

Open Data Economy: from wow to how

WOW

OD is a big opportunity, hundreds millions of dollars of potential sitting there.

Drivers of public data release: accountability, transparency and FOI, provision of 'fresh' inputs to researchers but it is also a way to create new businesses.

Several institutions have tried to quantify the impact of the reuse of PSI:

- Direct impact → Several tenths of billions of euros per year.
- Indirect impact (public data combined with other information sources) → Several hundreds of billions of euros per year.
- Induced impact (including also the efficiency of private and public processes and a general consumer surplus) → Around 1K billions of euros per year (source: McKinsey).

HOW

Deloitte: in 2012 says that is not possible to quantify the economic value of open data since no evidence is available yet. As a matter of fact, actual reuse is striving to take off.

Even though public sector information is superabundant, it is one of the largest underexploited assets of nowadays society.

Problem setting:

- The initial great deal of excitement cultivated a techno-enthusiastic view. A similar approach loses sight of the principles and mechanisms that allows the open data economy to scale and become relevant ("how").
- Scholars and practitioners have put an enormous effort into evincing "how much" value could be extracted from public data without considering "how" this could happen.
- The moderate pace of market expansion cannot be exclusively ascribed to the paucity of entrepreneurial maturity. The success of open data reuse passes through a vital input that is released upstream. Consequently, also the supply side has to be included into the big picture.

Goal

Get a unique and meaningful picture based on evidence and real-life experiences brought by the audience, understand barriers and address blind spots both on the demand side and on the supply side to catalyse the uptake of the open data economy.

The demand side of the Open Data economy. Elevator pitches: What is your business about? What are the key barriers?

Evidence of barriers and success – what needs to be done

- 1- **Compass publishing.** Customers: banks and other big companies. Use company register since 100 years, problems in getting those data as the ministry of justice trying to prevent to use it...law trial that cost billions (?)

- 2- **Miguel Garcia – Zabala Consulting – Coordinator of FINODEX incubator.** Looking for companies that want to reuse OD. Problems:
 - a. In ICT sector, people do not know what OD is, they don't know how to reuse. Awareness should be increased.
 - b. Problem with licensing. For some company anything published on the web is open data, company don't consider license restrictions. More awareness about licensing should be raised.
- 3- **OD Institute.** Promote reuse of OD. Quite often businesses are already relying on and using OD but they don't know. They don't know it until you explain they are using it. There is a need to focus on sustainability. Really valuable data still need to be published, that's where the biggest opportunity lies. We should there focus on deliver OD (supply side problem).
- 4- **Miguel – Portuguese state- funded OD initiative in transportation.** He is not using OD but he has been opening it. He aggregates data, standardizes it and makes it available. The most difficult part is to convince people of benefits stemming from OD. The main transport agency has a contract with a company to give info only to them (that company has an app that reuses them). It's hard to get access to this data. Success: they succeeded to make some apps to expand in Portugal some (not specified) offering access to OD.
- 5- **Simon – freelance Birmingham.** SW consultant service in the area of OD. They probably sitting in the wow. They found easy to do consultancy. Problem: turning nice idea into reality. Hard to get the data. Positive: they knew somebody that could get the data (informal approach).
- 6- **Helsinki community.** Very strong start up community. They are eager to use data, but Finnish market is small and not lucrative, moreover data is not standardized. To use data more intensively, it should be standardized.

Summarizing, there are 3 main barriers:

- 1- Awareness of what OD is. There is the need to educate potential entrepreneurs on what OD is (and is not).
- 2- Access to data.
- 3- Once OD is available, the problem is maintainability (in terms of data quality, frequency of update, formats, licenses).

Access and maintainability are problems inherited from **upstream**. To partially mitigate these problems, there is a role played by intermediaries (ex: Socrata, data markets).

Or the solution is in generation and release? There are some sort of enablers but is not known if they have market.

Pull model: there needs to be someone to request data... ("hungry people"). Business model works only in this way. The government doesn't need to do anything if people don't ask for it. PSI is already a billion dollar market (ex: GPS data is PSI). Pushing out is not enough... if nobody is interested, nobody is going to reuse.

Costly dataset: government knows which are the valuable data, in fact they sell those data for high price. Lowering down this price would be an "investment" (in the guise of revenue reduction).

Importance of citizen participation There should be the understanding of needs expressed by the citizens in order to see how OD could be appropriately used.

Possible solutions: 1) there are places in which OD is producing value... copy good idea. 2) hear the voice of citizens to understand what datasets/services could be useful, see what they are hungry for.

Upstream or downstream?

- Not good compliance with Open definition. There is a bottleneck on the **supply side**.
- Barriers to reuse **downstream**? Lack of ideas of entrepreneurs (government cannot do anything about it)
- Lack of data? Government should hear the voice of the entrepreneurs to understand what to open (upstream + downstream)
- Another problem: process of publication always changes...we don't have a constant way of approaching this open data process

Value of OD for public sector bodies in order that they publish?

What are the hard choices that public authorities (supply side) have to do while starting a OD program? Considering budget constraint, considering that they don't know the demand and the usefulness of that data...

Problem: PAs (Public Administrations) publish data but the benefits are somewhere else since PAs cannot internalize such benefits.

However, publishing OD could make PA communicate better between one another.

Not often data are published because of the benefits but because for obligation or because is fashionable...Today OD is seen as a cost by PA that don't have budget enough.

What are the situations in which there could be a tangible return that can be internalized by PA?

- "Information bus" connecting various departments by tearing down existing silos.
- OD could spot inefficiencies and other wastes by analysing data (that could also be outside the PA – free outsourcing).
- OD could also stimulate the action of the community (by leveraging "collective intelligence") targeting specific underserved (or unserved) needs.
- The government can limit itself to make available raw data (thus, achieving savings) and external innovators will build solutions on top of such data ("government as a platform", as theorized by Tim O'Reilly).

Could we spot cases in which the PA could think that publishing data could create internal benefits. Is a WIN-WIN situation (advantages on both sides) real or utopia?

- 1- Creation of jobs, tax revenue = from private company
- 2- Saving money due to improvement of efficiency: what type of saving, what guidelines to give to a policy maker?
 - a. There is no evidence but it's positive that they must publish (due to law enforcement) and benefits might show

Generally however it's hard to convince policy makers. PA is a huge machine in which it is really complex to convince people of indirect (and some direct) benefits of OD.

For example: some discovered benefits (already discovered by some start-up in the UK) but politicians don't react to this. Government reaction is the difficult and uncertain thing, not only showing the benefits of OD.

Conclusions

Barriers are mainly related to what happens upstream. We started from the demand side, we deep-dived into the supply side, we highlighted some of the key choices that need to be made: we need evidence to show to policy makers that opening data vaults is not a cost but it will bring benefits. Opening data will be beneficial for governments according to an inward orientation: no more open data as a fad or an obligation, but as a necessity.