Measuring the Digital Economy
Opportunities & challenges from new data sources & technologies

Jesus Vega Villa & Giuditta de Prato

Big Data for the Analysis of the Digital Economy and Society
JRC IPTS, September 22, 2015

Institute for Prospective Technological Studies - IPTS
Joint Research Centre - European Commission

Disclaimer: The views expressed are those of the presenter and may not in any circumstances be regarded as stating an official position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this presentation.
Mission:
To provide science-based responses to policy challenges that have both a socio-economic as well as a scientific/technological dimension.

This techno-economics research is carried out by multi-disciplinary teams, whose core activities are:
- to provide client-specific policy option analysis by conducting targeted studies;
- developing and running economic models;
- providing policy intelligence platforms.
Digital economy: our motivations

- Analysis of how (ICT) innovation & new technologies affect economy
- ICT industry, innovation, R&D, entrepreneurship, platforms..
We have accumulated data quantitative analysis on many established data sources:

- Eurostat data mainly (Natl Accounts, LFS, SES, SRD, SBS, CIS, Surveys)
- EC databases, FP7, H2020, etc
- Other data sources (OECD, private providers, NSOs)

We aim at complementing existing statistics and analyses with new data sources and methodologies to track the digital transformations affecting the economy

- by building new indicators hitting still unexplored dimensions,
- by better exploiting heterogeneous data
- getting more timely information

We still have many open questions..
### PREDICT Project

**Prospective Insights on R&D in ICT**

<table>
<thead>
<tr>
<th>Objectives &amp; Description</th>
<th>Policy Relevance</th>
<th>Reports</th>
<th>Scientific Publications</th>
<th>Events</th>
</tr>
</thead>
</table>

#### PREDICT 2015 REPORT: An Analysis of ICT R&D - EU & beyond

- [Online version of report](#)
- [Downloadable PDF version](#)

**ICT R&D performance across the world – patent-based evidence**

#### PREDICT 2014 REPORT: An Analysis of ICT R&D - EU & beyond

- [Online version of report](#)
- [Downloadable PDF version](#)

**ICT R&D performance across the world – patent-based evidence**

#### PREDICT 2013 REPORT: An Analysis of ICT R&D - EU & beyond

- [Online version of report](#)
- [Downloadable PDF version](#)

**ICT R&D performance across the world – patent-based evidence**

#### PREDICT 2012 REPORT: An Analysis of ICT R&D - EU & beyond

- [Online version of report](#)
- [Downloadable PDF version](#)

**ICT R&D performance across the world – patent-based evidence**
EIPE - Atlas of ICT activity in Europe

EIPE Composite Indicator
- ICT R&D Sub-indicator
- ICT Innovation Sub-indicator
- ICT Business Sub-indicator

Source: JRC-IPTS (EIPE Project)
We have accumulated data for quantitative analysis on many established data sources:
- Eurostat data mainly (Natl Accounts, LFS, SES, SRD, SBS, CIS, Surveys)
- EC databases, FP7, H2020, etc
- Other data sources (OECD, private providers, NSOs)

We aim at complementing existing statistics and analyses with new data sources and methodologies to track the digital transformations affecting the economy:
- by building new indicators hitting still unexplored dimensions,
- by better exploiting heterogeneous data
- getting more timely information

We still have many open questions..
GeoDIT

Geography of Digital Innovation and Technologies

✓ in the line of EIPE
✓ shifting toward the integration of new methodologies...
  ❖ Real time data indexing, retrieval and analysis
  ❖ Advanced visualisation / georepresentation
testing and applying technologies & methodologies

- on unstructured info from relational data
  - on patent related text information
  - on project abstracts and specifications
  - on related scientific production
- on other unstructured text resources
  - on firms related documents, IaD, etc

Text mining:
- POS Tagging
- Attribute detection
- Machine learning
- Thesaurus building

Semantic analysis

Pattern analysis & knowledge discovery, process mining

Network analysis
Transition toward BD

- adding other data sources
- investigating the dynamic networks

- of actors,
- technologies & products,
- and locations..

- and offering opportunities for different applications
We have accumulated data quantitative analysis on many established data sources:
- Eurostat data mainly (Natl Accounts, LFS, SES, SRD, SBS, CIS, Surveys)
- EC databases, FP7, H2020, etc
- Other data sources (OECD, private providers, NSOs)

We aim at complementing existing statistics and analyses with new data sources and methodologies to track the digital transformations affecting the economy
- by building new indicators hitting still unexplored dimensions,
- by better exploiting heterogeneous data
- getting more timely information

We still have many open questions..
### New data sources and data needs – Chair: A. de Panizza

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>09.45 – 10.00</td>
<td>Measuring the Digital Economy. Opportunities &amp; challenges from new data sources &amp; technologies - Jesus Vega Villa, JRC IPTS &amp; Giuditta De Prato, JRC IPTS</td>
</tr>
<tr>
<td>10.00 – 10.30</td>
<td>The EUROSTAT perspective - Albrecht Wirthmann, EUROSTAT</td>
</tr>
<tr>
<td>10.30 – 11.00</td>
<td>Big data in official statistics: ongoing work at Statistics Italy - Giulio Barcaroli, ISTAT &amp; Paolo Righi, ISTAT, IT</td>
</tr>
</tbody>
</table>

**11.00 – 11.30 Coffee**

### Opportunities for economic development – Chair: G. De Prato

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.30 – 12.15</td>
<td>Big Data for Development. New opportunities for emerging markets - Rohan Samarajiva, Board Chair LirneAsia, Colombo, Sri Lanka</td>
</tr>
<tr>
<td>12.15 – 13.00</td>
<td>The data isn't big enough. The story of the ever growing hollow haystack - Prasanna Lal Das, World Bank group, US</td>
</tr>
<tr>
<td>13.00 – 13.30</td>
<td>Data for development: an emerging opportunity - Nicolas De Cordes, Orange, FR</td>
</tr>
</tbody>
</table>

**13.30 – 14.30 Lunch break**

### Scanning industries and technologies – Chair: J.P. Simon

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.30 – 15.00</td>
<td>The challenges of web analytics - Giovanni Lorenzoni, Bitbang, IT</td>
</tr>
<tr>
<td>15.00 – 15.30</td>
<td>Beyond Bibliometrics: the view from the publishing industry - Stephane Berghmans, Elsevier, NL</td>
</tr>
<tr>
<td>15.30 – 16.00</td>
<td>Towards Data Market Places: Nature of Data, Exchange Mechanisms, Prices, Choices, Agents and Ecosystems - Renaud di Francesco, Sony Europe</td>
</tr>
</tbody>
</table>

**16.00 – 16.30 Coffee**

### Methodologies for economic analysis – Chair: G. De Prato

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.30 – 17.00</td>
<td>The Big data black box and the research agenda: applied economics in the Digital Data Era – Pablo de Pedraza, Webdatanet &amp; Univ. Amsterdam, NL</td>
</tr>
<tr>
<td>17.00 – 17.30</td>
<td>The Wageindicator experience – Stefano Visintin, Univ. Amsterdam &amp; Madrid, ES</td>
</tr>
</tbody>
</table>
Thank you!

Further information available at: http://is.jrc.ec.europa.eu/pages/Homepage.html

giuditta.de-prato@ec.europa.eu