Data Value Chain in Europe

FedCSIS, 09.09.2013, Kraków

European Commission, DG *CONNECT*
Márta Nagy-Rothengass, Head of *Data Value Chain Unit*

OUTLINE

1. Data value chain – Why does it matter?
2. Virtuous cycle of data
3. Future challenges and trends
4. Conclusion

EUROPEAN COMMISSION

The Communication Networks, Content & Technology Directorate General

Directorate Media and Data

G1: Converging Media and Content
G2: Creativity

G3: **Data Value Chain**

G4: Inclusion, Skills and Youth
G5: Administration and Finance

"DATA Value Chain" Vision

*aims at **extracting the maximum value** from data by building on the intelligent use of multilingual **data sources** across Europe.*

This will lead to

*- **increased business intelligence** and **efficiency** of private and public sectors*

*- **world class applications and services***

***- new business opportunities** involving SMEs*

*- new ways of tackling **societal challenges***

"Data value chain" in ICT policy context

- ***A strategy for smart, sustainable and inclusive growth***
- ***A vision to achieve high levels of employment, a low carbon economy, productivity and social cohesion, to be implemented through concrete actions at EU and national levels.***

- *One of the seven flagship initiatives of Europe 2020, set out to* ***define the key enabling role that the use of ICTs will have to play*** *if Europe wants to succeed in its ambitions for 2020.*
- ***The overall aim […] is to deliver sustainable economic and social benefits from a digital single market […]***
- ***Support research and innovation***

"The Commission is invited to ***make rapid progress in key areas of the digital economy to ensure the creation of the Digital Single Market by 2015, including […] the availability of public sector Information."***

Conclusions of the European Council (4 February 2011)

Main pillars of the "Data value chain" thinking

1. Creation of "data value" friendly **policy environment**

2. Development of **European Digital Service Infrastructure** (CEF) and **fostering new services** in relation to
   - Building and reuse of language resources
   - Open Data portals at local, regional, national and European level

3. **European Research and innovation support**

Questions around the Data Value Chain

Who/what is producing data in the EU?

How is data captured?

How is data shared?

How is data reused?

What is the relevant maturity of these technologies and processes?

How deep and diverse is the EU data supply chain?

How can EU help?

Where "data" comes from

Humans (data volumes growing slowly)

- **Professional**
- **Personal/social**

Machines (data volumes growing fast)

- **Earth observation (satellite/env monitoring)**
- **Scientific equipment (DNA sequencers, telescopes, particle accelerators, …)**
- **Transport (air, trains, ships, cars, …)**
- **Sensors (Internet of Things, RFID, Smart Cities)**
- **Industrial machinery**
- **Robots (UAVs, smart spaces)**

Data has become a resource

- **Types:**

Publicly funded data (like statistics, environmental and geographical data, meteorological data, business information, legal information), institutional and private data

- **Use and reuse**: Combination of different types of data (e.g. geo, traffic and tourism; business and open)

- Data activities generate **externalities**, positive (reuse) and negative (privacy): capture the good and avoid the bad

- **competitive advantage is** to offer the right data to the right people at the right time

"Big data": What is at stake?

- **3 V's** (volume, velocity, variety)

- **Value-generation** for the economy and society

- Smart, sustainable and inclusive **growth** and the creation of **jobs** in Europe

- **Automatized** data-generation, data-processing and value extraction: A revolution that will change society?

What are the challenges at EU level?

- Fostering a data-friendly **legal and policy environment**
- Developing an **efficient European data ecosystem**
- Supporting **competence** in the area of data
- Enhancing **interoperability**
- **Multilingualism**….

Main challenges for data business?

- Growth of organisational information, large and growing data quantity; quality?
- Multimodal information
- Unstructured data
- Heterogeneity of data and data sources
- Complexity
- Interoperability
- Enriching business data with open data
- External shocks, e.g. financial crisis

Decision makers rely on countless heterogeneous complex ICT systems …

Finance

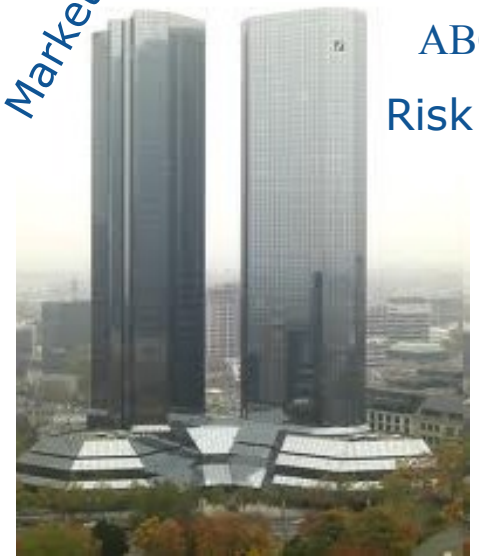
Stock markets

Accounting

Taxes *GL*

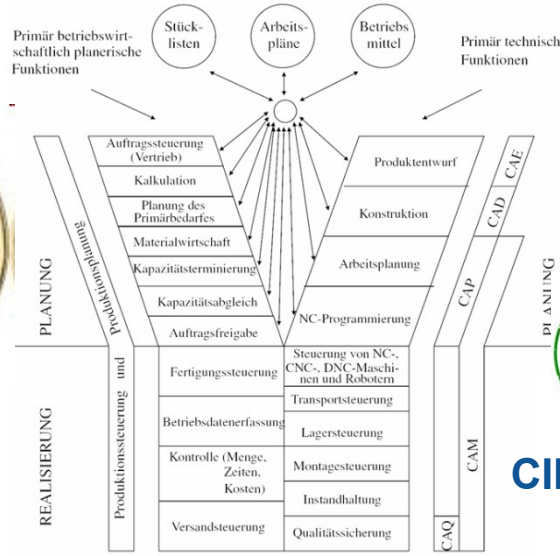
Cash-flow

Marketing



ABC / ABM

Risk analysis



PERT-CPM

Human resources

MTM

Assets management

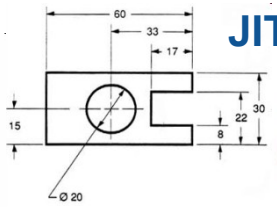
CRM

DRM

Services

Bank accounts

Industry



JIT

MPS

MRP

BOM

ERP

SPC

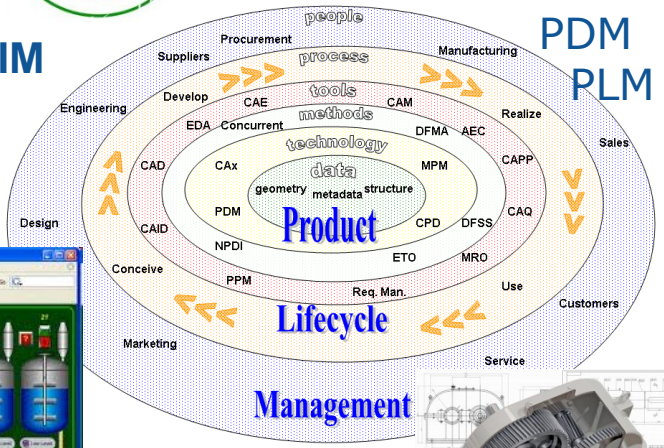
TQM

CIM

PDP

PDM

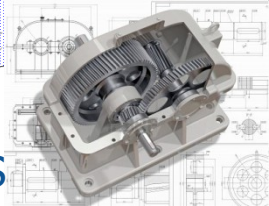
PLM



SCADA

STEP XPDL

OEE CMMS



… but structured information represents only a small fraction of the whole picture

15%

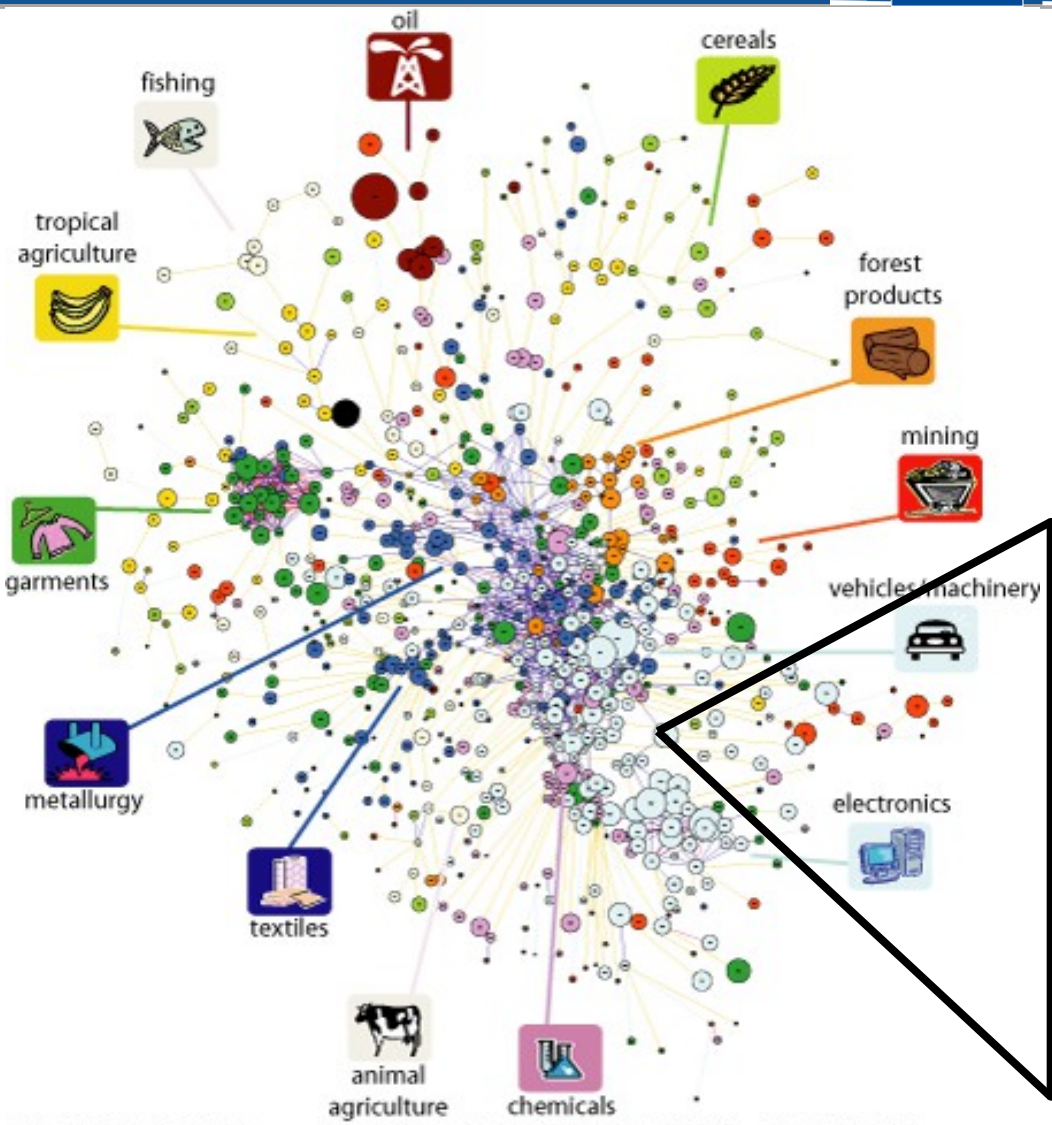
85%

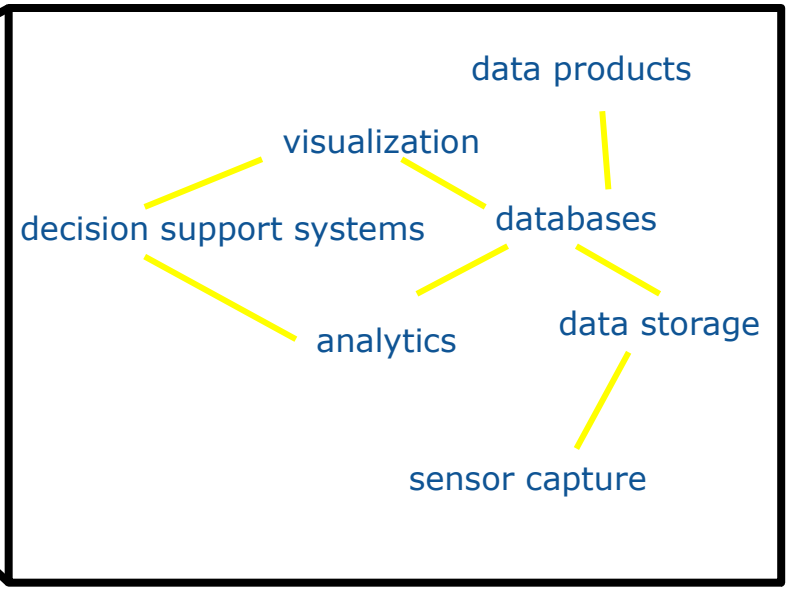
More than **85%** of all business **valuable information** exists in the form of e-mails, memos, notes from call-centres, news, user groups, chats, reports, web-pages, presentations, image-files, video-files, marketing material and news.

Source: Merrill Lynch

Data Supply Chain and Product Spaces

Need a high value added data cluster





Towards a virtuous cycle

Creation of "data value chain friendly" policy environment (1)

- Fostering of (Open) **Data policy**

- Adoption of the
  - revised Directive on the **re-use of Public Sector Information** (PSI)
  - **Commission decision on re-use of its own information**

- **Implementation of PSI** policy across Europe by ensuring compliance and the development of **soft law instruments** (e.g. guidelines on licensing and charging)

- **Stakeholder** involvement and engagement

Open (public) data: Why does it matter for Europe?

1. **Untapped business and economic opportunities**: data is the new gold; possible direct and indirect gains of €140bln across the EU27; Dutch geo-sector in 2008: 15.000 jobs

2. **Better governance and citizen empowerment**: open data increases transparency, citizen participation and administrative efficiency and accountability

3. **Addressing societal challenges**: data can enhance sustainability of health care systems; essential for tackling environmental challenges

4. **Accelerating scientific progress**: e-science essential for meeting the challenges of the 21st century in scientific discovery and learning.

Multilingual (Open) Data infrastructure (2)

- Development of **European Digital Service Infrastructure** and **fostering new services and opportunities** in relation to

  - Open Data portals at local, regional and national and European level

  - "Multilingual access to online services"

  - Leading by **best practice examples** …

… the *European data portals*

- ***European Commission data portal*** *(2012)*

- ***Pan-European open data portal*** *(2013)*

  Multilingual access point to data from across the EU, can be extended to **European Open Data Infrastructure** through **CEF** (Connecting Europe Facility)

- ***Benefits***
  - Scale
  - Interoperability of datasets
  - Easy to find across languages
  - Similar basic use conditions

Supporting Research and innovation (3)

which fosters the intelligent use, management and reuse of **complex and large amount of data**

for

- better decision making
- efficiency
- knowledge management
- extraction of embedded intelligence and data insights.

including

- R&D in Multilingual data and content analytics
- Innovation in Data driven intelligence and knowledge management in data intensive sectors

Main related research challenges in FP7

(5 calls; >80 projects; >600 contractors; 1500 FTE)

- Content creation & processing (including multimedia and games)
- Integration of social software & semantics
- Personalisation & summarisation
- Semantic foundations
- **Reasoning (temporal, dimensional and uncertainty, approximate & incomplete reasoning)**
- **Knowledge management in business & public-interest domains**
- **Coping with data explosion** ("big data" + real-time)

Financing and support measures of *R&D&I* enhancing new data-handling technologies:

- 2011-2013: ~ € 100 million
- **Special cluster**: Environmental and geographical data driven projects (eContentplus, CIP and FP7) including support of community building and exchange
- **Related calls of 2013:** 48 projects under negotiation

  FP7 SME - Digital Content Analytics

  FP7 ICT-2013.4.2 - Scalable data analytics

  CIP 2013: support for ***technology innovation and uptake*** (pilot actions, testing, showcasing innovative applications)

+ priority area envisaged for ICT <u>in Horizon 2020</u> (2014-2020)

FP7 data market experiment:
(SME call funding of 20 projects out of 344 proposals)

Key themes :

- **Bootstrapping a data economy**
- Community building and best practices
- **Sharing language resources**
- Building consensus and common services

Key dimensions:

data pooling for new service

\+ *focus on SME participation*

R+D in a nut shell:

Application domains:

Business sector
- **Marketing**
- **Logistics**
- **Product dev.**
- **Financial dev.**
- **…**

Personal and social sphere
- Entertainment
- Personal applications
- Social networks
- …

Public services
- Health
- Education
- Culture
- Emergency management
- …

Societal challenges
- Science
- Transport
- Environment
- Smart cities
- GIS
- …

Technology challenges:
- Big data analytics, scalability
- Semantics and reasoning; Collaboration tools
- Language Technologies (Speech recognition and dialogue systems, Machine translation, Analytics)
- Multimedia and multimodal content

Outlook: Horizon 2020

*Next Research, Development and **Innovation** programme of EU 2014-2020*
http://ec.europa.eu/research/horizon2020/index_en.cfm

- Excellence in Science

- Competitive Industries: leadership in enabling and industrial technologies (LEIT)

- Better Society: Societal challenges

***H2020: Competitive Industries**:*

*build leadership in enabling and industrial technologies, with **dedicated support for ICT**, nanotechnologies, advanced materials, biotechnology, advanced manufacturing and processing, and space, while also providing support for **cross-cutting actions** to capture the accumulated benefits from **combining several Key Enabling Technologies**;*

+ facilitate access to risk finance;

+ provide Union wide support for innovation in SMEs.

Stakeholder meeting point

European Data Forum 2013

http://2013.data-forum.eu/

INFO

European Data Forum 2014

19.-20.03.2014 in Athens

http://2014.data-forum.eu/

Call for contributions

european data forum

Conclusion

- Big data creates fantastic **opportunities** for new business but also some **threats** (Information overabundance, privacy)
- The business champions of the future will be the most successful companies in **coping with data flood and DVC**
- Yet **technology solutions lags far behind** the complexity of information problems
- **Legal issues** (privacy, licensing, reuse) needs European actions
- Towards a **better use of publicly funded data** in Europe
- The **EC is committed to support** to improve European data driven competitiveness

Further info

- ***ICT under FP7***
  http://cordis.europa.eu/fp7/ict/

- ***Experts data base:***
  https://cordis.europa.eu/emmfp7/

- ***Unit*** *– Data Value Chain (DG CONNECT G3)*

  URL*: http://cordis.europa.eu/info-management/*
  http://cordis.europa.eu/fp7/ict/language-technologies/

  eMail to**: cnect-g3@ec.europa.eu**

Thank you!

Re-use of open government data

- *Publicly funded data must be available for all*
- *What data are we talking about?*
  Statistics, environmental and geographical data, meteorological data, business information, legal information
- *Combination of different types of data (e.g. geo, traffic and tourism)*
- *EU-wide applications and services*
  Capitalise on the size of the internal market
- *Systems that facilitate decision making by companies*

Revised Directive enables easier re-use of PSI

- Creation of a **genuine right to re-use public data**: all public data not covered by an exception is to be re-usable
- Invitation for public bodies to make their **documents and data** available in a **machine-readable format** and together **with their metadata** where possible and appropriate
- **Charges** shall be **limited to the marginal costs of reproduction and dissemination** (does not apply to cultural institutions)**; Burden of proving compliance** with charging rules shifts to public bodies
- **Independent supervision (at MS level) of application of the rules is required**

CEF digital service infrastructure for data

- ***"Core service platform"***
  - Distributed system
  - Query and visualization tools
  - Open source
  - Governance model involving the data providers

- ***"Generic services"***
  - Aggregation of datasets
  - Interoperability of datasets
  - Interface to open data infrastructures in third countries
  - Data repositories and long-term preservation services