ESA Big Data related Initiatives

Michele Iapaolo, Joost van Bemmelen, Andrea Della Vecchia
European Big Data Value Forum 2017
Versailles, 23/11/2017

ESA UNCLASSIFIED - For Official Use

Outline

european space technology harmonisation

1. Introduction
2. Big Data from Space (Conference) – **BIDS 2017 Nov 28-30 Toulouse**
3. Big Data from Space Harmonisation / ESA Competence Domain 9
4. ESA Big Data activities
5. Opportunities and future activities

Big Data from Space

european space technology harmonisation

- **Earth and Space observation data** from *space-borne, aerial and ground-based sensors*
- Qualified as **Big Data** according to the **5-V paradigm**:
  1. ***Volume*** of sensed data (archived data reaching the Exabyte scale)
  2. ***Velocity***: new data is acquired almost on a continuous basis and with an increasing rate)
  3. ***Variety***: data acquired by various sensors with different techniques
  4. ***Veracity***: data associated accuracy and uncertainty, provenance and trust
  5. ***Value***: depends on our capacity to extract meaningful and useful information from them
- **Multi-Domain** by nature: *Earth Observation, Space Science, Spacecraft Operations, Navigation, Telecommunication, Space Situational Awareness, Secure Societies, etc.*

Big Data from Space - Key Issues

european space technology harmonisation

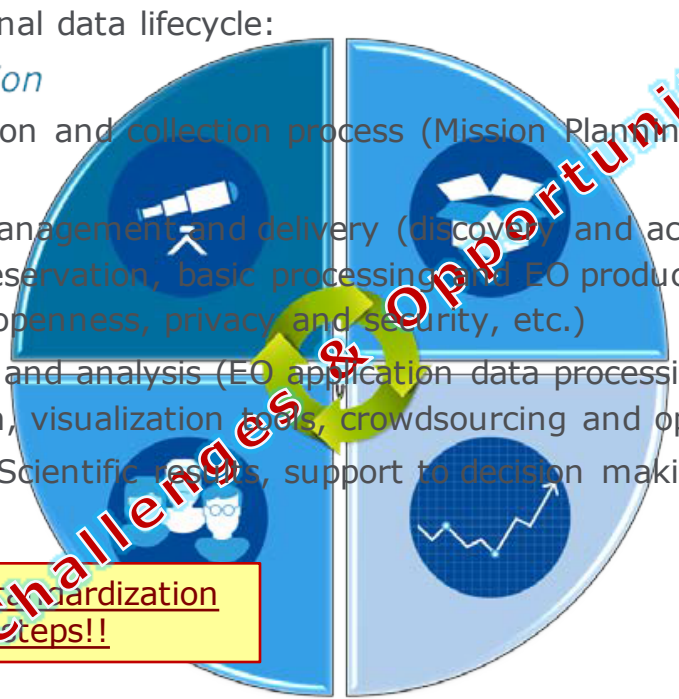
Main challenges introduced by Big Data (5-V paradigm):

	Traditional Data	Big Data
Volume	GigaByte (10^9)	TeraByte (10^{12}), PetaByte (10^{15}) and beyond
Velocity (Rate of generation and dissemination)	Per hour, per day	Faster
Variety (Structure and heterogeneity)	Structured	Semi-structured or un-structured, coming from heterogeneous sources
Veracity (Data Source Organisation and Trust)	Centralized	Distributed
Value (Added Value, Data Integration, Data Fusion)	Easy / Low value	Difficult / High value

Big Data from Space – Data Lifecycle

european space technology harmonisation

Big Data Model built around traditional data lifecycle:



- **Data Acquisition**: data acquisition and collection process (Mission Planning, Monitoring & Control functions, downlink capacity, etc.)
- **Data Organization**: storage, management and delivery (discovery and access functions, availability and retrieval, long-term archiving and preservation, basic processing and EO product generation, heterogeneous data sources, semantic access, data openness, privacy and security, etc.)
- **Data Analysis**: data processing and analysis (EO application data processing, analytics functions, processing chaining and orchestration, visualization tools, crowdsourcing and open science approaches, etc.)
- **Data/Information Provision**: Scientific results, support to decision making, delivery of services, etc.

Interoperability, Harmonization & Standardization and Security across the four steps!!

Big Data from Space Conference

european space technology harmonisation

- The advent of the Big Data from Space era has been confirmed by the success of the Big Data from Space event organised by the European Space Agency (ESA) in June 2013…

- …and the series of ***Big Data from Space Conferences*** (jointly organised by ESA, JRC and SatCen) in November 2014, March 2016, November 2017 → **BIDS 2017, 28-30 Nov @CNES, Toulouse**
  - Bring together researchers, system engineers, algorithm developers, industry, stakeholders and users interested in Big Data for the Space sector
  - Foster synergies and cross-fertilisation opportunities between the various domains, to jointly address common Big Data issues
  - Widening competences and expertise of universities, research institutes, labs, SMEs and all industrial actors in the field
  - Foster networking of experts and users towards better access and sharing of data, knowledge, tools and resources

Big Data from Space Conference

european space technology harmonisation

The BIDS Conference is helping:

- Foster and coordinate the European Research & Development activities to address forthcoming Big Data challenges, including "partnership models" for the actual exploitation of future Big Data flows
- Identify priorities and prepare implementation plans for research, technology development and innovation
- Establish and promote a close and permanent dialogue with key industry and SMEs actors
- Leverage innovation, spin-in & spin-off of technologies and business development
- Increase and promote the (potential) value stemming from the Big Data
- Promote the Data Science cross-cutting discipline, addressing the whole data lifecycle

Big Data Harmonisation

Big Data from Space is:

- Multi-domain; many key technologies are non-space driven
- Evolving fast and involving global actors
- New topic for harmonisation *(no previous roadmap exists):* benefit from Big Data across multiple space and non-space domains

➡ ***Mapping Meeting*** (held with national delegations) permitted to:

- Improve awareness of existing capabilities
- Identify technology gaps to address Big Data challenges
- Discuss key issues and R&D priorities
- Mapping the technology state-of-art and benchmarking the European industrial capability is extremely difficult!!

➡ **1st Big Data Roadmap** preparation (Feb-Nov 2017)!!

Recommendations

1. Support and catalyze the development and competitiveness of European actors (private companies, research institutes, university laboratories, etc.) and citizen awareness, enabling their capabilities to actually "exploit" Big Data
2. Agree on the logical functionalities ("building blocks") and the "glue" between the logical blocks: ***Collaboration*** and ***Interoperability*** aspects shall be reinforced, and guidance on such integration (building-up) process shall be provided
3. The ***Standardisation*** aspects are important as well: reuse and harmonise existing components, provide guidelines and recommendations, write best practices and advice Standardization Organizations (like e.g OGC)
4. Consolidate a first ***Roadmap*** built around few "Lines of Action" (***AIMS***)
5. Fast short-term validation and revisit cycle

Aims (Roadmap first draft)

european space technology harmonisation

A. Improve technology for data, information and knowledge discovery and accessibility (Volume, Velocity, Variety);
B. Improve technology for data analysis, visualization and use (Value);
C. Improve technology for Long Term Data Preservation;
D. Improve technology for Data Quality Assurance (Veracity) and Data Privacy, Security and Tracking;
E. Improve collaboration among European stakeholders considering interoperable functional components (e.g. for storage, discovery, access, visualization, analysis, etc.);
F. Promote and support use and commercial exploitation of data value;
G. Explore future opportunities for Space Applications;
H. Improve on-board data processing functions.

Big Data (ESA Competence Domain 9)

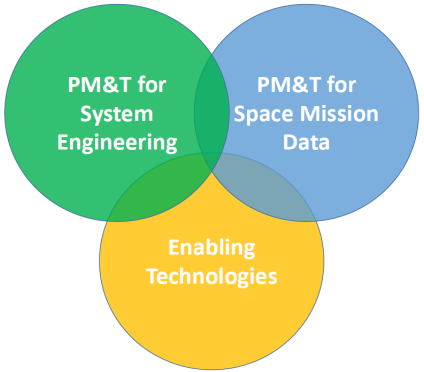
ESA Competence Domain 9 (CD09): ***Digital Engineering for Space Missions***

CD09 Objectives:

- Leverage the consistent use of *digital technology* to improve *efficiency* and *effectiveness* of the ***development processes***
- Targeting engineering, quality and management and the overall system aspects of a ***Space Mission Product Lifecycle***
- Analysing the ***state-of- the-art*** of these technologies in ESA and in the European space sector
- Identifying ***gaps*** and high-level ***needs***
- Defining ***strategic Lines of Action*** and ***priorities*** for R&D plans

The CD09 addresses the above objectives via its ***three areas***:

- Processes, Methods & Tools for ***System Engineering***
- Processes, Methods & Tools for ***Space Mission Data***
- ***Enabling Technologies*** (for Digital Engineering)



Big Data (ESA Competence Domain 9)

european space technology harmonisation

CD09 Technical Areas - *Processes, Methods & Tools for Space Mission Data*

Space Mission Data Management: ICT and SW *methods* and *tools* for *storage*, *organization*, *indexing*, *querying* (e.g. archives, catalogues, databases, brokers, etc.), *access* and *dissemination* of *Space Missions Data* (i.e.: HKTM data, payload data, value-added data and extracted information). It includes *harmonization* and *standardization* for *interoperability* of *data archiving*, *cataloguing*, *discovery*, *access* and *exchange services, long-term preservation of mission data*, along with associated documentation and processing SW.

Space Mission Data Processing and Analysis: ICT and SW *methods* and *tools* for:

- *Calibration*, *validation* and *conversion* of raw (payload) data into science data products using standard data formats;
- Data *pre-processing* and *preparation* (e.g. Extraction, Transformation and Loading) before the analysis;
- Data *processing* and *analysis workflows* and *orchestration*: algorithms chaining, data flows, processing tasks activation, etc.;
- Data *analytics* and *exploitation*: feature extraction, data mining, data fusion, data correlation, machine learning, artificial intelligence, data diagnostics, trends analysis, behaviour prediction, etc.

Space Mission Data Visualization and User Interfaces: ICT and SW *methods* and *tools* for *visualization* and *interaction* with Space Mission Data and extracted information. It includes data visualization, visual analytics, man-machine interfaces, virtual reality, augmented reality, etc.

Space Mission Data Security: ICT and SW *methods* and *tools* to guarantee and maintain the *protection* and *security* of space missions data assets across systems. The general aspects of security are discussed in the Enabling Technologies area.

Big Data (ESA Competence Domain 9)

european space technology harmonisation

Processes, Methods & Tools for Space Mission Data - Lines of Action

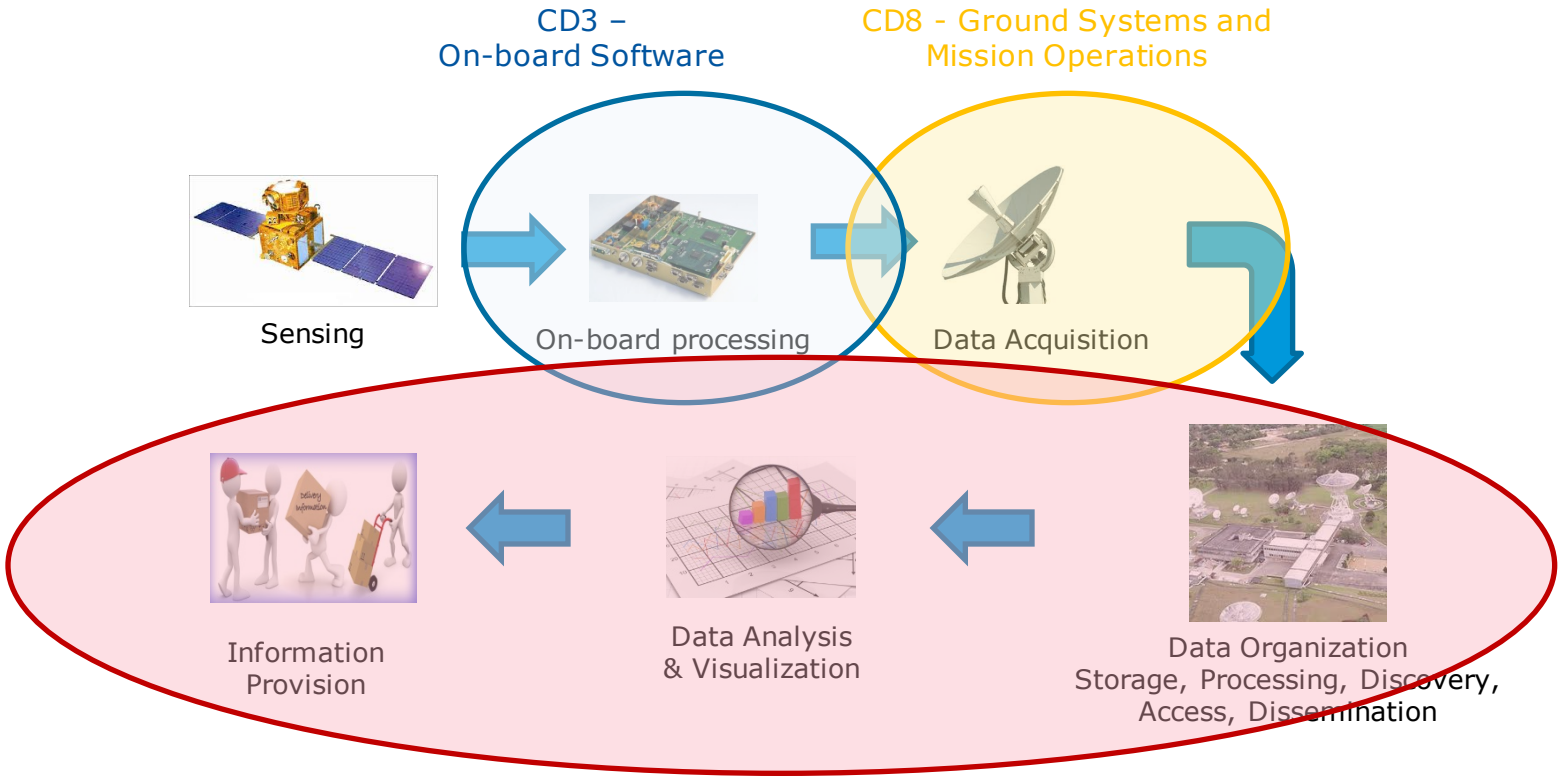
Space Mission Data Management:
- Reference distributed architectures for space mission data
- Data discovery, access and dissemination
- Scalable data storage and processing frameworks, massive data analysis methodologies and infrastructures
- Harmonisation, Standardisation and Interoperability of data and components

Space Mission Data Processing and Analysis:
- Data search and data mining tools
- Data analytics tools
- (Distributed) data processing workflows and orchestration

Space Mission Data Visualization and User Interfaces:
- Improve the effectiveness of interfaces (e.g. usability, ergonomics, robustness, etc.)
- Improve the responsiveness and adaptability (e.g. multi-modality, user interactions, etc.)
- Improve the performance (e.g. size of the data volumes, timeliness of interactions and visualization, etc.)
- Improve and develop guidelines for future data visualization prototypes and products to improve the human-system interaction

Space Mission Data Security:
- Cryptographic Key Management Infrastructures to ensure security and trust
- Physical Layer Security to ensure transmitter identification and protect confidential communications
- Data Link Security / Network Security / Software Security

Building Blocks



Planned and Upcoming Projects - GSTP

european space technology harmonisation

Project	CD09 Technology Area / Line of action	BIDS Harmonisation Aim
Data Analytics for EO (DANEO - GSTP)	Space Mission Data Processing, Analysis and Understanding / Data analytics tools	AIM B: Improve technology for data analysis, visualization and use
New Generation Virtual Reality for Big EO data sets (VIREO - GSTP)	Space Mission Data Visualization and User Interfaces / Improve the effectiveness of interfaces; Develop future data visualization prototypes	AIM B: Improve technology for data analysis, visualization and use AIM F: Promote and support use and commercial exploitation of data value
Microservices to Applications - a Generic Interdomain Composer (MAGIC - GSTP)	Space Mission Data Processing, Analysis and Understanding / Data processing workflows and orchestration; Interoperability of components	AIM B: Improve technology for data analysis, visualization and use AIM E: Interoperable functional components
BlockChain for Space Activities (BC4SA - GSTP)	Space Mission Data Security / Ensure security and trust	AIM D: Improve technology Data Privacy, Security and Tracking AIM F: Promote and support use and commercial exploitation of data value
New EO value chains using HC, CS and Big Data Analytics (NEOVAC - GSTP)	Space Mission Data Processing, Analysis and Understanding / Data analytics tools	AIM B: Improve technology for data analysis, visualization and use AIM F: Promote and support use and commercial exploitation of data value

Endorsed by CD09
Published in GSTP Compendium 2017

Additional Big Data Initiatives @ ESA

european space technology harmonisation

- ***Exploitation Platforms (TEP, MEP, REP):*** Community-oriented user front-end for Big Data environments, with both data and processing resources (Coastal, Forestry, Geohazards, Hydrology, Polar, Urban, Food // Proba-V MEP // ESPOSS – Black Sea // IPT Poland)
- ***Data and Information Access Service (DIAS)*** for Sentinel and Copernicus service data on a cloud platform
- ***Datacube*** facilities to be deployed on top of DIAS → Provide operational Datacube service
- ***Network of EO Resources*** → Integration of EO Big Data resources into a common ecosystem
- ***Global Earth Observation System of Systems (GEO/GEOSS) Initiative***: ESA Contributions to the GEOSS Platform, consisting of the GEOSS Portal, GEOSS Discovery and Access Broker, the Yellow Pages, and the Status Checker (from ESA, CNR-IIA, University of Geneva and USGS respectively), and co-funding from the EC under H2020 co-funded activity EDGE (European Direction in GEOSS Common Infrastructure Enhancements).

Interoperability and Standardization

european space technology harmonisation

Data Catalogues:

FedEO (Federated Earth Observation missions access): a unique entry point (gateway) to a growing number of scientific catalogues and services for, but not limited to, European and Canadian EO missions:

- Provide brokered discovery, access and ordering capability to European/Canadian EO missions data based on HMA (Heterogeneous Missions Accessibility) interfaces;
- Implement the OpenSearch OGC (Open Geospatial Consortium) and other interfaces for an increased number of discoverable and accessible EO data collections;
- Implement the OpenSearch OGC interfaces for interfacing with CEOS Community Catalogues and Clients
- Server and client enhancements

Processing and Visualization Components: WPS, WMS, WFS, WCS, etc.

Conclusions

1. ESA is involved and promoting various initiatives in the Big Data field

2. Through the Big Data Harmonisation and Competence Domain 9 ESA is fostering common approaches and collaboration to address Big Data issues

3. ESA is trying to create consensus and agreement around the most needed Lines of Actions

4. Big Data from Space Conference → **BIDS 2017 Nov 28-30 Toulouse** to promote knowledge exchange and partnerships