

# Open Data: A necessary step to Open Access

**Timos Sellis**

Institute for the Management of Information Systems

“Athena” Research Center

[timos@imis.athena-innovation.gr](mailto:timos@imis.athena-innovation.gr)

# Open Data

- Data is **open** if it can be freely **used, reused,** and **redistributed** by **anyone** (*opendefinition.org*)
- Open Data go way back, in 1957 (World Data Center system)
- Much activity and publicity lately in Web 2.0/3.0 (e.g. [data.gov](http://data.gov), [data.gov.uk](http://data.gov.uk))

# Context

## Politics

*Public data means online data in real time. We must **democratize** data. People want the raw data. The public can handle the data and create **ideas** and move government forward. This movement is the **third economic revolution** in the USA: the agricultural revolution, the industrial revolution, and the **transparency of information revolution***

*Vivek Kundra, USA  
Chief Information Officer*

## S&T

*Numerous scientists have pointed out the irony that right at the historical moment when we have the technologies to permit worldwide availability and distributed process of scientific data, broadening collaboration and accelerating the pace and depth of discovery.....*

***we are busy locking up that data and preventing the use of correspondingly advanced technologies on knowledge***

*John Wilbanks,  
Executive Director  
Science Commons*

# Driving factors

## Political

- Transparency
- Accountability
- Freedom of information
- Tax payers money
- Cost cutting
- Productivity gains for the public sector

## Technological

- Crowdsourcing
- Mashups
- Social Networks
- Open standards
- Web 2.0
- Linked Data

# Another perspective: sustainable growth

- A KPMG study on the Canadian Spatial Data Infrastructure (2001), concluded that **closed, restricted data has major economic harm.**

*"the consequences [of cost recovery] for businesses are **higher marginal costs, lower research and development investments and threatened marginal products.** The results for consumers are **negative:** higher prices and reduced products and services. The overall economic consequences... are fewer jobs, reduced economic output by almost **\$2.6 billion** and a lower gross domestic product."*

# Convergence

- This mix of technological, political, and economical support is unique
- Excellent opportunities for:
  - Sustainable growth and business development
  - Advances in research and innovative products
  - Cross disciplinary research agendas

# Learning from geospatial data

# Our experience

- IMIS/RC “Athena” is at the forefront of the open data movement in Greece, with emphasis on geospatial data
  - Geodata.gov.gr (open geospatial data)
  - GeoDataCamp (hachathons to open up data)
- Technical challenges and solutions are **similar**
  - More on that later



# Geospatial Data

- A great source for examples
  - Account for roughly 80% of government data
  - International, EU (e.g. INSPIRE Directive), and national efforts to promote reuse and openness (Spatial Data Infrastructures)
  - Well documented use cases for ROI and CBA.

# The ecosystem in a nutshell

- Standardization
  - ISO, Open Geospatial Consortium, INSPIRE Directive, OSGeo
- Spatial Data Infrastructures from mid 80's
  - Lots of experience acquired
- Vendor support
  - 99% of commercial offerings are compliant
- Open Source
  - Catalogues, services, ready to use SDIs

# Across the EU

- Cataluña
  - 1,5M investment in SDI (2002-2006)
  - 500h per month -> 2,6M in efficiency savings
- Lombardia
  - 4M investment in SDI (2006-2008)
  - 3M per year cost savings (11% time savings, 17% cost savings)
- Denmark - In 2002 the Danish government decided to provide all addresses, free of cost
  - ROI is 70:1 (\$250K to publish the information, \$18M direct economic benefits)

# The steps for open data

# Open Data is the last step

- Consider Open Data as the **last step**
  - Where you simply choose an appropriate license
- The main focal point is how to make your resources (data & services) **discoverable**
- **Step 1**
  - Standard metadata
  - Unique resource identifiers (data & services)
  - Catalogue

# Step 2

- After the first step, each one has a catalogue (repo) with his/her resources. Moving on:
- **Step 2**
  - Catalogues to aggregate metadata across various catalogues
  - Harmonized search and APIs
  - User Management
- We now have an infrastructure to support discovery across all resources

# Step 3

- Now decide on the business models, i.e. licensing
  - Harmonized licenses (standardized) is a plus
- Sell?
  - Electronic marketplace
- Pool resources?
  - Decide common sharing within a group?
- Open?
  - Various shades of “*openness*”

# Contributing to the discussion



# Questions

- What is the market **value** of Open Data?
  - Establishing a data economy on OD should first focus on identifying exactly where we stand, at an international level. We need to study and estimate the market value.
- What is the **potential** of the market?
  - Given the needs, priorities, and goals, is there room for growth? If so, how can growth be achieved? Do we need regulation, or the market forces can pursue this goal alone?
- Are open data **discoverable**?
  - To promote the data economy we should first ensure that OD resources can be found, even on the lowest technical level. Where can a potential client search for all available OD (data and services)?

# Questions

- What do we **gain** by **closing** data?
  - Are business models based on closed and heavily guarded OD actually successful? Are we losing opportunities for growth by not systematically exploiting common sharing and synergies?
- What do we **lose** by **opening up** data?
  - Can the direct income lost compensated by direct or indirect financial gains?