



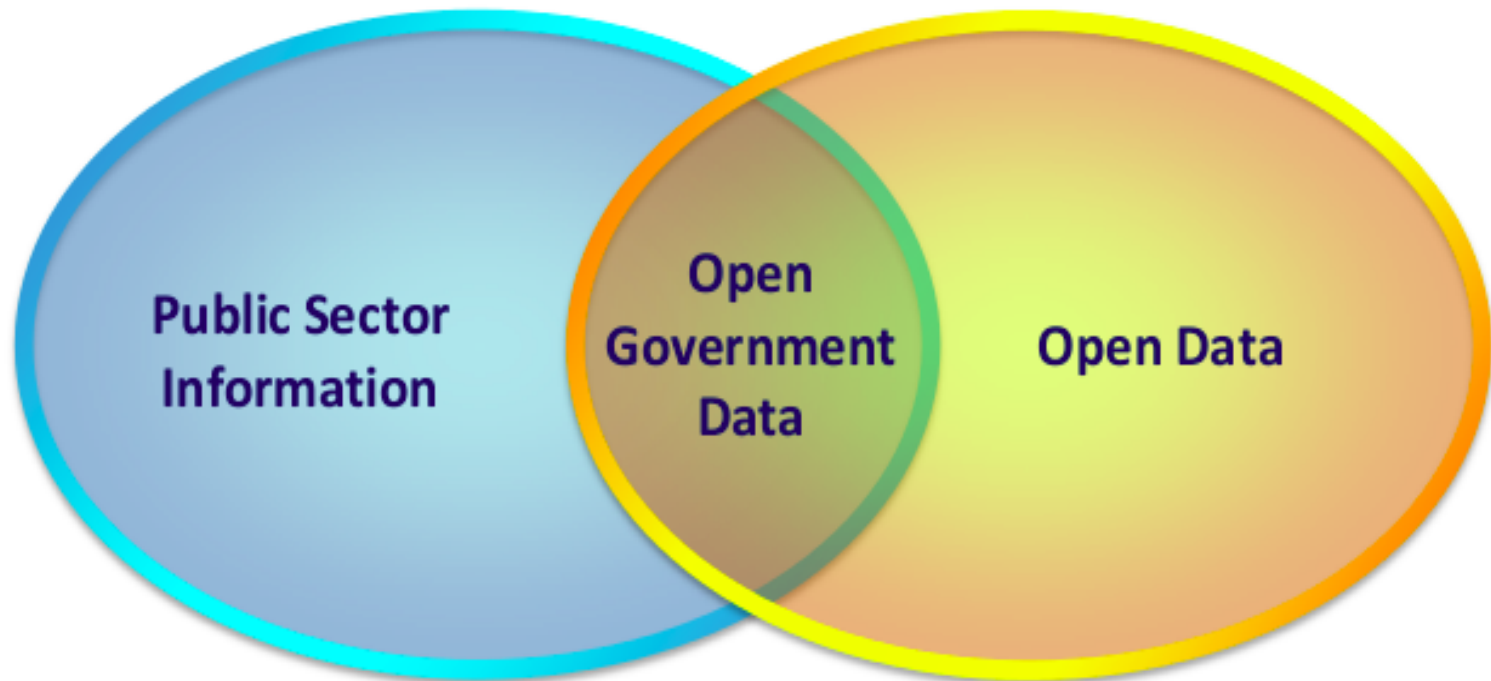
Nexa Center for Internet & Society

Politecnico di Torino

Studying the Internet, exploring its potential & experimenting new ideas

Open data e business models

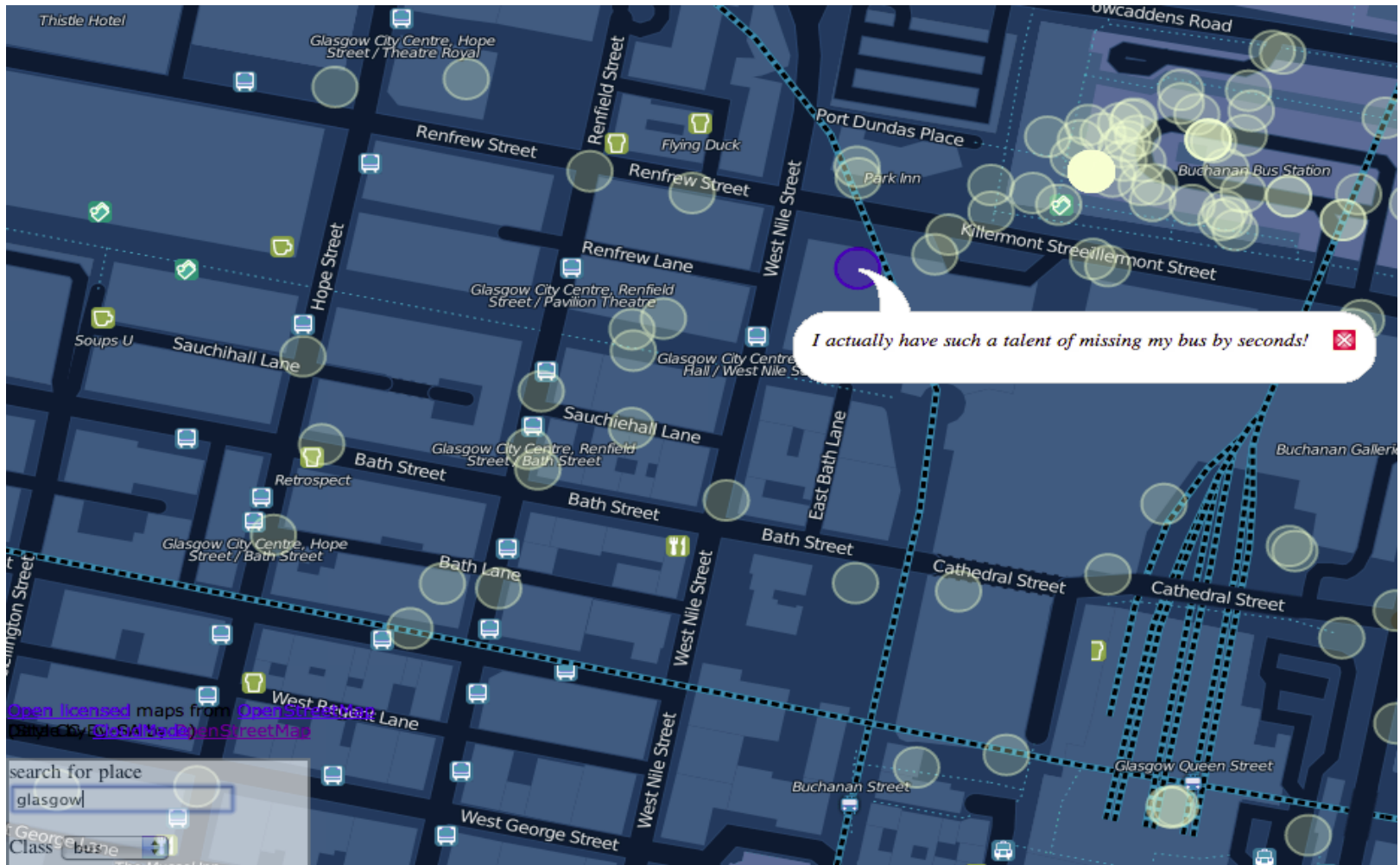
Il quadro di riferimento



Introduzione

- L'informazione del settore pubblico
 - *Upstream vs downstream*
 - Quali principi di tariffazione?
 - “Freemium”, ma con eccezioni...
- Beni pubblici ed effetti di rete
- Qualche censimento
 - ITA: <http://it-city.census.okfn.org/>
 - USA: <http://us-city.census.okfn.org/>
- Open Government vs Open Innovation

Placr



Placr

- Data service company founded in 2010 (Case study [here](#)).
- Aim: to create a single UK source of transport information.
- 2 key services:
 - **Transport API**, or 'cash and carry' service (data as a service, **freemium**)
 - placr.mobi consumer web app (Local authorities pay and place their label, it's free for end-users)
- Key advantages:
 - N°1 in integrating different data sources (live time tables)
 - Real time data.

Placr

- Which challenges?

Placr

- Which challenges?
 - certain **unreleased Oyster data** would tell us in real time how busy the London transport network is;
 - producing value, and therefore **revenue**, from the data;
 - ‘Some of the data streams that come to us need a lot of tidying up’ says Jonathan, ‘so we need to act swiftly and provide the **quality assurance** our customers demand’.

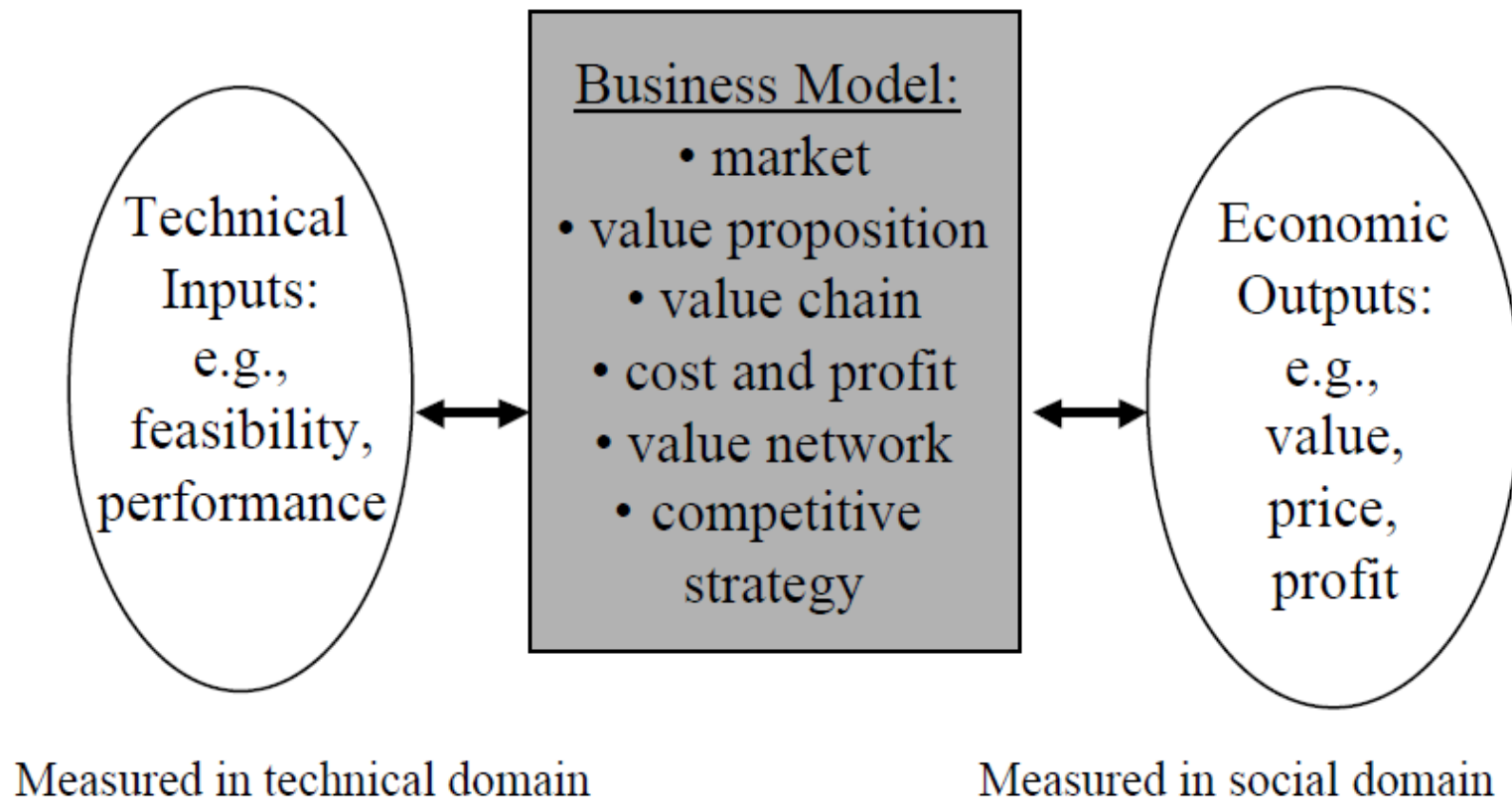
Cos'è un modello di business?

A business model is a conceptual tool that contains a set of **elements** and their **relationships** and allows expressing a company's logic of **earning money**. It is a description of the **value** a company offers to one or several **segments of customers** and the architecture of the firm and its network of **partners** for creating, marketing and delivering this value and relationship capital, in order to generate profitable and sustainable **revenue streams** (Osterwalder, 2004).

Altre definizioni

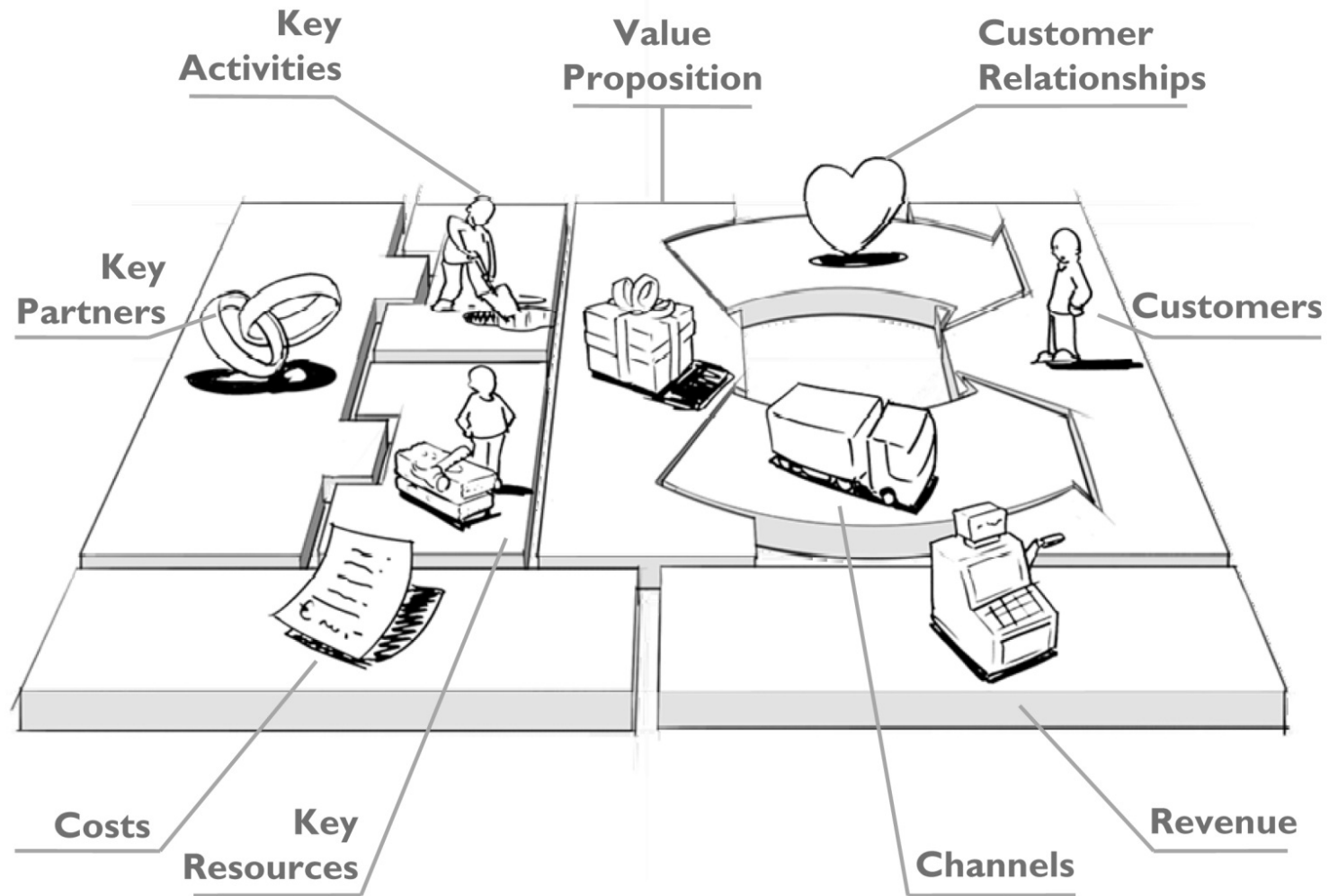
Autore	Definizione	Fonte
Paul Timmers	<i>'...the architecture for the product, service and information flows, including a description of the various business actors and their roles and a description of the potential benefits for the various business actors and a description of the sources of revenues.'</i>	'Business Models for Electronic Markets', 1998
Alexander Osterwalder	<i>'... a description of the value a company offers to one or several segments of customers and of the architecture of the firm and its network of partners for creating, marketing, and delivering this value and relationship capital, to generate profitable and sustainable revenue streams.'</i>	'Business model generation'
Christoph Zott e Raphael Amit	<i>'The business model is a structural template that describes the organization of a focal firm's transactions with all of its external constituents in factor and product markets'</i>	<i>Strategic Management Journal</i>
Steve Blank	<i>'A business model describes how your company creates, delivers and captures value... shows all the flows between the different parts of your company... how the product gets distributed ... how money flows back... company's cost structures..how company can work with other companies or partners to implement your business...'</i>	'The startup owner's manual'
Mark Johnson	<i>"Business model innovation is the key to unlocking transformational growth—but few executives know how to apply it to their businesses."</i>	'Seizing the White Space'

Le componenti fondamentali



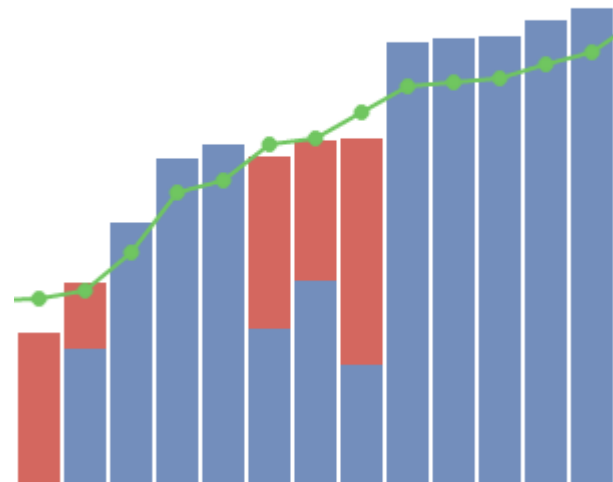
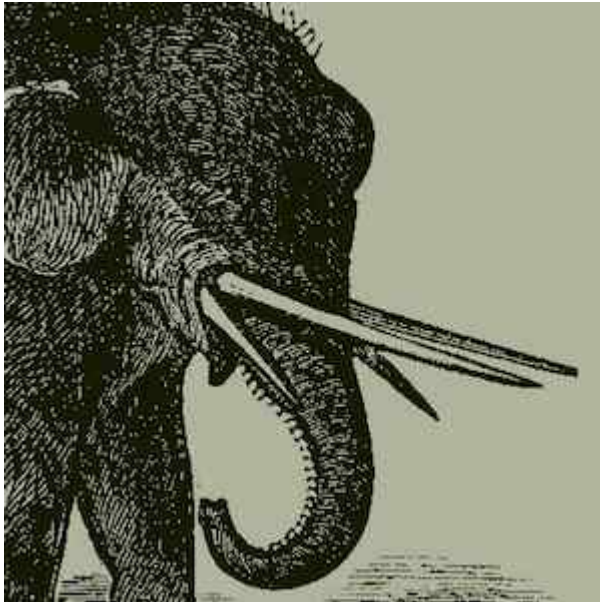
Chesbrough and Rosenbloom (2002)

Il business model canvas



drawings by JAM

Mastodon C



Mastodon C

- **Data-expert company** founded in 2012 (Case study [here](#)).
- “We built a dashboard to help other developers choose infrastructure to minimise their carbon emissions, **using open weather and power grid data**”.
- Mastodon C uses whatever data sets a client already has, combined with public data and social media.
- 2 key services:
 - customisable big data platform, 'Kiki' (using Hadoop);
 - advanced analytics and development, ‘Chaos to Insight’.

Mastodon C

- Which challenges?

Mastodon C

- Which challenges?
 - to innovate and to build **data products that clients really value**;
 - make people understand the **value the open-source technology** we use and how to use it.



Nexa Center for Internet & Society

Politecnico di Torino

Studying the Internet, exploring its potential & experimenting new ideas

**Public Sector Information (PSI):
i mercati fluidi**

PSI business models (Ferro, Osella 2012)

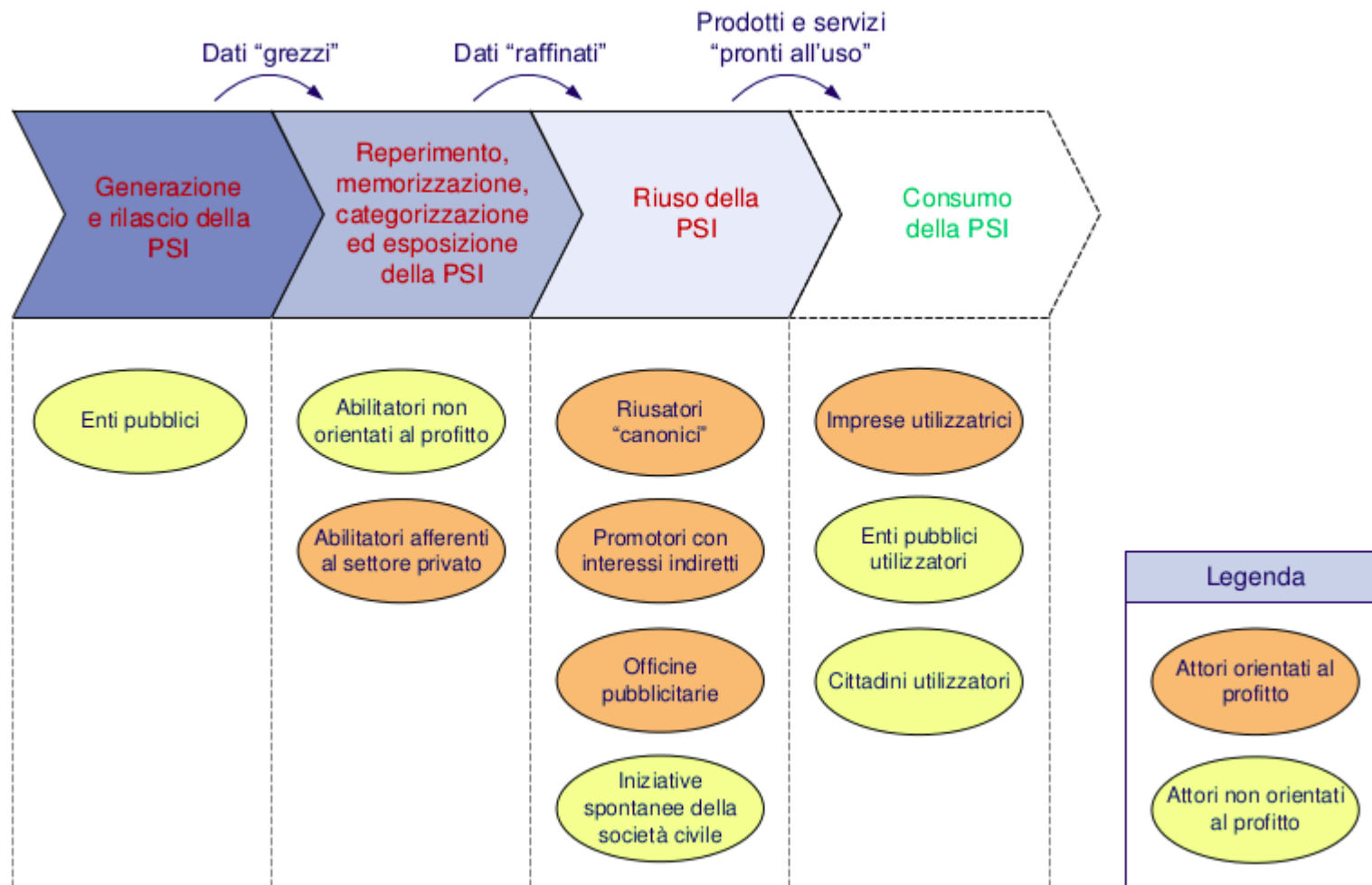
- **Obiettivo**

- Comprendere quali siano i meccanismi sottesi alla creazione del valore nell'ambito della “Public Sector Information” (specie quella pubblicata come 'open data').

- **Metodologia**

- 139 imprese recensite >> 13 casi di studio analizzati attraverso un modello di analisi “proprietario”.

Catena del valore



Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riutilizzo dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

3 dimensioni di analisi

Types of Data Elaboration	
<input type="checkbox"/>	Data aggregation
<input type="checkbox"/>	Data structuring and classification
<input type="checkbox"/>	Data geo-referencing
<input type="checkbox"/>	Data validation
<input type="checkbox"/>	Data mash-up
<input type="checkbox"/>	Visual analytics

Role of PSI in the Value Proposition	
<input type="checkbox"/>	Final good itself
<input type="checkbox"/>	Key ingredient of a product
<input type="checkbox"/>	Key ingredient of a service
<input type="checkbox"/>	Marginal ingredient of a product
<input type="checkbox"/>	Marginal ingredient of a service

Price Mechanisms		
<input type="checkbox"/> Premium	<input type="checkbox"/> Freemium	<input type="checkbox"/> Free
<input type="radio"/> À la carte	<input type="radio"/> Feature limited	<input type="radio"/> Advertising powered
<input type="radio"/> Subscription fee	<input type="radio"/> Time limited	<input type="radio"/> Cross subsidization
<input type="radio"/> Royalties	<input type="radio"/> Size limited	<input type="radio"/> Zero marginal cost

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riuso dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

PSI business models (Ferro, Osella 2012)

- **Risultati**

- Sfruttamento commerciale della PSI come **realità non marginale in Europa**, con due idealtipi: **mercati fluidi vs mercati oligopolistici**.
- Differenti posizionamenti: ad es., è possibile distinguere tra **riutilizzatori *tout court***, e **abilitatori / infomediari**.
- **8 modelli di business archetipali** a partire dall'incrocio di tre dimensioni.

I) Prodotti / servizi premium

- Elevato valore intrinseco dei dati, forniti a pagamento (ad es. Con tariffazione con componente fissa + variabile)
- Target per lo più B2B.
- Logiche di medio/lungo termine che vanno oltre il mero principio transattivo.

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riuso dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

I) Prodotti / servizi premium

Types of Data Elaboration	
<input checked="" type="checkbox"/>	Data aggregation
<input checked="" type="checkbox"/>	Data structuring and classification
<input checked="" type="checkbox"/>	Data geo-referencing
<input checked="" type="checkbox"/>	Data validation
<input checked="" type="checkbox"/>	Data mash-up
<input checked="" type="checkbox"/>	Visual analytics

Role of PSI in the Value Proposition	
<input type="checkbox"/>	Final good itself
<input checked="" type="checkbox"/>	Key ingredient of a product
<input checked="" type="checkbox"/>	Key ingredient of a service
<input checked="" type="checkbox"/>	Marginal ingredient of a product
<input checked="" type="checkbox"/>	Marginal ingredient of a service

Price Mechanisms		
<input type="checkbox"/> Premium	<input type="checkbox"/> Freemium	<input type="checkbox"/> Free
<input checked="" type="checkbox"/> À la carte	<input type="checkbox"/> Feature limited	<input type="checkbox"/> Advertising powered
<input checked="" type="checkbox"/> Subscription fee	<input type="checkbox"/> Time limited	<input type="checkbox"/> Cross subsidization
<input type="checkbox"/> Royalties	<input type="checkbox"/> Size limited	<input type="checkbox"/> Zero marginal cost

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riuso dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

I) Voglio il ruolo

- Piattaforma online che consente a insegnanti precari ed al personale scolastico di poter consultare velocemente tutti i dati relativi alle graduatorie.
- Per la legge sulla Privacy, avrebbero dovuto chiedere consenso al trattamento dati a 300'000 precari...
 - Ma il Garante della Privacy ha riconosciuto il valore del servizio (e, soprattutto, la compatibilità del suo oggetto con lo scopo con cui i dati vengono raccolti in origine).
- **Intervista ai fondatori**

I) Hospital Registers

- High quality medical market information.
- Access to detailed, accurate and up-to-date medical information on over 19,000 European hospitals in 25 countries plus a further 6,000 hospitals in the USA (*subscription based*).
 - Information present as downloadable tables, charts & interactive maps.
- Modelli di tariffazione: <http://www.walnutmedical.info/prices.html>

2) Prodotti / servizi freemium

- Fruizione gratuita delle funzionalità di base, pagamento per le funzionalità più avanzate.
- Approccio prevalentemente B2C, con logiche di gestione del cliente di breve/medio termine.
- Canali web e mobile (app).

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riuso dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

2) Prodotti / servizi freemium

Types of Data Elaboration

- Data aggregation
- Data structuring and classification
- Data geo-referencing
- Data validation
- Data mash-up
- Visual analytics

Role of PSI in the Value Proposition

- Final good itself
- Key ingredient of a product
- Key ingredient of a service
- Marginal ingredient of a product
- Marginal ingredient of a service

Price Mechanisms

- | | | |
|--|--|---|
| <input type="checkbox"/> Premium | <input type="checkbox"/> Freemium | <input type="checkbox"/> Free |
| <input type="radio"/> À la carte | <input checked="" type="radio"/> Feature limited | <input type="radio"/> Advertising powered |
| <input type="radio"/> Subscription fee | <input checked="" type="radio"/> Time limited | <input type="radio"/> Cross subsidization |
| <input type="radio"/> Royalties | <input checked="" type="radio"/> Size limited | <input type="radio"/> Zero marginal cost |

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riuso dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

2) Real-time transit data in Boston

- Massachusetts Bay Transit Authority making real-time public transport data available to third parties.
- Apps created on top of these data:
http://www.mbta.com/rider_tools/apps/
- Price depending on the number of API calls from developers (free of charge up to a certain amount).

3) Simil open source

- Offerta di prodotti, servizi o set di dati non “confezionati”, senza il pagamento di alcuna tariffa.
- Pagamenti per servizi a valore aggiunto o per modifiche alla licenza originaria.
- Sovvenzione incrociata: le entrate da ascrivere a queste linee di business supplementari vengono impiegate per coprire i costi dei servizi gratuiti.

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riuso dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

3) Simil open source

Types of Data Elaboration

- Data aggregation
- Data structuring and classification
- Data geo-referencing
- Data validation
- Data mash-up
- Visual analytics

Role of PSI in the Value Proposition

- Final good itself
- Key ingredient of a product
- Key ingredient of a service
- Marginal ingredient of a product
- Marginal ingredient of a service

Price Mechanisms

<input type="checkbox"/> Premium	<input type="checkbox"/> Freemium	<input type="checkbox"/> Free
<input type="radio"/> À la carte	<input type="radio"/> Feature limited	<input type="radio"/> Advertising powered
<input type="radio"/> Subscription fee	<input type="radio"/> Time limited	<input checked="" type="radio"/> Cross subsidization
<input type="radio"/> Royalties	<input type="radio"/> Size limited	<input type="radio"/> Zero marginal cost

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riutilizzo dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

3) OpenCorporates

opencorporates The Open Database Of The Corporate World 

Search

We have information on 3,896,898 companies

Filter by jurisdiction

- 41,002 Bermuda
- 32,620 Jersey
- 3,823,276 United Kingdom

Filter by data held

- 13,607 Company Address
- 4,955 Financial Transaction
- 88,803 UK Gazette Notice
- 960 Government Framework Supplier
- 12,685 UK Health & Safety Notice
- 748 Public Statement
- 724 Vat Number

Use Google Refine to match companies in your data

See screencast now

transactions MOO 09 Sep 2010 [Previous](#)

2267 rows

Show as: rows records Show: 5 10 25 50 rows

Cluster	Item	Expense type	Supplier
	1	Fees for professional services	A.A. SECURITIES LIMITED A.A. SECURITIES LIMITED (21) Credit note 1000 Search by name
	2	Purchase of inventory and services	AM PHARMACEUTICALS LTD AM PHARMACEUTICALS LIMITED (21) AM PHARMACEUTICALS LIMITED (20) Credit note 1000 Search by name
	3	Purchase of inventory and services	AM PHARMACEUTICALS LTD AM PHARMACEUTICALS LIMITED (21) AM PHARMACEUTICALS LIMITED (20) Credit note 1000 Search by name
	4	Purchase of inventory and services	AM PHARMACEUTICALS LTD AM PHARMACEUTICALS LIMITED (21) AM PHARMACEUTICALS LIMITED (20) Credit note 1000 Search by name
	5	Purchase of inventory and services	AM PHARMACEUTICALS LTD AM PHARMACEUTICALS LIMITED (21) AM PHARMACEUTICALS LIMITED (20) Search by name

3) OpenPolis



open**municipio**
Informati, monitora e intervieni nei lavori del tuo municipio



4) “Rasoio e lamette”

- Intermediari che facilitano l'accesso a risorse PSI / Open Data da parte di sviluppatori.
- Inizialmente un prodotto viene venduto ad un prezzo molto basso al fine di alimentare il successivo acquisto di un bene complementare (con domanda poco elastica).
- Esempio: set di dati pubblici su piattaforme cloud e rendendoli accessibili via API da chiunque, tariffando poi il solo utilizzo della capacità computazionale necessaria per il processamento dei dati.
- Anche qui sovvenzione incrociata.

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riuso dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

4) “Rasoio e lamette”

- ### Types of Data Elaboration
- Data aggregation
 - Data structuring and classification
 - Data geo-referencing
 - Data validation
 - Data mash-up
 - Visual analytics

- ### Role of PSI in the Value Proposition
- Final good itself
 - Key ingredient of a product
 - Key ingredient of a service
 - Marginal ingredient of a product
 - Marginal ingredient of a service

- ### Price Mechanisms
- | | | |
|--|---------------------------------------|--|
| <input type="checkbox"/> Premium | <input type="checkbox"/> Freemium | <input type="checkbox"/> Free |
| <input type="radio"/> À la carte | <input type="radio"/> Feature limited | <input type="radio"/> Advertising powered |
| <input type="radio"/> Subscription fee | <input type="radio"/> Time limited | <input checked="" type="radio"/> Cross subsidization |
| <input type="radio"/> Royalties | <input type="radio"/> Size limited | <input type="radio"/> Zero marginal cost |

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riuso dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

4) Public data sets su Amazon Web Services

- <http://aws.amazon.com/public-data-sets/>

The screenshot shows the AWS website page for 'Public Data Sets on AWS'. The page features the AWS logo and navigation links at the top. A sidebar on the left contains 'Related Resources' such as 'Public Data Sets Catalog', 'Subversion Forum', 'Community Forum', 'Amazon EC2', and 'AWS Management Console'. Below this is a promotional banner for 'Numerate' with the headline 'Reduce Your EC2 Bill' and a sub-headline 'See how customers like Numerate reduced their Amazon EC2 bill by 40% in just 5 days using Spot Instances'. The main content area is titled 'Public Data Sets on AWS' and contains the following text:

Public Data Sets on AWS provides a centralized repository of public data sets that can be seamlessly integrated into AWS cloud-based applications. AWS is hosting the public data sets at no charge for the community, and like all AWS services, users pay only for the compute and storage they use for their own applications.

Previously, large data sets such as the mapping of the Human Genome and the US Census data required hours or days to locate, download, customize, and analyze. Now, anyone can access these data sets from their Amazon Elastic Compute Cloud (Amazon EC2) instances and start computing on the data within minutes. Users can also leverage the entire AWS ecosystem and easily collaborate with other AWS users. For example, users can produce or use prebuilt server images with tools and applications to analyze the data sets. Users can also discuss best practices and solutions in the dedicated Public Data Sets Forum.

By hosting this important and useful data with cost-efficient services such as Amazon EC2, AWS hopes to provide researchers across a variety of disciplines and industries with tools to enable more innovation, more quickly.

At the bottom of the page, there is a section titled 'This page contains the following categories of information. Click to jump down:' with two links: 'Available Public Data Sets on AWS' and 'How to Share a Public Data Set on AWS'.

5) Piattaforme orientate alla domanda

- Obiettivo: fornire agli sviluppatori consistenti facilitazioni nell'accesso a risorse PSI, le quali vengono memorizzate su server intermedi proprietari.
- Dati catalogati mediante metadati, armonizzate in termini di formati ed esposte attraverso API, rendendo agevole il reperimento dinamico dei dati in maniera on-demand.
- “One-stop shops”.
- Tariffazione tipicamente freemium.

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riuso dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

5) Piattaforme orientate alla domanda

Types of Data Elaboration

- Data aggregation
- Data structuring and classification
- Data geo-referencing
- Data validation
- Data mash-up
- Visual analytics

Role of PSI in the Value Proposition

- Final good itself
- Key ingredient of a product
- Key ingredient of a service
- Marginal ingredient of a product
- Marginal ingredient of a service

Price Mechanisms

- | | | |
|--|--|---|
| <input type="checkbox"/> Premium | <input type="checkbox"/> Freemium | <input type="checkbox"/> Free |
| <input type="radio"/> À la carte | <input checked="" type="radio"/> Feature limited | <input type="radio"/> Advertising powered |
| <input type="radio"/> Subscription fee | <input type="radio"/> Time limited | <input type="radio"/> Cross subsidization |
| <input type="radio"/> Royalties | <input type="radio"/> Size limited | <input type="radio"/> Zero marginal cost |

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riuso dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

5) Data markets

	Azure	Datamarket	Factual
<i>Data sources</i>	Broad range.	Focus on industry and country data, + others.	Mostly geodata.
<i>Free data</i>	Yes	Yes	No
<i>Delivery</i>	OData API.	API, downloads.	API, downloads for heavy users.
<i>Data publishing</i>	Via database connection, or web service.	Upload or web/database connection.	Upload or web service.
<i>Data reselling</i>	Yes, 20% commission on non-free datasets	Yes. Fees and commissions vary.	No

6) Piattaforme orientate all'offerta

- In questo caso paga la PA.
- Interfacce applicative (API) + interfacce grafiche (GUI).
- Esempio di mercato “multi-versante”.
- Riferimento a:
<https://nexa.polito.it/nexacenterfiles/lezione-05a-piattaforme.pdf>

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riuso dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

6) Piattaforme orientate all'offerta

Types of Data Elaboration	
<input checked="" type="checkbox"/>	Data aggregation
<input checked="" type="checkbox"/>	Data structuring and classification
<input type="checkbox"/>	Data geo-referencing
<input type="checkbox"/>	Data validation
<input type="checkbox"/>	Data mash-up
<input checked="" type="checkbox"/>	Visual analytics

Role of PSI in the Value Proposition	
<input type="checkbox"/>	Final good itself
<input type="checkbox"/>	Key ingredient of a product
<input checked="" type="checkbox"/>	Key ingredient of a service
<input type="checkbox"/>	Marginal ingredient of a product
<input type="checkbox"/>	Marginal ingredient of a service

Price Mechanisms		
<input type="checkbox"/> Premium	<input type="checkbox"/> Freemium	<input type="checkbox"/> Free
<input type="radio"/> À la carte	<input type="radio"/> Feature limited	<input type="radio"/> Advertising powered
<input checked="" type="radio"/> Subscription fee	<input type="radio"/> Time limited	<input type="radio"/> Cross subsidization
<input type="radio"/> Royalties	<input type="radio"/> Size limited	<input type="radio"/> Zero marginal cost

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riuso dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

6) Open data sw platforms



ckan



An Infrastructure for Open, Linked
Governmental Data Provision towards
Research Communities and Citizens



Co-funded by the European Union



7) Promozione del marchio

- Offerta di prodotti o servizi basati su PSI / Open data senza realizzare diretto ricavo;
 - l'attività promozionale è volta a generare ricadute positive in senso lato.
- Open data come strumento di attrazione.
- Eventualmente in modalità 'white-label' (8° archetipo).

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riuso dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

7) Promozione del marchio

Types of Data Elaboration

- Data aggregation
- Data structuring and classification
- Data geo-referencing
- Data validation
- Data mash-up
- Visual analytics

Role of PSI in the Value Proposition

- Final good itself
- Key ingredient of a product
- Key ingredient of a service
- Marginal ingredient of a product
- Marginal ingredient of a service

Price Mechanisms

- | | | |
|--|---------------------------------------|---|
| <input type="checkbox"/> Premium | <input type="checkbox"/> Freemium | <input type="checkbox"/> Free |
| <input type="radio"/> À la carte | <input type="radio"/> Feature limited | <input checked="" type="checkbox"/> Advertising powered |
| <input type="radio"/> Subscription fee | <input type="radio"/> Time limited | <input type="radio"/> Cross subsidization |
| <input type="radio"/> Royalties | <input type="radio"/> Size limited | <input checked="" type="checkbox"/> Zero marginal cost |

Fonte : Ferro, E., & Osella, M. (2011). Modelli di Business nel Riuso dell'Informazione Pubblica. Osservatorio ICT della Regione Piemonte.

7) Promozione del marchio

- IBM City Forward
- IBM Many Eyes
- Google Public Data Explorer



Nexa Center for Internet & Society

Politecnico di Torino

Studying the Internet, exploring its potential & experimenting new ideas

PSI: i mercati oligopolistici

L'informazione commerciale

- Sistema delle Camere di Commercio (CCIA), raggruppate in Unioncamere.
 - Il Registro delle Imprese, gestito da Infocamere, è la principale banca dati.
- Il Registro delle Imprese è interrogabile da singoli utenti (in parte a pagamento).
- Vi è un numero limitato di distributori (57 nel 2012) che, oltre ad assolvere alcuni **vincoli** di capacità tecnologica e commerciale, corrispondono una fee annuale per il riutilizzo dei dati in blocco.
 - Mercato a valle dell'ordine dei 600 milioni di Euro / anno (a fronte di ca. 30 milioni / anno incassati da Infocamere dai distributori)
 - Il sistema camerale percepisce diritti di segreteria anche per l'iscrizione dei dati da parte delle imprese (per obbligo di legge).

I dati territoriali / catastali (1/2)

- Principale soggetto detentore: Agenzia del Territorio.
- Riutilizzatori: cittadini per la singola visura, ma soprattutto imprese/studio professionali, e istituti di credito.
- Tre tipologie di rilascio:
 - 1) dati grezzi relativi a operazioni commerciali immobiliari;
 - 2) “ricerca continuativa per via telematica” (piattaforma SISTER);
 - 3) Osservatorio del Mercato Immobiliare.

I dati territoriali / catastali (2/2)

- Tariffazione
 - gratis per le PA;
 - download in bulk: quota fissa + abilitazione annuale
 - 2006: maggiorazione per riutilizzi commerciali (di fatto)
- Rischi di “doppia marginalizzazione” (a discapito dell'utente finale), oppure di “foreclosure” del mercato a valle.



Nexa Center for Internet & Society

Politecnico di Torino

Studying the Internet, exploring its potential & experimenting new ideas

Data Ingredients: open data and smart disclosure. The case of energy efficiency.

Smart disclosure

- Dalla “product-attribute information” alla “product-usage” information.
- C'è chi investe per acquisire e gestire dati di consumo (cosa che noi non facciamo).
- Esempio: «Record, Evaluate and Compare Alternative Prices» (RECAP) model (UK).
- <http://www.supermoney.eu/>

Smart disclosure



Objectives

- (1) Compare different government-enabled data disclosure policies, highlighting complementarities.
- (2) Analyse the key aspects of a gov-enabled data disclosure initiative (Green Button on energy consumption).
- (3) Isolate key data ingredients to make Green Button-based apps actually viable.

Open stuff

- **Open government data:** everyone free to reuse it for any purpose, subject at most to ~~share-alike~~ attribution;
 - e.g., a Geo-JSON file describing points of interest in a given area.
- **Smart disclosure:** timely release of consumption data in standardized, machine-readable formats;
 - e.g., my hourly energy consumption in XML format.
- **Open corporate data:** same as open government data, but provided by companies;
 - e.g., a .csv containing accounting data of an electric utility.



	Open government data	Smart disclosure	Open corporate data
Example (in the energy field)	Dataset reporting energy efficiency levels of buildings in a given area.	Granular personal consumption patterns of a single consumer served by a specific utility.	Accounting data of a energy utility company.
Government role(s)	Direct promoter.	Identifier of the opportunity; challenger for the industry; direct promoter (where industries are regulated).	Challenger for the industry.
Prominent objectives	Government transparency and accountability; economic growth.	Consumer empowerment; market efficiency.	Corporate accountability; corporate social responsibility.
Drivers for the promoters	Policy agenda; competitive regulation; preserving credibility towards citizens; direct benefits from new services.	Regulatory constraints; improved reputation; increased knowledge on consumer behaviour; consumer retention.	Regulatory constraints; improved reputation; opportunities of collaboration with third parties.

Open government data**Smart disclosure****Open corporate data****Prominent technical aspect(s)**

(Legal and technical) interoperability; standardisation of data formats; quality of service.

Standardisation of data formats.

(Legal and technical) interoperability; standardisation of data formats.

Publication means

Erga omnes through open data portals, as static datasets, linked data (e.g., RDF), and/or open data services (e.g., via APIs).

To each consumer as static files (eg., XML), and/or as dynamic data via APIs.

Erga omnes through dedicated websites, as static datasets, linked data (e.g., RDF), and/or, -but less frequently- open data services (e.g., via APIs)

Prominent legal/regulatory aspect

Intellectual Property Rights (including database rights); previous agreements with third parties.

Privacy and data protection; competition.

IPRs and privacy.

Complementary tools

- **Used jointly**, different types of data may serve the same purpose
- For example, to create a service that enables consumer switching from an energy provide to another, one needs, at least:
 - **actual energy consumption data** of a user over a relevent period of time;
 - information about **buildings efficiency** (to set the benchmark);
 - and **energy rates**.
- These three layers of information are disclosed by different actors, and refer to different paradigms: smart disclosure, open government data, and open corporate data respectively.

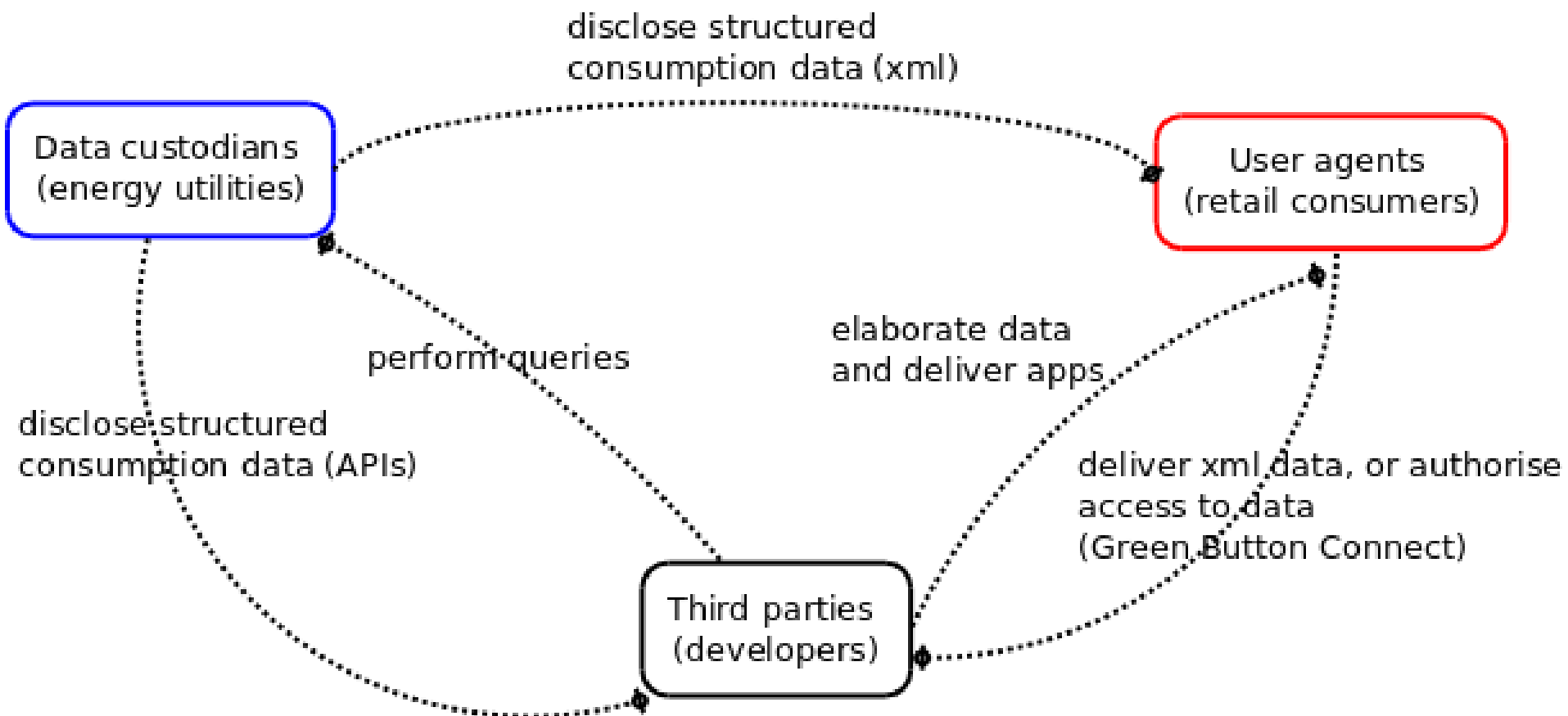
Green Button

- An initiative promoted by the U.S. Department of Energy in 2011, aimed at **disclosing to energy customers their digital consumption data**;
- The U.S. Government:
 - defined guidelines and **standards** for data publication;
 - engaged with utilities to make them **adopt common practices**;
 - set the stage for the creation of **apps**, e.g., through competitions.
- As of early 2013, 58 energy utilities serving around 36 million customers implemented or committed to Green Button.

A short step back

- Governmental policies set aim at **rationalising energy usage**. **Data play a key role** in this.
- Smart metering allows to **manage load balance more effectively**.
- Incentives for utilities to disclose granular consumption data might instead **get weaker**.
- In this respect, gov moral suasion can play a role... but **standardisation is the actual answer**

Back to Green Button



Energy apps

- Stemmed from Green Button, the U.S. initiative of smart disclosure of energy data to consumers (involving 50+ utilities to date).
- 17 breakthrough applications that, for instance:
 - benchmark your consumption against the one of other buildings;
 - project your expenditure with other energy providers;
 - calculate your energy score;
 - suggest possible energy savings.
- Which data do you need to do that?

Main findings | I

- 3 main categories of apps:
 - **suggesting efficiency-improving changes** in the consumption behaviour of their user (“Enabling consumption choices”);
 - **benchmarking** and scoring residential energy efficiency (“Scoring”);
 - **deriving trends and hints** from consumption data (“Creating insight”)
- All apps in this sub-sample perform some kind of data analysis and integration.
- 82% of them are also relying on at least one open government data source as a key ingredient.

Type of Green Button-based app	Enabling consumption choices	Scoring	Creating insight
Number of apps in the sample	8	4	5
Most used open government data	Electricity cost/price data (3)	Buildings performance data (2)	Average energy consumption (2); Weather data (2).
Most used proprietary/other data	Real estate data (3) Utility rates (2)	(none)	
Also retrieving data via utilities' APIs - occurrences	2	2	1
Enabling social interaction between users – occurrences	3	2	1

Main findings | 2

- Apps suggesting to a user a more efficient consumption behaviour also benefit from average retail electricity cost/price information
- Energy efficiency 'scoring' apps also need, at least, structured and updated information on buildings performance
- Value-added services that derive insights from consumption data frequently rely on average energy consumption information.

Green Button data

Standardized XML describing hourly consumption of a user (or meter reading through smartphone)

U.S. Building performance database

Age, size, number of occupants, etc., and shows us the annual electricity usage for 200K US buildings

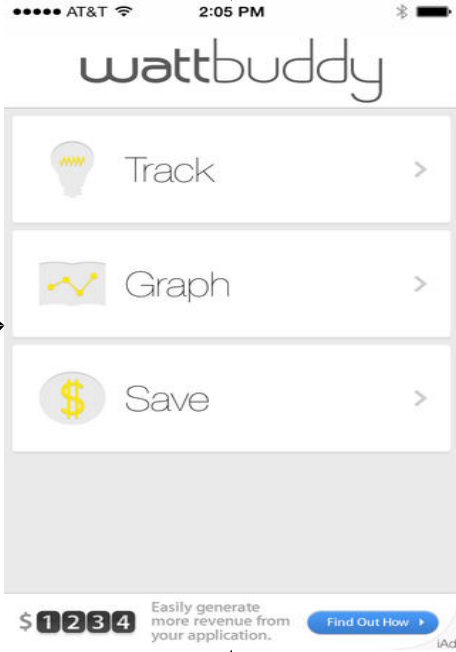
Avg. electricity cost In the user's area

to tell you how much you are probably spending right now

Utility rates

to calculate possible savings

Specs of the particular house



your energy usage

over time

by looking at competitive offers in unregulated markets

Smart disclosure

Standardized XML describing hourly consumption of a user (or meter reading through smartphone)

Federal open government data

U.S Buildings performance database : Age, size, number of occupants, etc., and shows us the annual electricity usage for 200K US buildings

Federal open government data

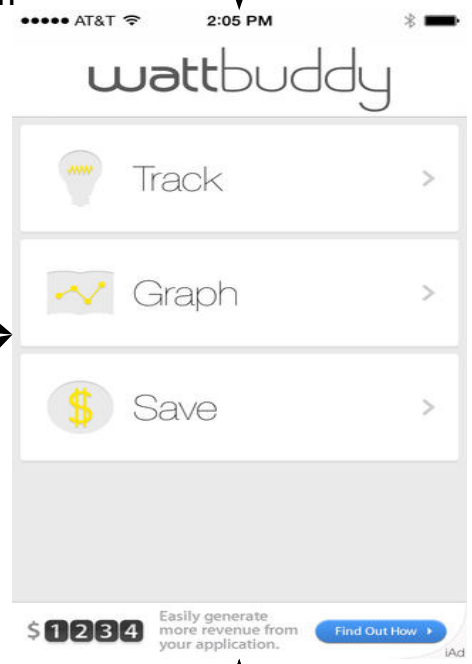
Avg electricity cost in the user area, to tell how much a consumer is currently spending

Utility rates, to calculate possible savings

Local open government data

Other (proprietary) data

Specs of the particular house (via Zillow)



your energy usage

over time

by looking at competitive offers in unregulated markets

	USA (Federal agencies)	EU (Open Data portal, Eurostat, EU agencies)	Italy (Open Data portal, national agencies)	UK Open Data portal
Energy consumption	.csv	DFT, SDMX-ML, TSV	.csv, SDMX-ML	.csv, .xls
Buildings performance	tabular (API queries available)	poor match	no match	.pdf
Electricity cost/price	.xls	.csv, .xls	.html (poor match)	.pdf
Commercial offers in unregulated markets	available only in two States (as tabular data)	no match	no match	no match

	USA (Federal agencies)	EU (Open Data portal, Eurostat, EU agencies)	Italy (Open Data portal, national agencies)	UK Open Data portal
Census / demographic information	.csv	.csv, .xls	csv, SDMX-ML	.csv, .xls, .html
Weather data	.xml via RestFUL APIs	ASCII	csv, SDMX-ML	RDFa
Buildings energy certificates	single queries	no match	.csv	Existing, but unpublished
Energy incentives	OpenXML; JSON	no match	.xml	.xml