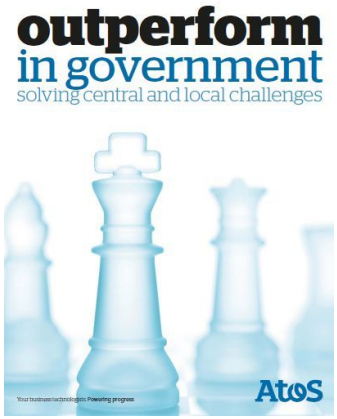
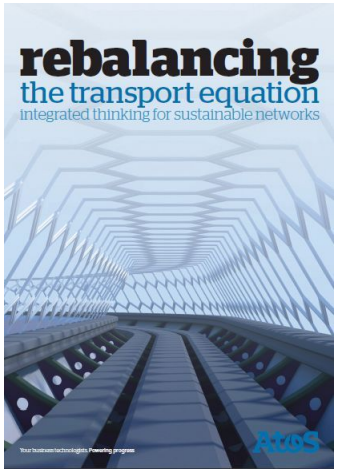
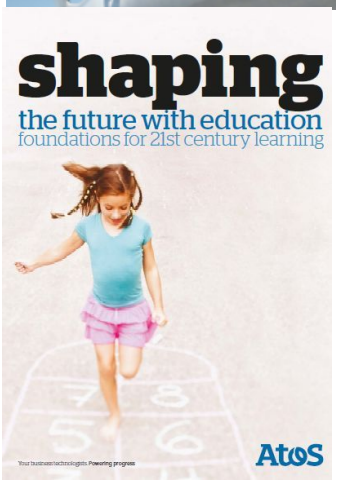




Text Analytics and Big Data

META-FORUM 2012
Brussels, 20th June 2012

Tomás Pariente
Atos Research & Innovation



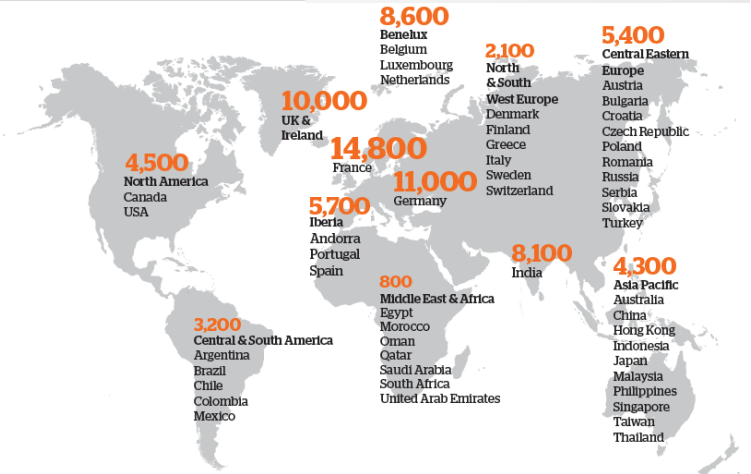
Atos
Worldwide IT Partner

Table of Contents

Atos:
The company

20/06/2012
Tomás Pariente

Atos is an international **information technology services company**,
delivering **hi-tech transactional services**, consulting, systems integration and managed services.



- Annual revenues of € 8,7 billion (pro-forma 2010).
- Over 78,500 business technologists worldwide in 42 countries.
- Worldwide headquarters in Bezons / Paris, France.
- Established on July 2011 with the integration of Atos Origin and Siemens IT Solutions and Services

Atos:
From critical IT to business support

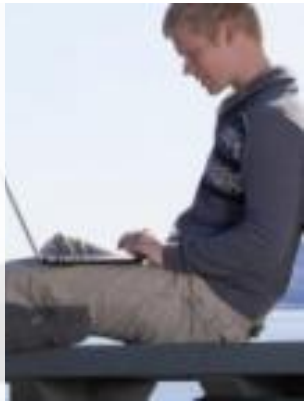
20/06/2012
Tomás Pariente

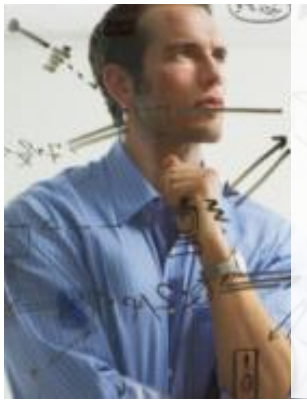
Consulting & Technology Services

Transforming business through **innovation** leveraging Information Technolo

Delivering seamless business systems

Systems Integration





Transactional Services

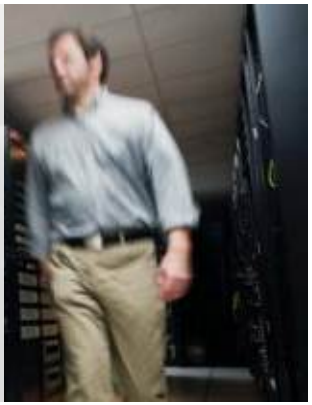
Advancing busin strategy through **innovative** processing solutions

Innovation

forming IT infrastructure and business operations

revenues

Managed Services



Atos:
From vision to Innovation

20/06/2012
Tomás Pariente

Innovation is in our company DNA

Innovation can be seen everywhere and particularly:

- in our set of Innovative Offerings growing over 10% a year.
- in our innovation policy involving every single employee's ideas through our social platforms.
- in the works of our scientific community actively upgrading our vision of the technologies for the future.
- in **the Research and Innovation group of Atos,** applying the latest research results to opportunities where clients need solutions that go beyond market's offer.

This is how we are able to constantly propose new solutions and ideas to our clients, and to help them invent their own new way towards growth and profitability.

Atos Research & Innovation Main Activities

20/06/2012
Tomás Pariente

- Atos R&D hub
- Involved in projects that combine advanced technological developments with commercial exploitation
- Usually working in international partnerships
- A source of innovative ideas to be used by the company's sales and technical staff

R&D&I since 1987

180 ongoing Projects

Of which, 105 EU R&D&I

Research »

- Project management
- Technical consulting
- Technology development
- Market studies & project results exploitation plans

R&D &I

Innovation »

- Bring research outcomes to Atos' customers
- Support the sales force as "technical backstage"
- Business development

Consulting »

- Strategic R&D&I consulting
- Emerging technologies expertise
- Technology watch

Atos Research & Innovation
Why we are here: Knowledge Lab

- Main capabilities:
  - Semantic Technologies
  - Language Technologies
  - Big Data
  - Linked Data

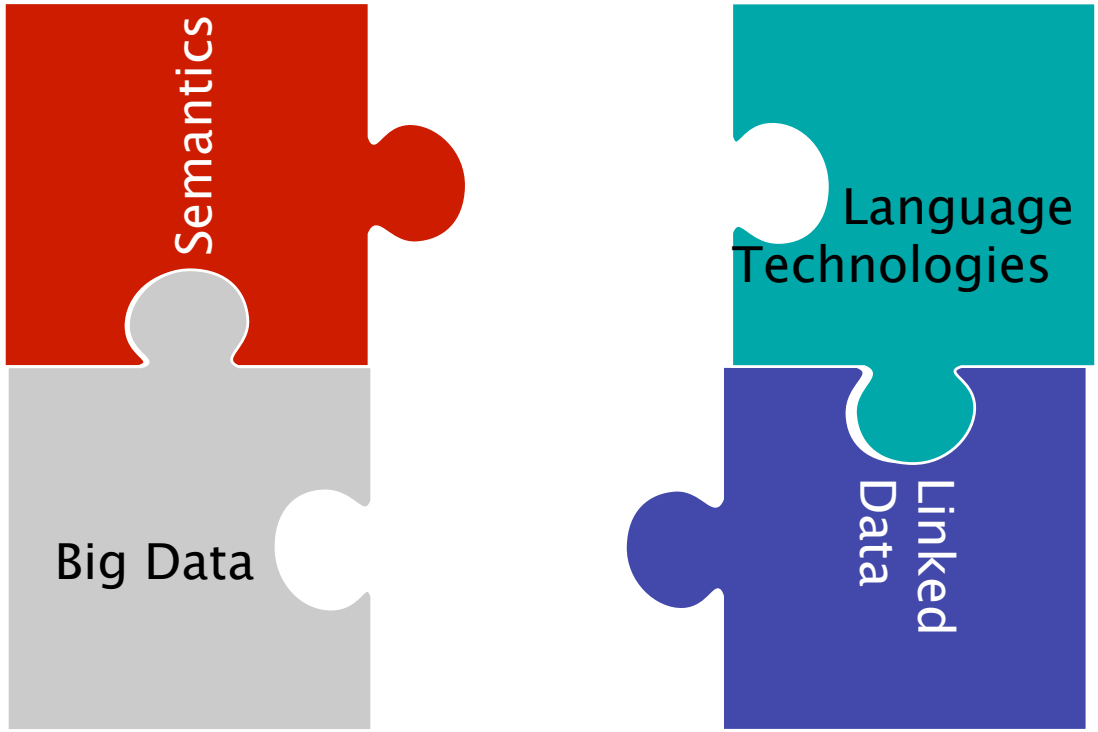
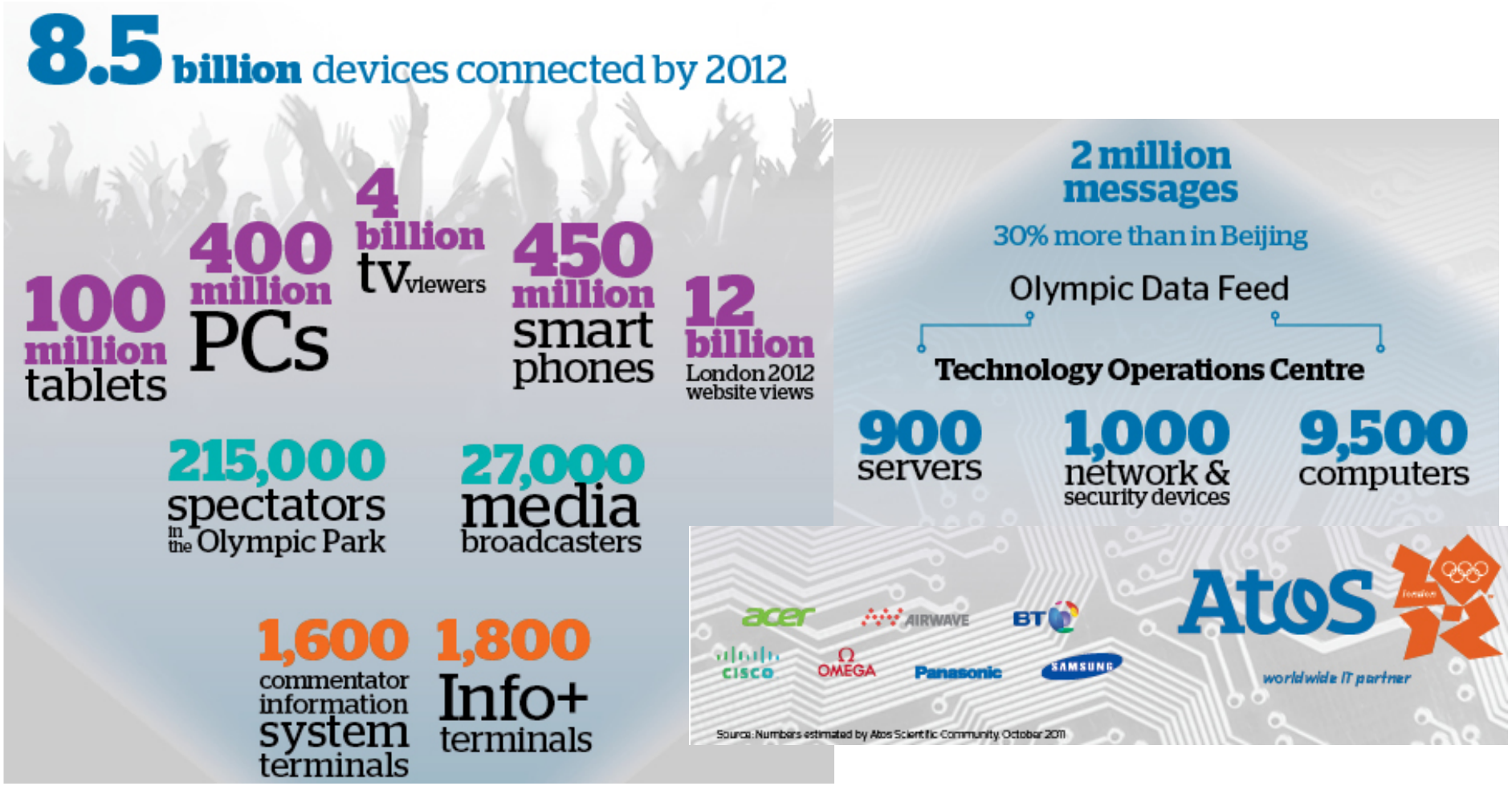


Table of Contents

A use case in Atos: Olympic Games
(increasing demand of data processing, storage and innovative applications)

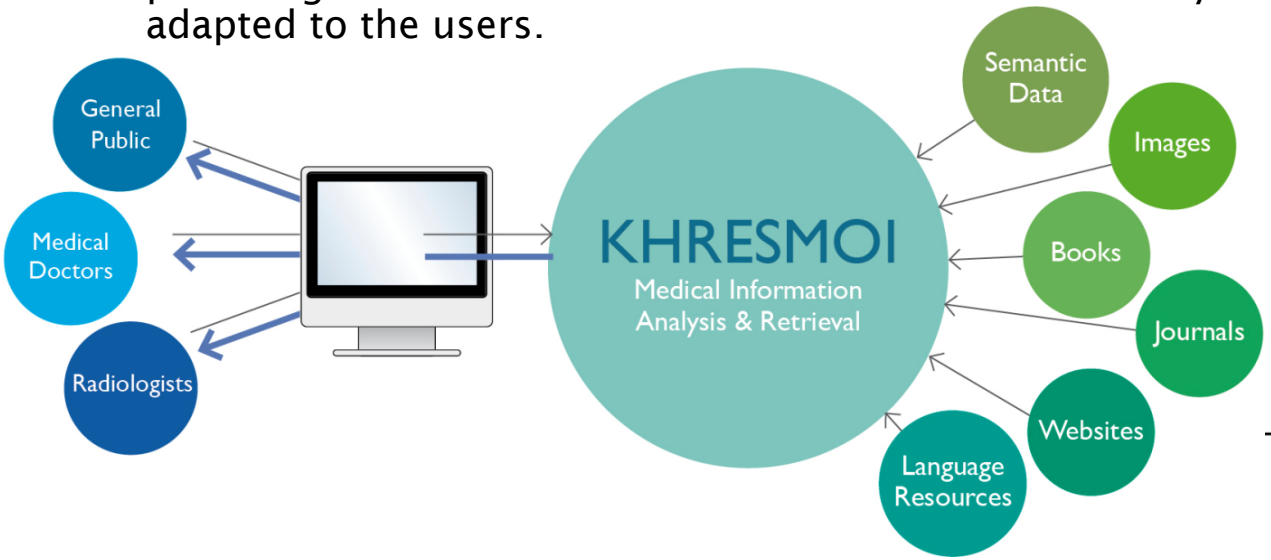
20/06/2012
Tomás Pariente



Khresmoi Project

20/06/2012
Tomás Pariente

- KHRESMOI aims to build a **multi-lingual**, **multi-modal search** and access system **for biomedical information** and documents. The system will allow access to biomedical data:
  - from many sources,
  - analyzing and indexing multi-dimensional (2D, 3D, 4D) medical images,
  - with improved search capabilities due to the integration of technologies to link the texts and images to facts in a knowledge base,
  - **in a multi-lingual environment,**
  - providing trustable results at a level of understandability adapted to the users.



Integrated Project

Sep 2010 - Aug 2014

http://www.khresmoi.eu/

A growing universe of unstructured data

… created and accessible from almost everywhere

A growing universe of unstructured data

FIRST Project

*is to **make available the relevant information***
*of the **financial information space***
(including unreliable, unstructured, sentiment sources)
to the decision maker in near-real time
in an automated way

http://project-first.eu

FIRST
Mining the Web for financial texts

20/06/2012
Tomás Pariente

Some numbers
- 176 Web sites
- 2,671 RSS sources
- ~40,000 documents per day
- >8,000,000 documents and growing
- ~13,000 tweets per weekday, for ~1,800 NASDAQ/NYSE stocks ($GOOG, $MSFT…)

Natural Language preprocessing and entity extraction

Streaming

CONTENT

Cleaning

HTML

Financial terms, Companies, Instruments …

Sentiment

DSS

About Financial objects

Table of Contents

BIG: Key facts

- Type of project: Coordination Action (CA)
- Duration: 26 months
- August 2012 – October 2014
- Budget: 3,055 Meuro
- Funding: 2,5 Meuro
- Consortium: 11 partners

Overall objective

Address **technical**, **business** and **policy** aspects of **Big Data** with the aims of shaping the future of the area, positioning it in **H2020** and bringing the necessary **stakeholders** into a self-sustainable **industrially-led initiative** to **enhance EU competitiveness** taking full advantage of Big Data.

Project objectives

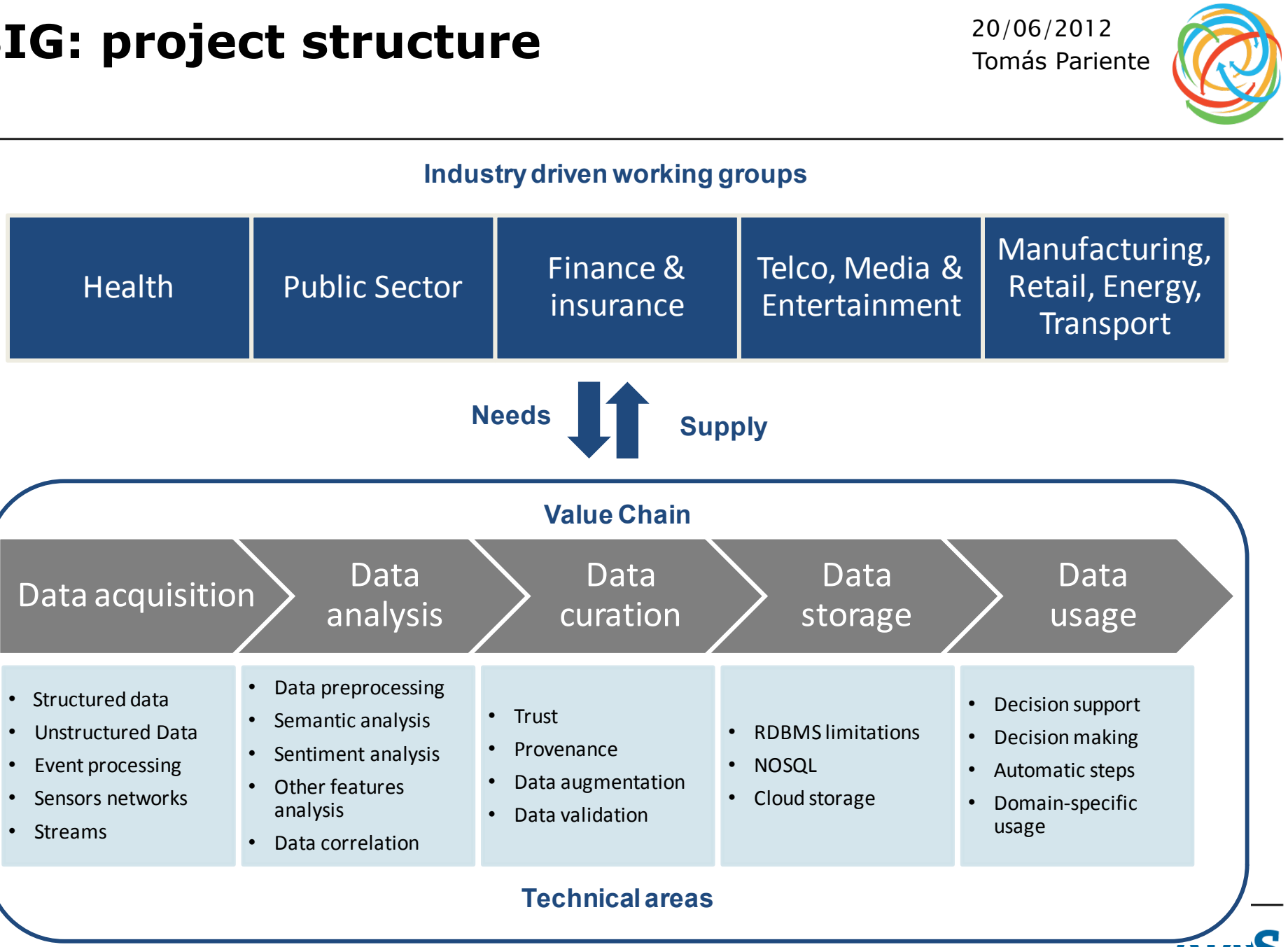
- **Main Missions**
  1. Build a self-sustainable Industrial community around Big Data in Europe
     - **Technical level** establishing the proper channels to gather information
     - **industrially-led initiative** to influence adequately the decision makers
  2. Promote adoption of earlier waves of big data technology
  3. Tackle adequately existing barriers such as policy and regulation issues

- **Concrete Objectives** (and outputs from BIG project)
  - **Define Stakeholders and players** in the value chain (D2.3 Sector's Requisites).
  - Elaborate **a clear picture of existing technological trends and their maturity** (D2.2 Technical white papers)
  - Acquire a **sharp understanding of how big data can be applied** to concrete environments/sectors (D2. 4 Sector's Roadmap)
  - **Disseminate results and involve** different stakeholders (D3.4 Project Dissemination Reports and D3.5 Stakeholder engagement activities)
  - **Define priorities based on expected impact** (D.2.5 Integrated Roadmap)
  - **Contribute to EU competitiveness and position** it in Horizon 2020 (D4.2 IPR, Standardization Recommendations)

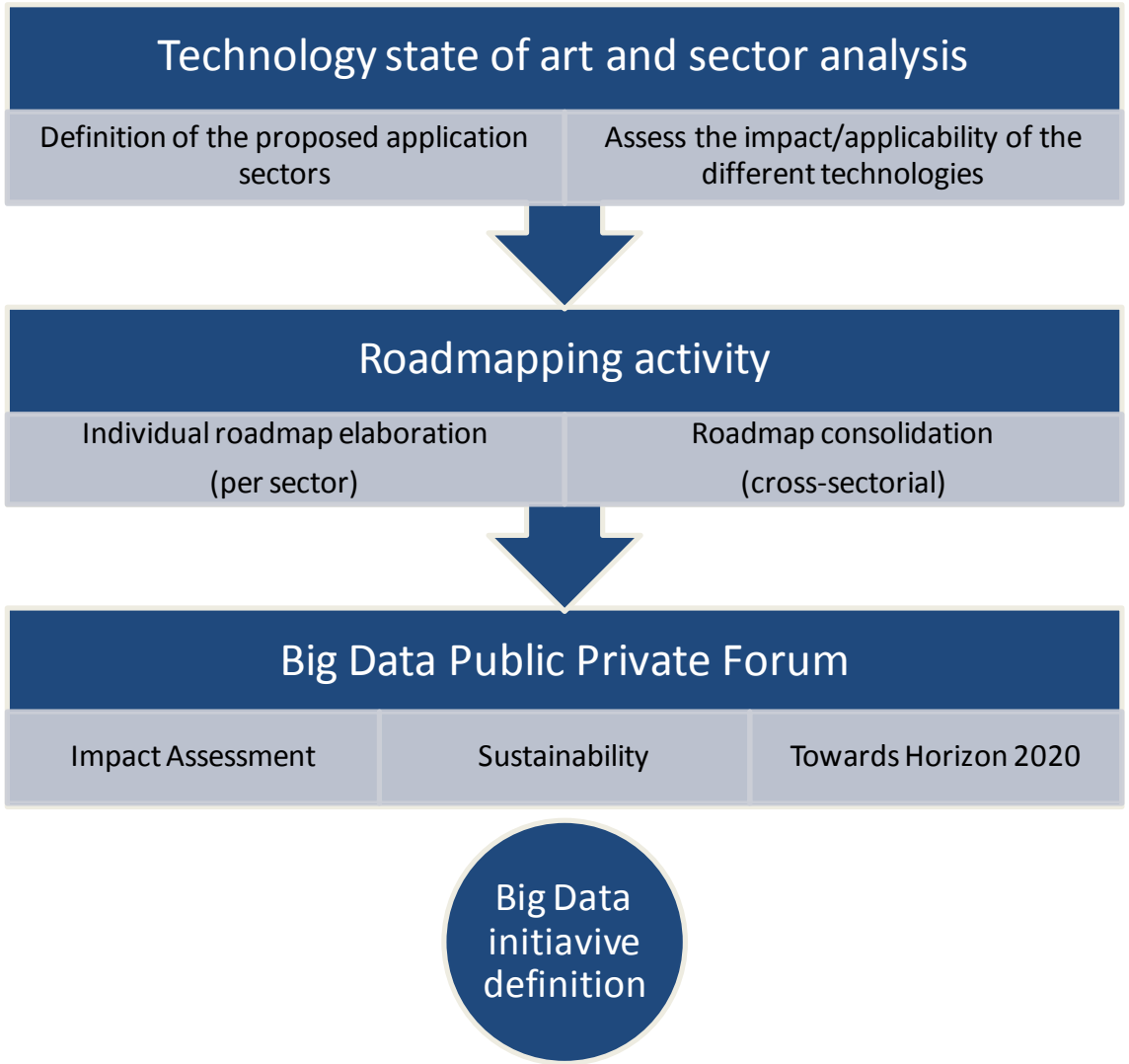
BIG: approach (I)

- Not only technology, but also **business**, policy and regulation;
- Not only generic plans for research, **but specific plans for adoption** for those sectors that are positioned for greater gains from the use of Big Data;
- Not only theoretical activities including roadmaps, coordination and dissemination aiming at future actions, but also **actions in the course of the project** to foster understanding and adoption of current technology solutions;
- Not only development activities in the a limited timeframe (the duration of the project), but the creation of an operational framework (including stakeholder engagement and leadership, organizational structures and technical infrastructure) as a starting point for future work that will go **beyond the project duration**

BIG: project structure



BIG: major activities



BIG Philosophy

- Presence of the **right profiles** in the consortium to drive this process to the level of influence and impact we are aiming for
- Involve **stakeholders from EU industry** and officers at the right level
- An **open philosophy** will be applied to all the documents generated by the project, which will be made public to a **wider community for active contribution and content validation.**
- BIG is by nature a **cross-disciplinary** initiative with many angles.
- Reach a coherent and sensible result that satisfies the research community and high level decision makers at the same time. Thus a **top-down and bottom-up approach** have been defined.

BIG: methodology

What is there (in terms of technology)?	Which is the level of maturity?	Which support actions are needed?	Which metrics should be used?	Which are the impacts ?
What is needed (domain requirements)?	Can it be implemented?	How can it be done?	Which is the actual situation?	What are the residual challenges?
What benefits will it bring to the stakeholders?	What is the time to market?	When can it be done?	What do we want to achieve?	Highlight barriers, strengths, future directions…
	Is there any kind of restrictions?	Links between topics (technology/ sectors needs)?	How to fulfill the existing gap (cost/ timeframe…)?	

Atos Research & Innovation

Tomás Pariente

Thank you

For more information please contact:

tomas.parientelobo@atosresearch.eu

6/26/12