangladesh”

Dati e metadati




Inferenze indesiderate (?)

- **How browsers affect your job prospects**
 - “(...) those applicants who have bothered to install new web browsers on their computers (such as Mozilla's Firefox or Google's Chrome) perform better and stay in their posts for 15% longer, on average”.
- **Networked employment discrimination**
 - algorithmic assessments of workforce data to identify the variables that lead to stronger employee performance;
 - automated systems, algorithms, and existing networks systematically benefits some at the expense of others, often without employers even recognizing the biases of such mechanisms.

The dark market for personal data

**ATTENTION DATA BROKERS
PLEASE BE ADVISED OF THIS
UNIQUE CASH OPPORTUNITY**

**MY NAME FOR SALE:
“Sam Potts”**



**100%
UNVETTED
CONSUMER
DATA**

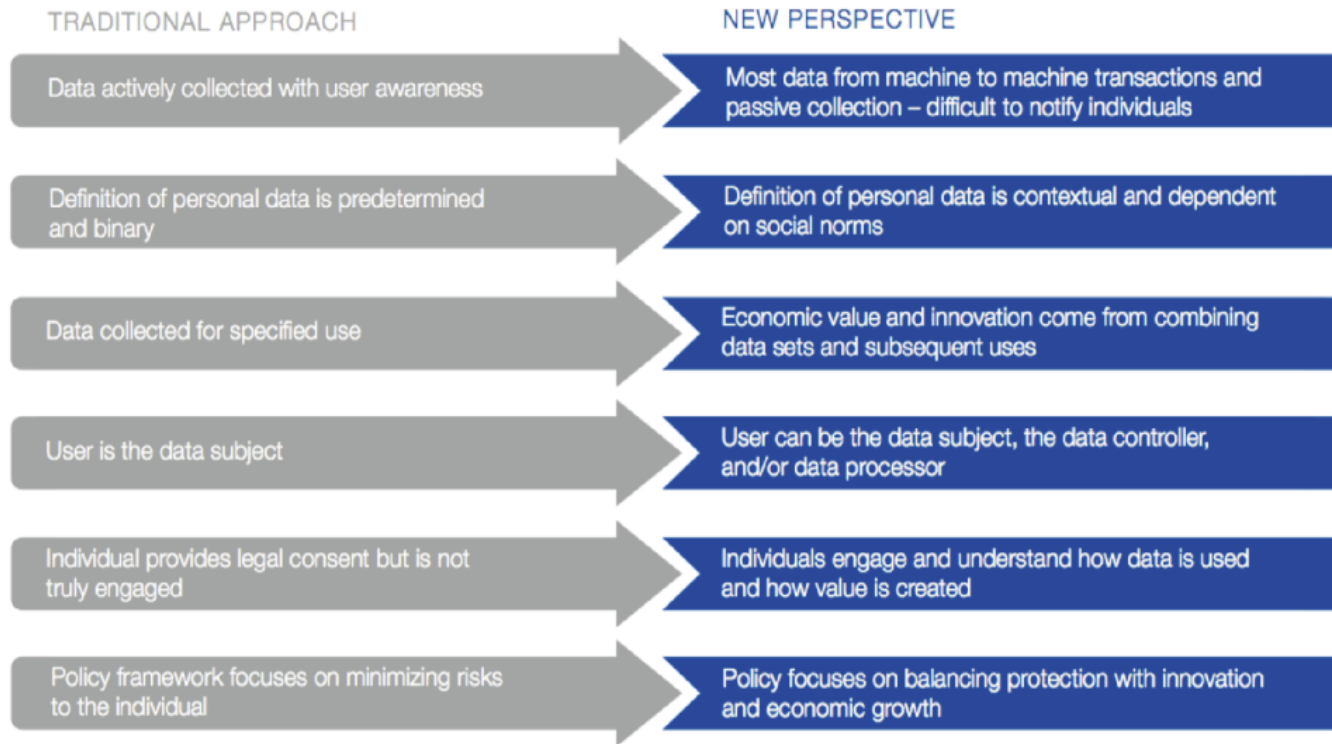
**ALL
INFORMATION
MORE OR LESS
ACCURATE**

**FOR INCLUSION IN ANY OF YOUR LISTS
— SUCH AS —**

NEARSIGHTED	\$ 0.10
CASUAL SMOKER FOR 10+ YEARS	\$ 0.12
BALDING (SLOWLY)	\$ 0.19
HAD PNEUMONIA ONCE	\$ 0.23
SPLINTER STILL STUCK IN MY HAND FROM 1978	\$ 0.35
BUNIONS	\$ 0.46
AGING SENIOR (FEELS LIKE ONE, AT LEAST)	\$ 0.52
TERRIBLE FINANCIAL PLANNER	\$ 0.58
IMPULSE BUYER (WHY NOT?)	\$ 0.64
NEVER TOOK CALCULUS	\$ 0.73

NY TIMES di oggi:
« The reputation business
is exploding »

the new role of data



Source: World Economic Forum and The Boston Consulting Group

<http://www.weforum.org/issues/rethinking-personal-data>

Un “habeas corpus” digitale?

- Right to see what data any register stores and request the rectification, actualization or even the destruction of the personal data held.
- Right to appeal against any violation of the above.
- Right to reconcile and reuse personal information.

Privacy vs open data

(“Privacy by design”)

- In genere, i dati personali possono essere pubblicati se:
 - esiste un consenso informato (revocabile); OR
 - c'è un'obbligazione giuridica; OR
 - i dati sono stati anonimizzati.
 - Esempio: database incidenti
- Dati sensibili
- Impact assessment
- Distinzione sempre più lasca tra chi raccoglie i dati e chi li gestisce

Dati e tracce

- **Personal data**

“any information relating to an identified or identifiable individual (data subject)” (OECD)

- **Digital footprint**

All data left behind by users within their use of digital services, and the personally unique arrangement that makes someone identifiable just upon the specific combination of her system information.

- **“Internet of Things”**

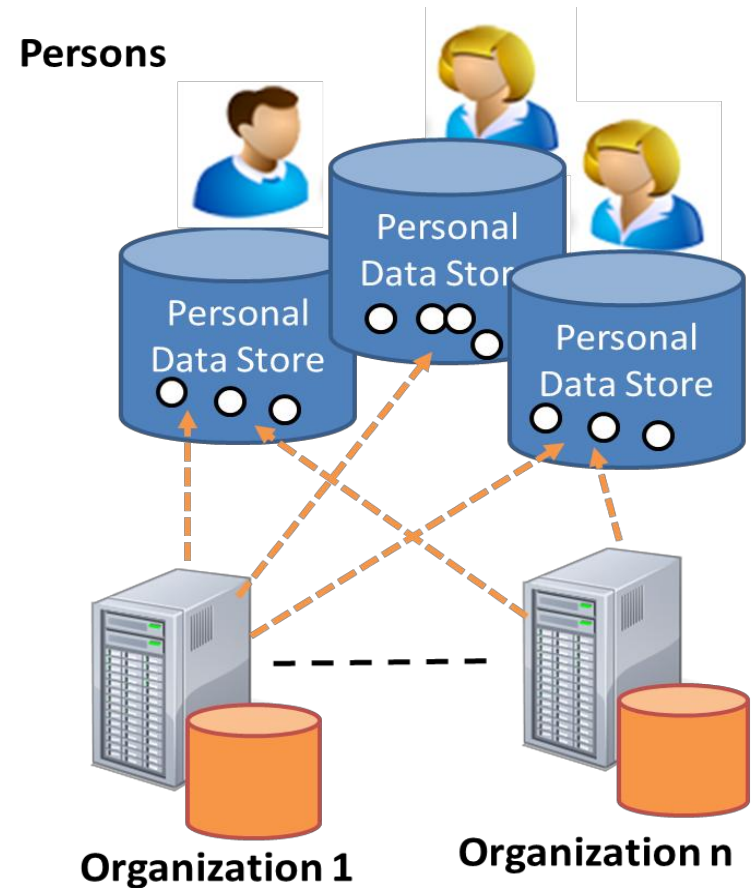
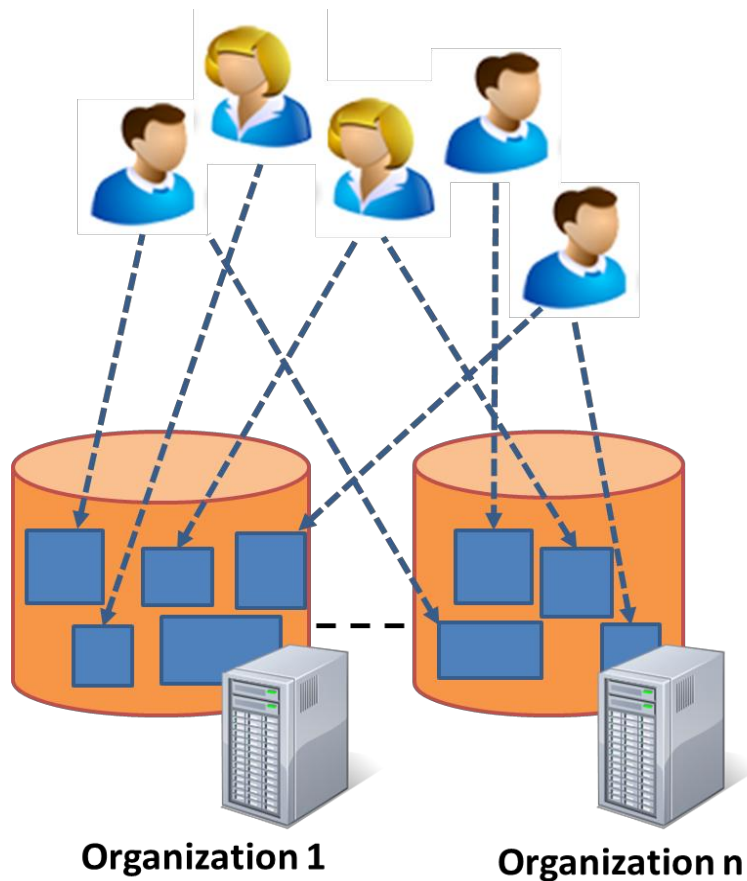
Digital footprint: tipologie (OECD)

- Data can be **voluntarily shared** by a consumer.
- Data can be **observed or recorded**, with or without consumers' knowledge, or explicit consent.
- Data that can be **inferred**.
- In other cases (Acquisti, 2010), personal data remains protected, either because it was intentionally not disclosed by a consumer, or because the service provider is not able to access it.

Personal data vaults

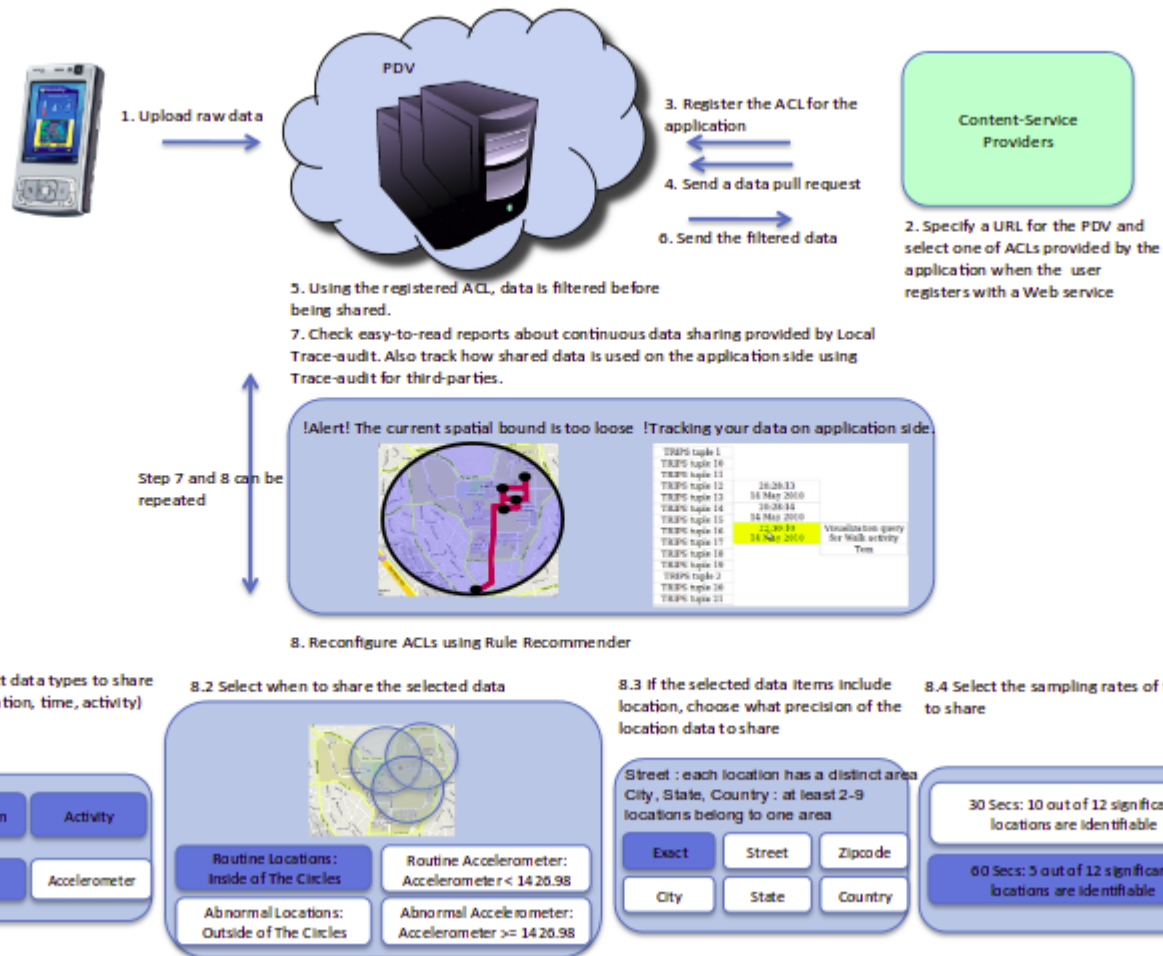
- Primato degli individui sul controllo dei (propri) dati.
- Alto livello di accessibilità e leggibilità dei dati.
 - E condivisione selettiva con terzi.
- Relazioni / accordi di lungo termine.
- E.g., **Personal**

Approccio centralizzato vs approccio “utente-centrico”



Fonte: Moiso, C., & Minerva, R. (2012, October). Towards a user-centric personal data ecosystem: The role of the bank of individuals' data. In Intelligence in Next Generation Networks (ICIN), 2012 16th International Conference on (pp. 202-209). IEEE.

Possibili casi d'uso



Fonte: Mun, M., Hao, S., Mishra, N., Shilton, K., Burke, J., Estrin, D., ... & Govindan, R. (2010, November). Personal data vaults: a locus of control for personal data streams. In Proceedings of the 6th International Conference (p. 17). ACM.

Possibili casi d'uso

- Automatic and controlled personal information exchange
 - e.g., hotel check-in
- Vendor Relationship Management
- Fair personal data marketplace
- “Personal” application marketplace
 - come Green Button

Fonte: Moiso, C., & Minerva, R. (2012, October). Towards a user-centric personal data ecosystem The role of the bank of individuals' data. In Intelligence in Next Generation Networks (ICIN), 2012 16th International Conference on (pp. 202-209). IEEE.

Il valore dei dati personali (dal punto di vista degli individui)

- Organisations are increasingly in the position of holding information about individuals.
- A single consumer can experience **both benefits and costs** from the disclosure of personal information (Hann, 2007).
- An organisation gains clear advantages from being able to increase its knowledge about consumers' identity and behaviour. Still, it may deem detrimental for its own business to design **too invasive policies** (Acquisti, 2010).
- What seems to matter is that personal data (regardless of its amount) is used in an **appropriate context** (Nissebaum, 2004)

Un possibile paradosso

- Spiekermann et al. (2001) suggest that **even privacy conscious individuals are likely to share sensitive information** with strangers, in particular online.
- Acquisti and Grossklags (2005): almost 90% of individuals declared to be concerned about privacy, but more than 20% of the people in the same sample admitted to have disclosed their **social security number** for discounts or better services.
- Beresford et al. (2012): in a context of equal prices, the choice between two different firms to purchase a DVD seemed to give a **virtually zero premium to the more privacy-friendly vendor**. Nevertheless, 75% of the participants indicated that they have a very strong interest in data protection.
- *Declared vs. revealed* preferences.

Informazione contestuale

- “[i]ndividuals make privacy-sensitive decisions based on multiple factors, including (but not limited to) what they know, how much they care, and how costly and effective they believe their actions can be.” (Acquisti and Grossklags, 2005).
- Esperimento.

Endowment effect

- Willingness to pay (**WTP**) vs Willingness to accept (money) (**WTA**).
- More than half of subjects endowed with a completely anonymous \$10 gift card rejected an offer of \$2 to reveal their future purchase data.
- By contrast, fewer than 10% of subjects endowed with the identified \$12 card chose to give up \$2 to protect future purchase data.
- These results imply that consumers endowed with different levels of privacy protection valued the same good, i.e. future privacy protection, in a substantially different way.
- Most of the consumers whose privacy was fully protected did not consider \$2 as a price high enough to sell their future purchasing data, while nine out of ten of the not protected individuals estimated \$2 as a too high price to buy protection for future purchase data.

Il ruolo dei disclaimer

- Tsai et al. (2010), different sets of participants were asked to test a new search engine whose results were annotated with icons.
- 2 products: low privacy concern (batteries), high privacy concern (sex toy), using their credit card.
- Some participants exposed to icons describing the site privacy policy.
 - Others not, as a control condition.
- Participants in the privacy information condition were more likely to make purchases from websites offering medium or high levels of privacy (even at higher prices).
- Quite surprisingly, the premium to privacy was similar in purchasing the batteries and the sex toys.

Eterogeneità degli utenti

- Hann et al. (2007) find three different clusters in terms of privacy evaluation
 - ‘privacy guardians’, i.e. people who attach a relatively high value to information privacy [70 – 85%];
 - ‘information sellers’, who tend to give away personal information with little regard for privacy policies;
 - ‘convenience seekers’, people who prefer convenience with little regard for money or website privacy policies.

Privacy come bene di scambio (1/2)

- The existence of a privacy statement makes people more available to share data, even for free (Hui et al., 2007; Tsai et al., 2010; Hann et al., 2007; Tucker, 2012). Joinson et al. (2010) find the existence of a **compensatory relationship between privacy and trust**.
- Several studies (Hann et al., 2002; Hann et al., 2007; Tucker, 2012) show that individuals are willing to trade personal information in order to gain **convenience in terms of time saving and personalisation** of web services
 - Bear in mind that persons differ in terms of opportunity costs.

Privacy come bene di scambio (2/2)

- **Social rewards** are attractive in balancing privacy concerns and governing individuals' behaviour as well (Jiang et al., 2013).
- Hann et al. (2007) find that a sufficiently large **monetary reward** significantly increased the relative attractiveness of a website regardless of its privacy policy (WTA).
- WTP
 - Krasnova et al. (2009) observe that a user would be ready to pay between 14 and 17 Euros/year, if the social network refrained from using her demographic information for personalised advertising.