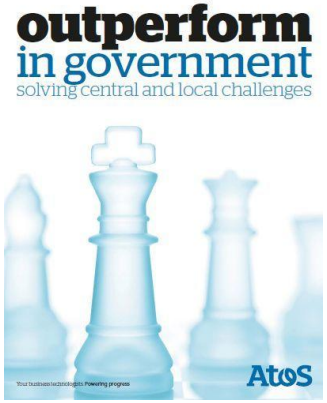
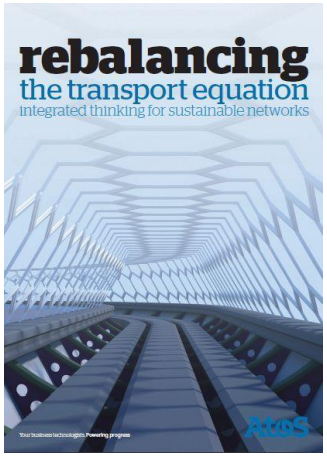
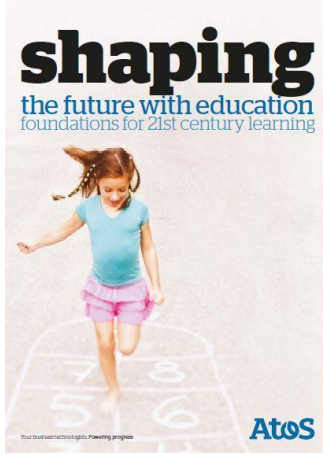




BIG
(Big Data Public Private Forum)

European Data Forum
Copenhagen, 7th June 2012



Nuria de Lama

Representative of Atos Research & Innovation to the EC

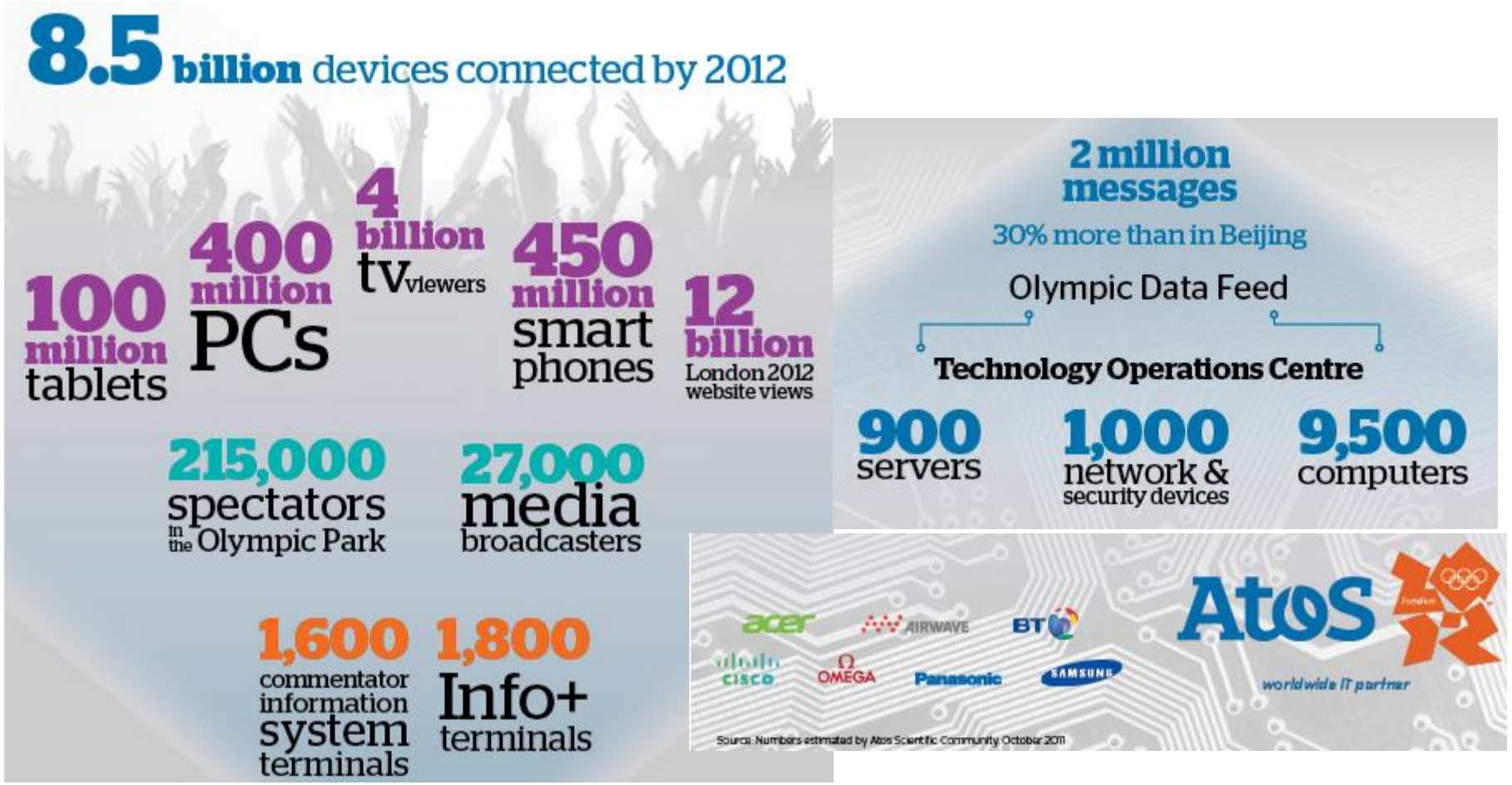
Atos
Worldwide IT Partner

Table of Contents

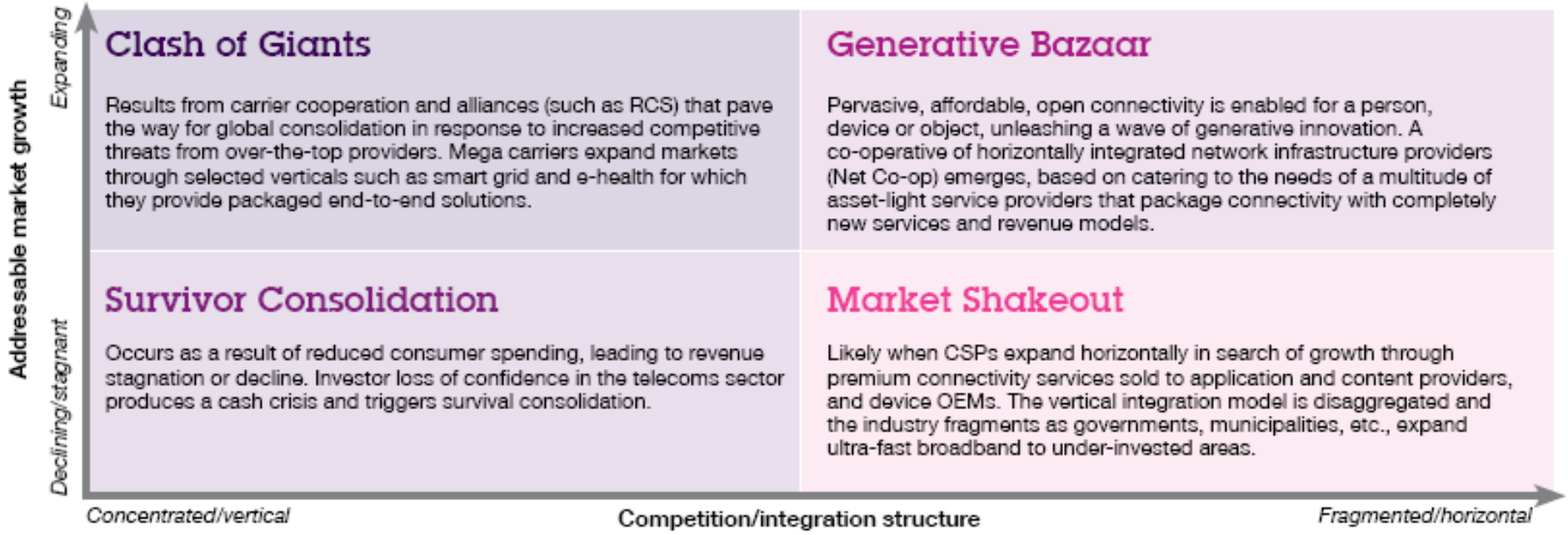
Table of Contents

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

07/06/2012
Nuria de Lama



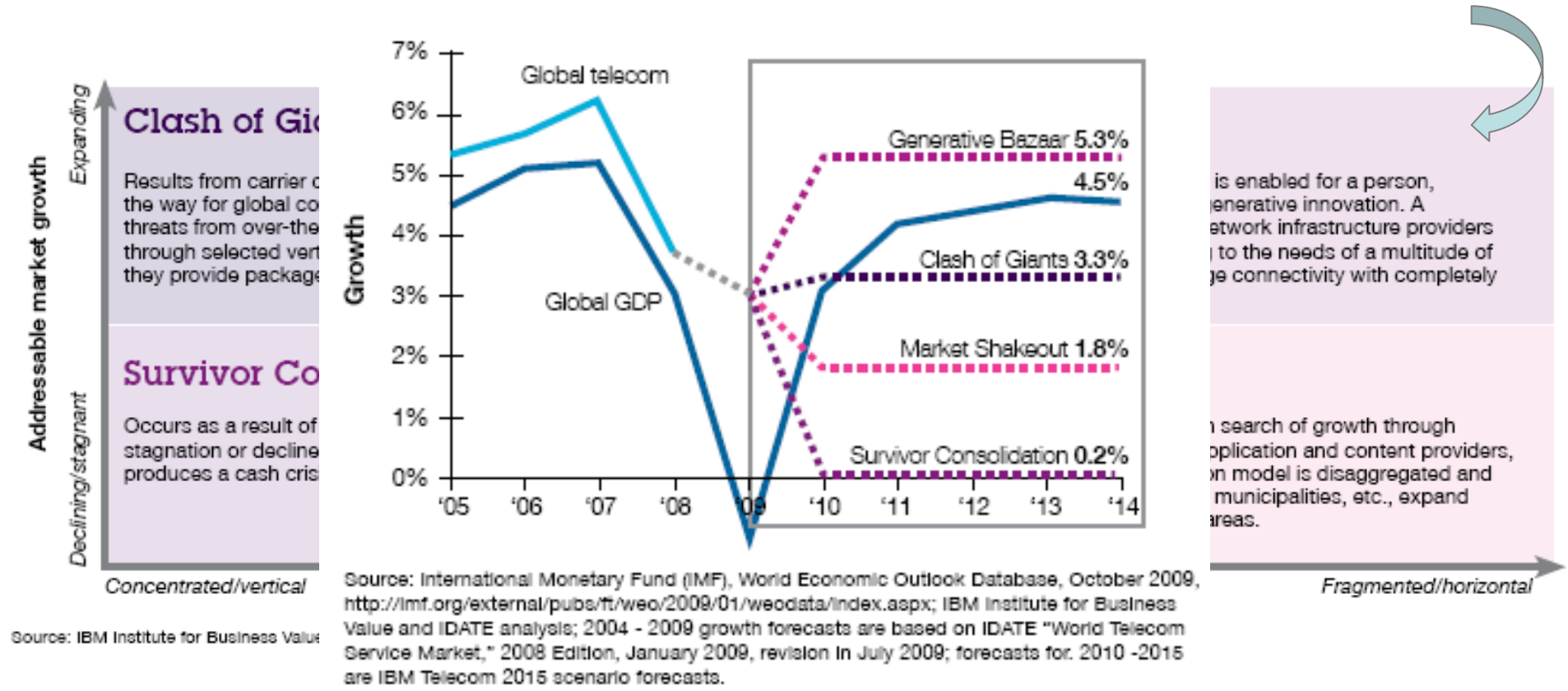
Business potential of openness and collaboration



Source: IBM Institute for Business Value analysis.

Business potential of openness and collaboration

07/06/2012
Nuria de Lama

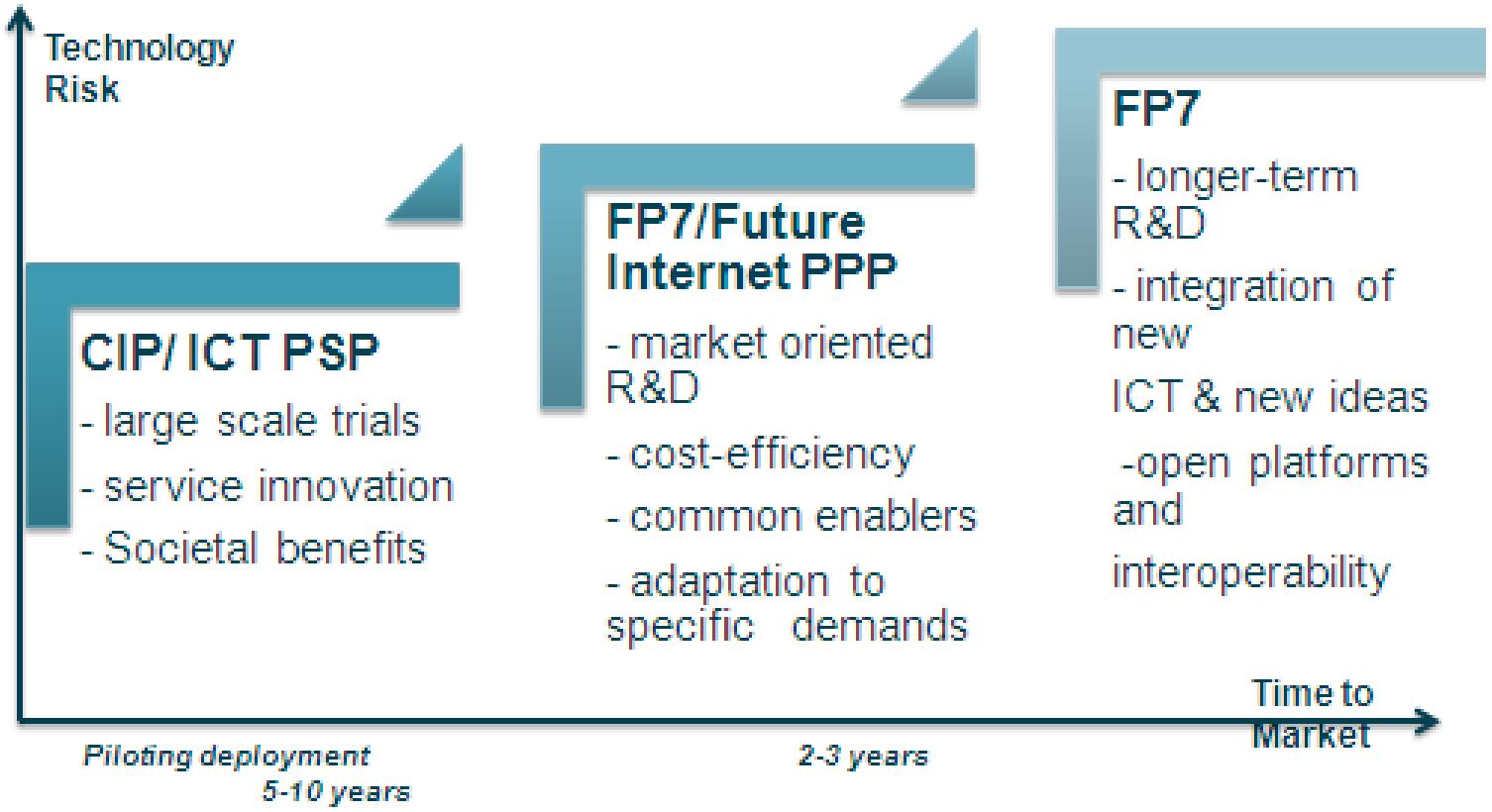


Source: IBM Institute for Business Value

Source: International Monetary Fund (IMF), World Economic Outlook Database, October 2009, http://imf.org/external/pubs/ft/weo/2009/01/weodata/index.aspx; IBM Institute for Business Value and IDATE analysis; 2004 - 2009 growth forecasts are based on IDATE "World Telecom Service Market," 2008 Edition, January 2009, revision in July 2009; forecasts for 2010 -2015 are IBM Telecom 2015 scenario forecasts.

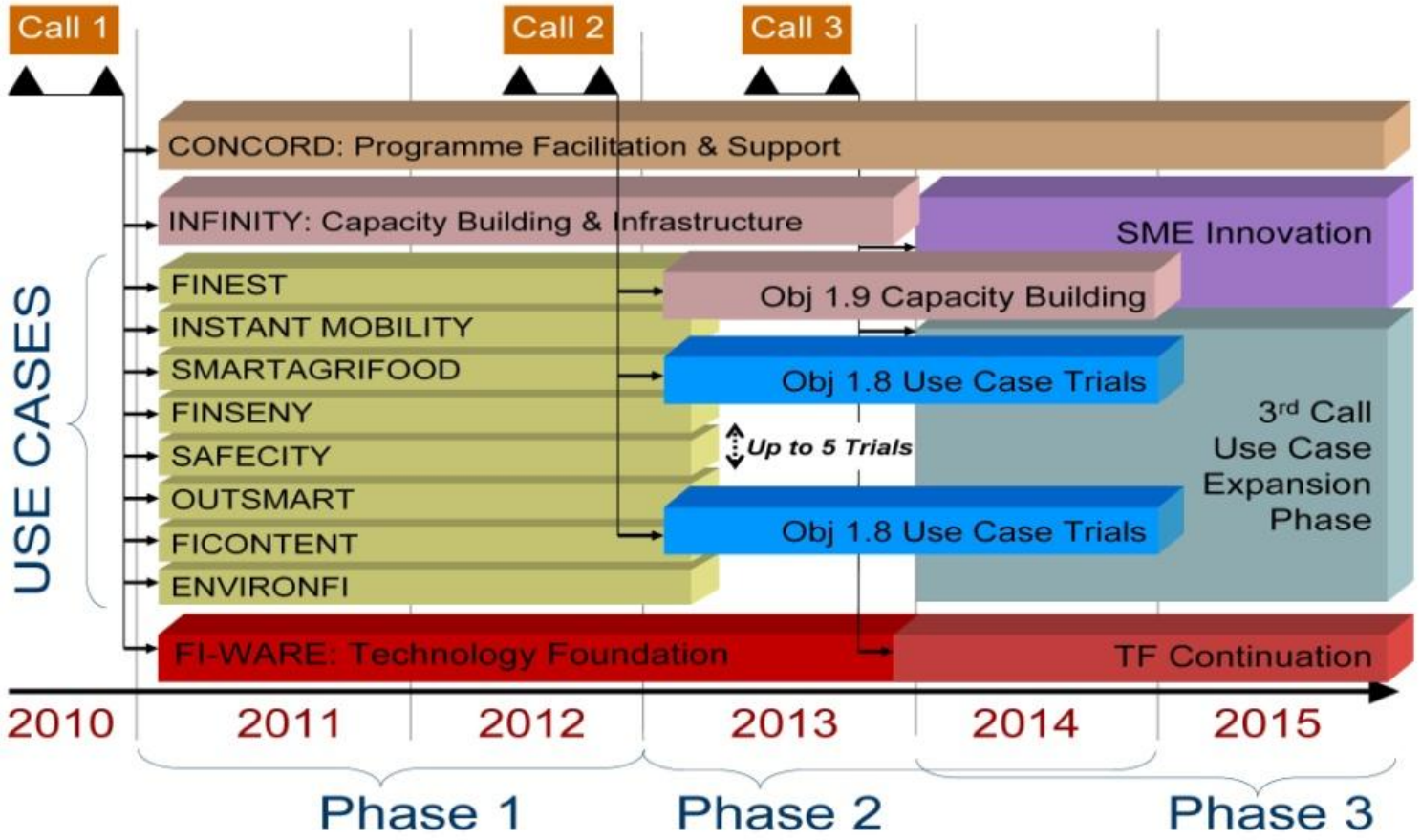
Table of Contents

The FI PPP: Towards an innovation landscape



FI PPP Programme Implementation: Technology aligned with needs

07/06/2012
Nuria de Lama



9

FI-WARE: Some figures and data

Main data

26 partners
5 Universities
4248 Person Months
(excl. open calls)

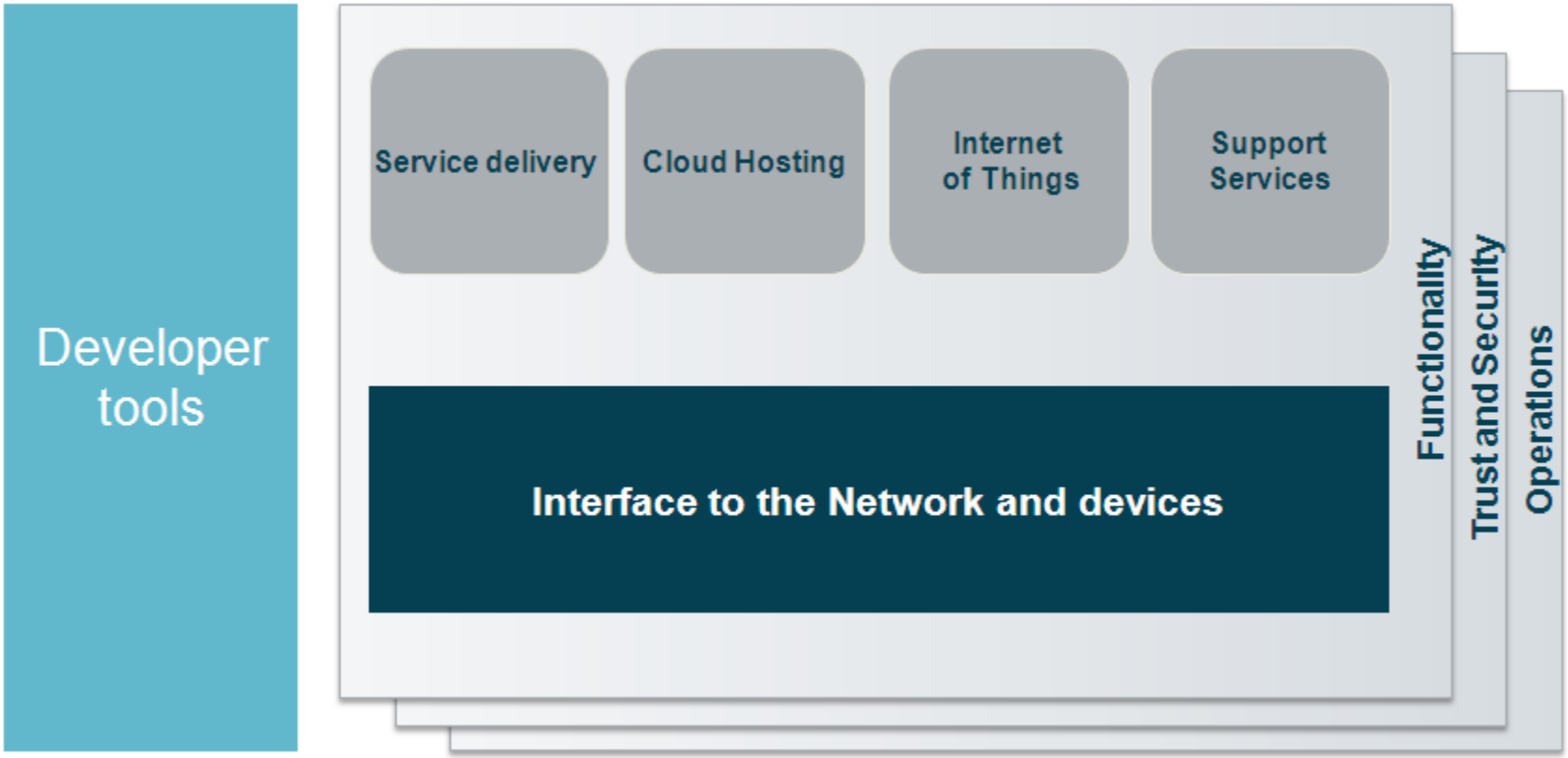
Total Funding 41 M€
Open calls 12,3 M€
Total budget 66,4 M€

Three years duration

FI-WARE: Collaboration with Usage Areas

07/06/2012
Nuria de Lama

FI Core Platform Architecture: main chapters



FI-WARE and Big Data

07/06/2012
Nuria de Lama

- The FI-WARE project addresses, among others, the following Data Management topics:
  - Publish/ Subscription
  - Complex Even Processing (CEP)
  - Multimedia analysis
  - Unstructured data analysis
  - Meta-data pre-processing
  - Semantic annotation and application support
  - **Big Data analysis**, based on
    - Hadoop (heterogeneous MapReduce platform mainly used for **ad-hoc data exploration of large sets of data**.)
    - MongoDB: a scalable, high-performance, open source, document-oriented database.
    - SAMSON Platform: high-performance streaming MapReduce platform that is used for the **near-real time analysis of streaming data**.

Table of Contents

BIG: Key facts

- Type of project: Coordination Action (CA)
- Duration: 24 months
- Budget: 3,055 Meuro
- Funding: 2,5 Meuro
- Consortium: 11 partners

Overall objective

Address technical, business and policy aspects of IIM and 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

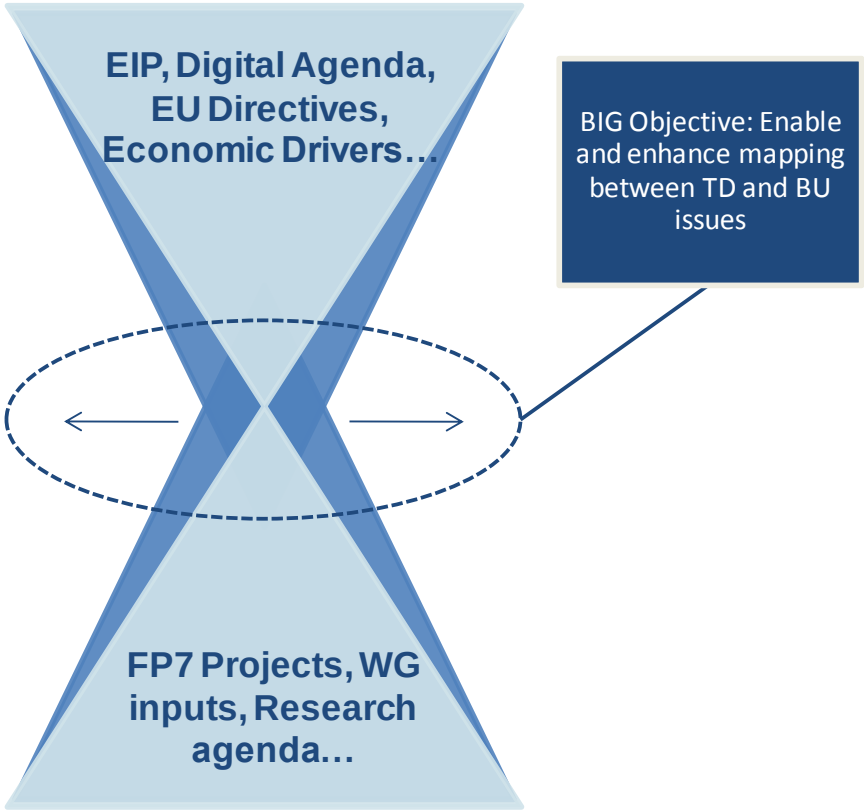
- **Concrete Objectives** (and outputs from BIG project)
  - **Define Stakeholders and players** in the value chain (D2.3 Sector's Requisites).
  - Elaborate **a clear picture of existing technological trends and their maturity** (D2.2 Technical white papers)
  - Acquire a **sharp understanding of how 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)

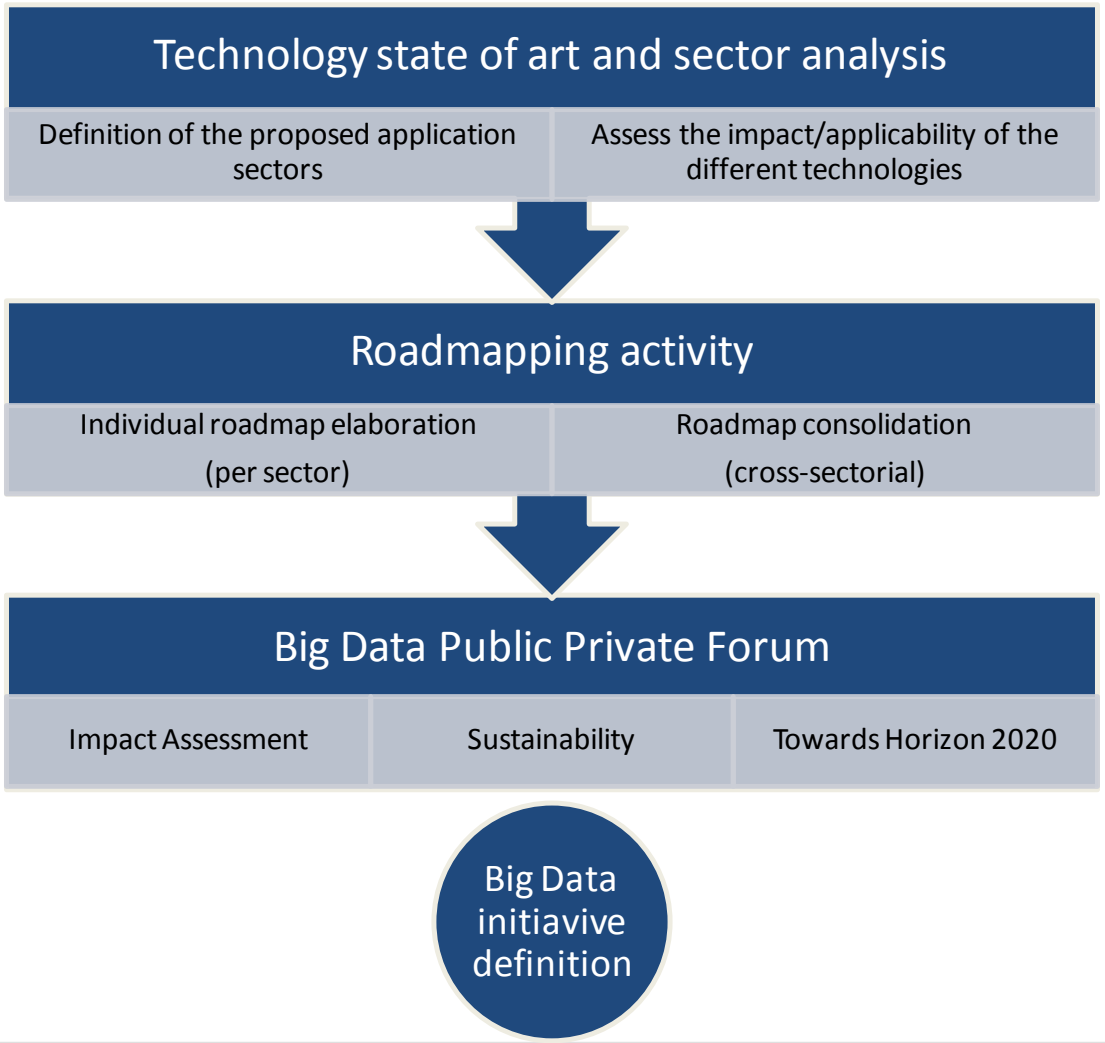
- Not only technology, but also **business**, policy and regulation;
- Not only generic plans for research, **but specific plans for adoption** for those sectors that are positioned for greater gains from the use of Big Data;
- Not only theoretical activities including roadmaps, coordination and dissemination 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: approach (II)

- Presence of the right profiles in the consortium to drive this process to the level of influence and impact we are aiming for
- 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: major activities



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…)?	

BIG: project structure

Industry driven working groups

Health	Public Sector	Finance & insurance	Telco, Media & Entertainment	Manufacturing, Retail, Energy, Transport

Needs ⇅ Supply

Value Chain

Data acquisition	Data analysis	Data curation	Data storage	Data usage
• Structured data • Unstructured Data • Event processing • Sensors networks • Streams	• Data preprocessing • Semantic analysis • Sentiment analysis • Other features analysis • Data correlation	• Trust • Provenance • Data augmentation • Data validation	• RDBMS limitations • NOSQL • Cloud storage	• Decision support • Decision making • Automatic steps • Domain-specific usage

Technical areas

Atos Research & Innovation

Nuria de Lama

Thank you

For more information please contact:

Ph: + 34 91214 9321
nuria.delama@atosresearch.eu

18/06/2012