european space technology harmonisation

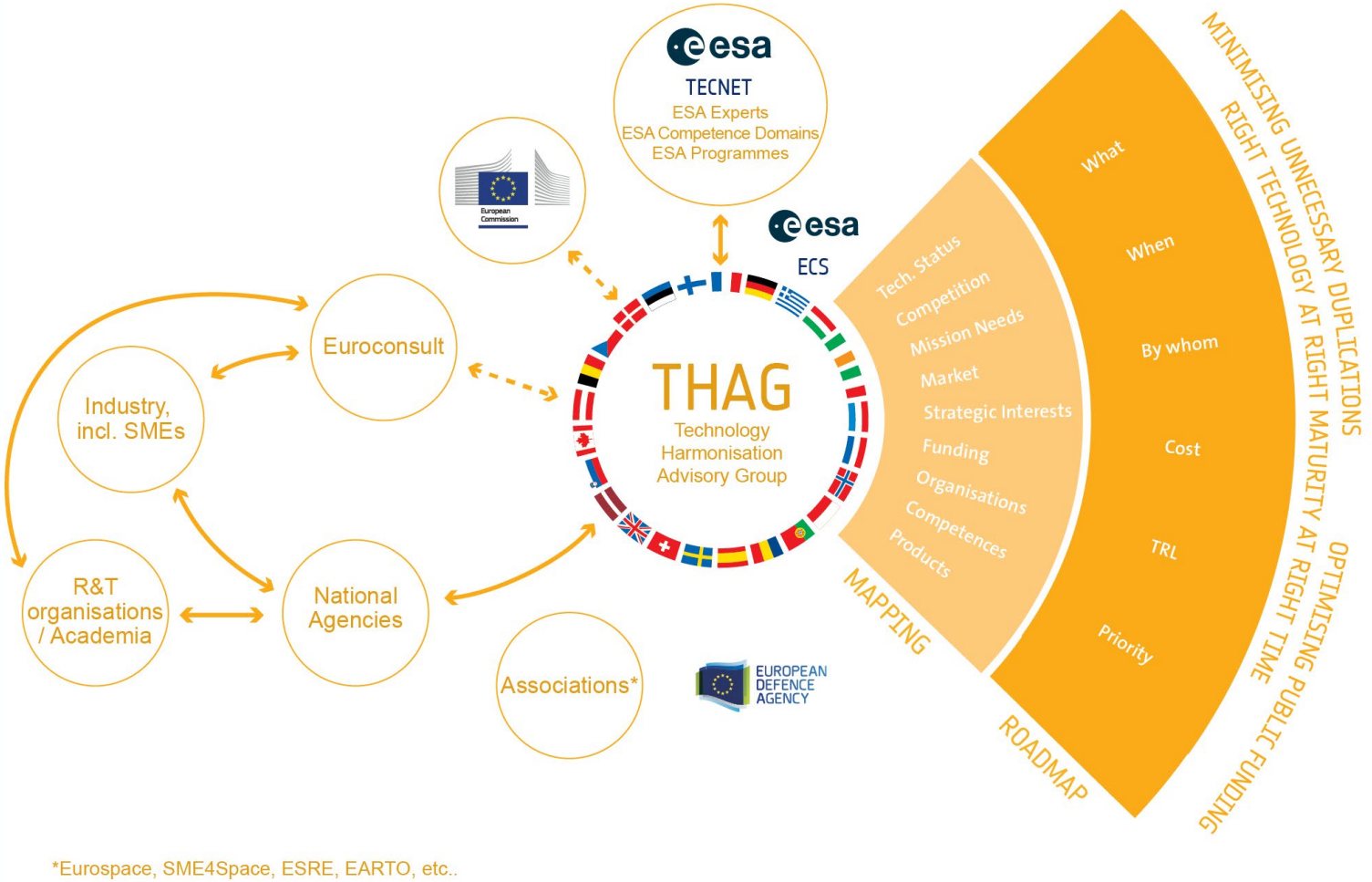
BIG DATA FROM SPACE (BDFS)

Harmonisation 2023 Cycle 2 – Current Status

Francesco Sgaramella

13/11/2023

BDFS Harmonisation Overview



*Eurospace, SME4Space, ESRE, EARTO, etc..

Implementation & Tracking

Big Data from Space (BDFS)

Big Data are characterized by:

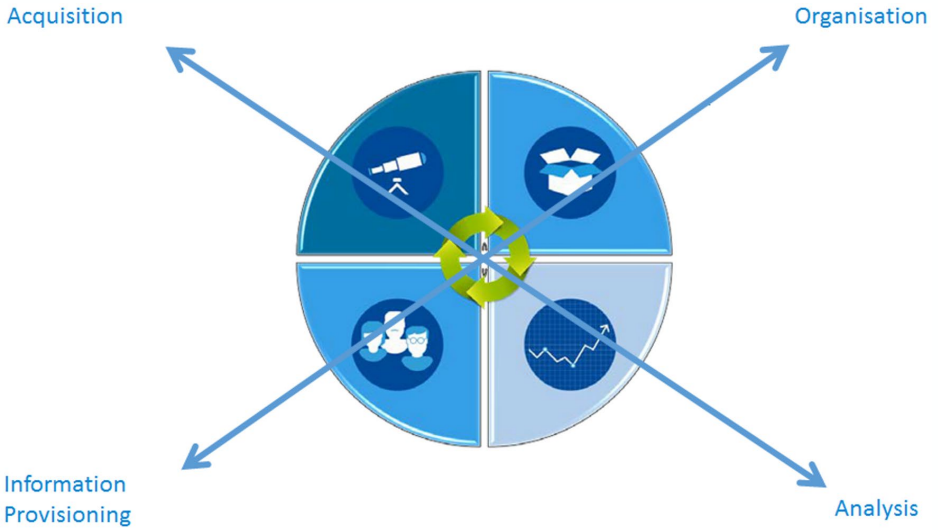
- The sheer volume of generated data – TB to EB
- Data high velocity – high frequency streaming data
- Data variety – structured, unstructured, multimedia, …
- Data veracity – consistency, completeness, …
- Data variability – data changing constantly
- Data visualization – dashboard, graph, …
- Value – business decisions from big data

BDFS refers to Earth and Space observations (EO, SCI), spacecraft housekeeping and telemetry data (OPS), Space-related applications (NAV, TELECOM, etc.), downstream applications.

Big Data from Space Lifecycle

BDFS lifecycle allows to:

- Deriving common technology requirements
- Eliciting the cross-cutting enabling technologies
- Highlighting the research areas requiring higher priority
- Reusing and adapting cross-industry and potentially Open Source components
- Ensuring interoperability, data protection and IPR
- Managing Data Provenance
- Generating and delivering integrated, meaningful and reliable information



Key Issues

- Allow effective access to simplify and integrate data management, analysis and delivery across a choice of endpoints spanning on-premises and multiple cloud environments
- Allow storage, organisation, preservation, retrieval and processing of data streams
- Allow access and exploitation of heterogeneous data through harmonized and standardized mechanisms and protocols
- Ensure data quality, integrity, provenance and trust, across organizational boundaries and along processing chains towards value-added products
- Extract information and inject it in downstream applications and services for research purposes and societal benefit (data science, ecosystem of distributed data exploitation platforms, support to engineering and modelling)
- Define and enforce widely recognized data usage licenses and data governance

Proposed Development Approach

Ref.	Aim Title	Aim Description
AIM A	Improve technologies for data, information and knowledge collection, discovery and access (Volume, Velocity, Variety, Variability)	To improve technologies and develop suitable infrastructures for the collection, organization, access and provision of data, related information and extracted knowledge (also permitting to share them across boundaries and organizations), facing the expected huge volumes of data, the velocity characterizing their production (data streams) and their variety and variability.
AIM B	Advance technologies for data analysis, visualization and use (Visualization , Value)	To advance on methods and tools dedicated to the analysis of such volumes of heterogeneous data, their visualization (also for interactive analysis) and their final use in practical applications (which is their actual value), developing the necessary basic algorithms, software frameworks and computing infrastructures for their deployment.
AIM C	Develop technologies for Data Quality Assurance and Tracking (Veracity), Privacy, Security and Protection	To investigate and develop mechanisms and technologies to ensure the quality of data and derived information, their provenance and integrity, in order to create trust on them when used in relevant applications and to avoid their leakage or tampering. Developed technologies shall guarantee a suitable balance between openness of data for their use and their protection, security and property right acknowledgement.

Proposed Development Approach (Cont'd)

Ref.	Aim Title	Aim Description
AIM D	Develop and promote interoperable functional components and federation of environments and platforms	To promote the development of interoperable functional components at European level, to foster the integration between systems and the federation of distributed environments and platforms, permitting to share data sources and processing resources.
AIM E	Improve methods and technologies to integrate space-derived data and their downstream applications into the user domain information ecosystem and promoting merge with non-space data	To significantly expand the use of space data for downstream commercial exploitation they have to become part of the usage domain (e.g. agriculture) big data ecosystem to be easily integrated with non-space data (e.g. IoT devices in a 5G connected environment). This implies exploring use of non-space ICT (Information and Communication Technologies) advance for space data to adopt new approaches for on-demand production, data discovery and data access, e.g. inherited from initiatives like GAIA-X.
AIM F	Develop and validate innovative technologies to improve on-board data processing functions	To develop and validate innovative technologies to improve space on-board data handling and transfer capabilities, processing functions, data storage and compression, payload data processing.

Conclusions – BDFS strategic aspects

- Big Data are today the core of businesses and industry growth to improve competitiveness in existing markets and to open new ones – this is applicable to the Space Industry as well
- Satellite services revenue is the largest space industry segment, such as satellite broadcasting and broadband services, Earth Observation, GNSS. The capabilities to handle and exploit Big Data are required to ensure such revenue
- There is a strategic interest for all European space industries to pull on new resources to establish the appropriate concepts, architectures, infrastructures and tools at the required time to market, in order to be able to keep the pace of the global market in the Big Data field
- Non-dependence on critical space technologies of European Institutions and Industry is a key factor. Big Data technologies, enabling the capability to acquire, manage, process and exploit Big Data, are a very crucial part of such non-dependence

Thank you for your attention