Earth Science community insight in support of ESA DP

Veronica Guidetti

European Space Agency

PV2009 Conference, Session *Adding value to data: Earth Science*

Outline

- Presentation context
- Out to the Earth Science community!
- Next steps

Outline

- Presentation context
- Out to the Earth Science community!
- Next steps

Some things we know very well… 1/2

- *Space data are a humankind asset*, their preservation is a challenging responsibility for Space Agencies and data owners
- DP must be seen as including data integrity, enabling data
- means, algorithms! We want to re-process long time series!

> *"In building a climate data record of sea level I am constantly faced with unearthing altimeter **data and auxiliary data from missions long deceased**. Old FTP archives / web pages, etc. disappear, and basic knowledge about the available data withers."*

Some things we know very well… 2 / 2

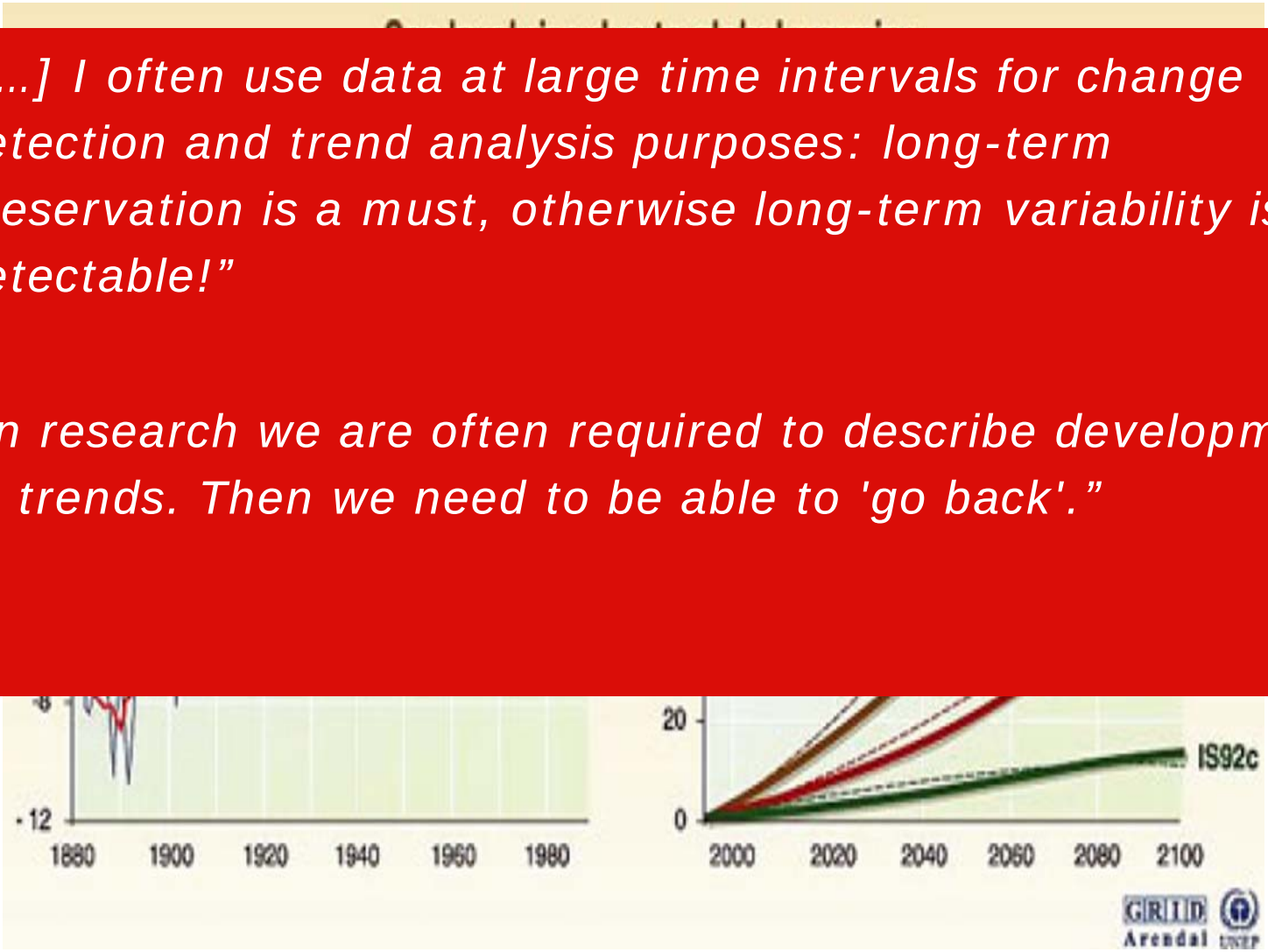
- Analysing the current state of the Earth, its environment and its variability over time requires a very large number of observations, usually impossible to resample, therefore global and complete measurements need to be performed

- **Easy access to global time-series**, their long-term preservation, availability of related scientific and technical knowledge, processors/algorithms, auxiliary data, and facilities for recovering, processing and calibrating archived data are all essential parts for generating consistent ECVs…

Scientific applications

- *“[…] I often use data at large time intervals for change detection and trend analysis purposes: long-term preservation is a must, otherwise long-term variability is not detectable!”*

- *“In research we are often required to describe development or trends. Then we need to be able to 'go back'.”*



And some things we do not know!

- What end users of our observations think about historical data! Did they ever need to access/use them? What could they
- are
- mig
- Pote
- can
- Whe
- hist
- is it

availability of current and future data? …

"Although GOME-1 is continuing to provide data after June2003, ***resources are limited or do not exist for producing and in particular recalibrating the data****. In general the calibration of the data both pre flight and in orbit, in particular dealing with degradation, has a high cost. The agencies in general do not have a sufficient budget for such matters."*

So comes the today context

- A programme to ensure preservation and long term availability of earth observations and Earth Science data has been approved
- ESA co-funds the Alliance for Permanent Access
- ESA started its Climate Change Initiative which requires processing of global historical time-series, regular re-analysis of archived data, periodic re-processing of basic datasets
- Some parallel experimentation is on-going at infrastructure and community level (EC projects)
- Big(!) amount of new data with Earth Explorers and Sentinels

A further action taken

- ESA's partnership in PARSE.Insight. Aligned with ESA's role of coordinator at European level to ensure preservation and accessibility for ESA and Member States' EO data in the long-term

- User consultation issued.
  Target: to get an insight about community's awareness on DP and current/envisaged exploitation of historical environmental data streams, including opportunity of experience, i.e. *requirements*.
  Result: high and active participation, further requirements on the table

ESA's objectives in Parse.Insight

- To get an insight in the Earth Science community, understanding EO data users' standpoint and requirements about historical space and non space data exploitation

- To provide input to LTDP initiatives, e.g. technology studies + dataset/information identification

- To contribute with further information to fill the gap between
  - EO data generation, archiving and maintenance
  - EO data exploitation

- To raise awareness on current DP activity in Earth Science

Outline

- Presentation context
- Out to the Earth Science community!
- Next steps

To start…

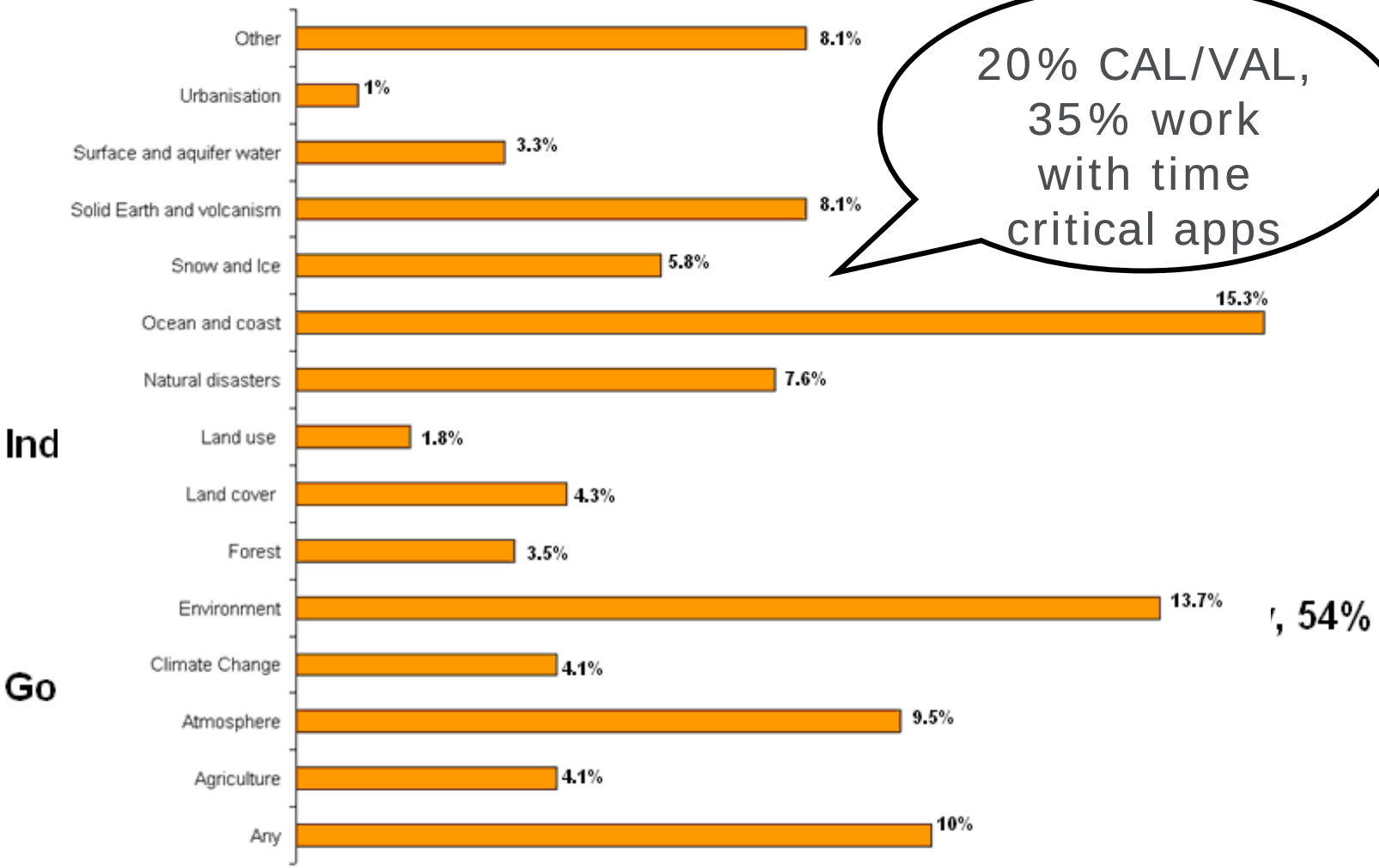
"I consider long-term preservation of earth satellite data to be critical and must be done at nearly any cost.

I actually think it should be literally against the law to not preserve data.

I believe the costs for long-term storage and continued accessibility should be coming down.

ESA should step up to the plate and take care of this very important issue."

Who participated



Focus on data

- Participants considered as ‘environmental data’ all types of earth observations + other type of information sources (documentation, presentations, technical reports…)
- Physical samples, in-situ and remotely sensed data classified per:
  - model output and synthesized products derived from these data
  - products developed through interpretation of original or synthesized products
  - products whose quality is based on experimental capabilities/algorithms

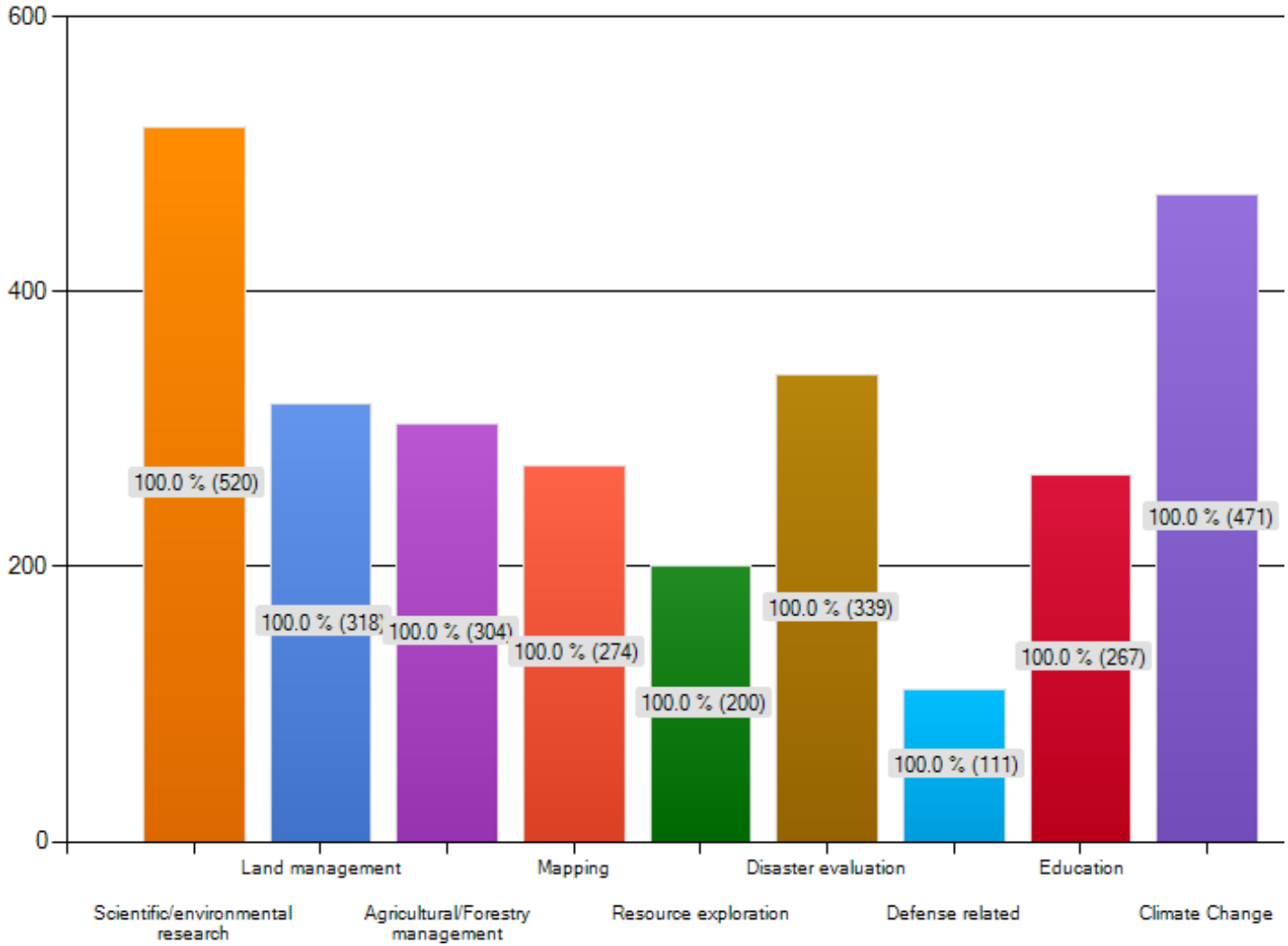
Quicklook at the result

- Familiarity with DP is there (but the 13%)
- High majority constantly access historical data (>85%)
- 46% required at least once to access them
- 25% report on cases on data losses or unavailability for *ERS1/ERS2, Landsat, Geosat, ENVISAT/ASAR-MERIS, ATSR-1/2, SeaStar/SeaWiFS, and SEASAT 1978, in-situ data*
- Suggested for preservation: original data and optical /multispectral radiometry products (land theme). Synthesized, interpreted or experimental products, radar imagery, atmospheric data and radar altimetry (oceans, air and cryosphere)
- Familiarity with standards: HDF, netCDF , ISO 19xxx, INSPIRE, OGC…

Reported experiences with past data

Mission / Sensor	Applications
ENVISAT/AATSR,ASAR, GOME,MERIS, SCHIAMACHY/L0-L1 data	Measurement of slow deformation rates around earthquake faults Analysis of tectonic loading of faults or the understanding of landslide history
ERS1/ERS2, SLC	
MSG	Mapping of historical land use/land cover Vegetation phenology Forest health development
PROBA/CHRIS	
Landsat(s)/MSS from 70s	
Terra-Aqua/MODIS	Tracing gases for data assimilation, time series analysis, climate model evaluation, trend analysis and for ascertaining seasonal changes and background levels of atmospheric pollutants
NOAA/AVHRR	
SEASAT 1978	
RADARSAT Constellation	Coast line monitoring over decades
QuickBird	
SPOT(s)	Variability of oceanographic conditions and productivity Use of time series of water quality, plankton development, seasonal aspects
Atmospheric, ocean colour, SST, LAI, NDVI time series data, in-situ data, LR, MR and HR data	Study of solar variability and stratospheric ozone CAL/VAL processing Interferometry

Where DP is perceived to add value



Data losses and threats to DP

- Aging and degraded/damaged media, hard disks failures and inadequate backups, failures in storage and data transfer
- Lost media, equipment to read data files no longer existing, disappearance of old cassette reader devices, files corruption
- Char ble anyr etadata, no re softv
- Resp

"SAR data from the Shuttle Imaging Radar mission C (SIR-C) that flew on the Shuttle twice in 1994 is effectively lost because the computer hardware to read and process the original data is obsolete and no new system has been created to read the data."

Ideal IT (scientific data) infrastructure

- Helps guard against some of the identified threats
- Improves data access/distribution systems
- Allows to compute on-demand and e.g. visualize domain specific long-term trends (desertification, ozone levels, climate change...), data intensive processing
- Allows to try new data assimilation methods and apply models
- Allows to ignore data location and maintenance
- Allows to take advantage of computational resources not accessible otherwise (transparency, clouds...)
- Enables access to whatever type of data type
- Provides end-to-end interoperable services
- Provides quality information

What we can derive, critical analysis

- Community requires similar services
  - New programs together with evolving user needs impose the implementation of similar services both for historical and current missions, like DP, access and distribution (HMIs), use of metadata standards, performance, on-demand processing, validation, interoperability..

- Focus is on end-to-end (operational) systems
  - So far projects have focused on different aspects of DP
  - Scattered initiative only
  - Main focus has been on DRs only but they constitute one component within a larger infrastructure required in operational scenarios

Outline

- Presentation context
- Out to the Earth Science community!
- Next steps

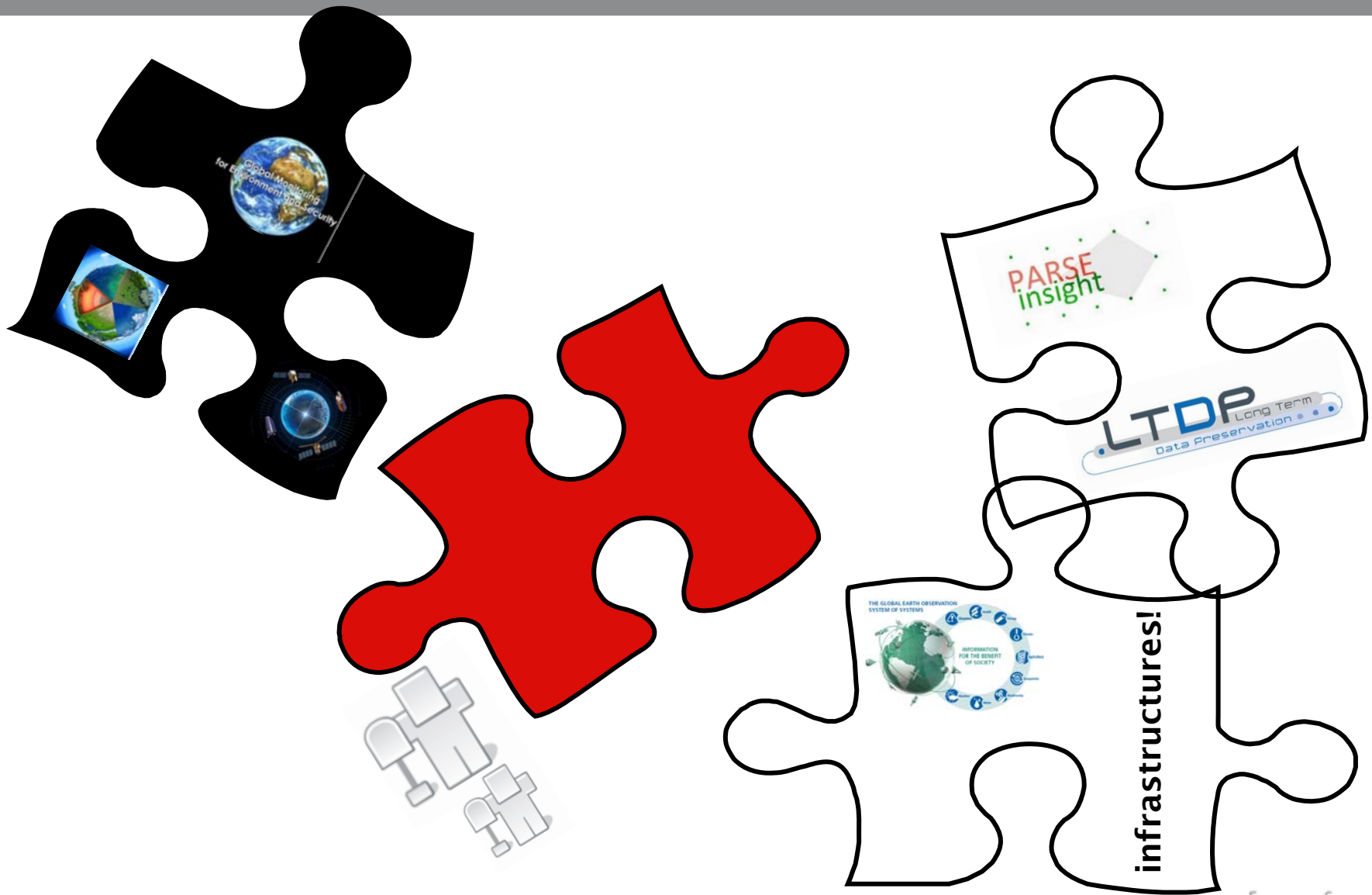
Drivers

- In-line with data sharing principles discussed at EC, GEO/GEOSS and (inter-)national space agencies level, the community require liberal data distribution policies and **missions' long term commitment**, **to enable operational data exploitation**
- In addition to new data the scientific community want to access historical time series of earth observations
- The community aim at a more active involvement in the process not only via reporting experiences and suggestions, but as customer of a global information system
- Data users require timely solutions to current infrastructures' constraints

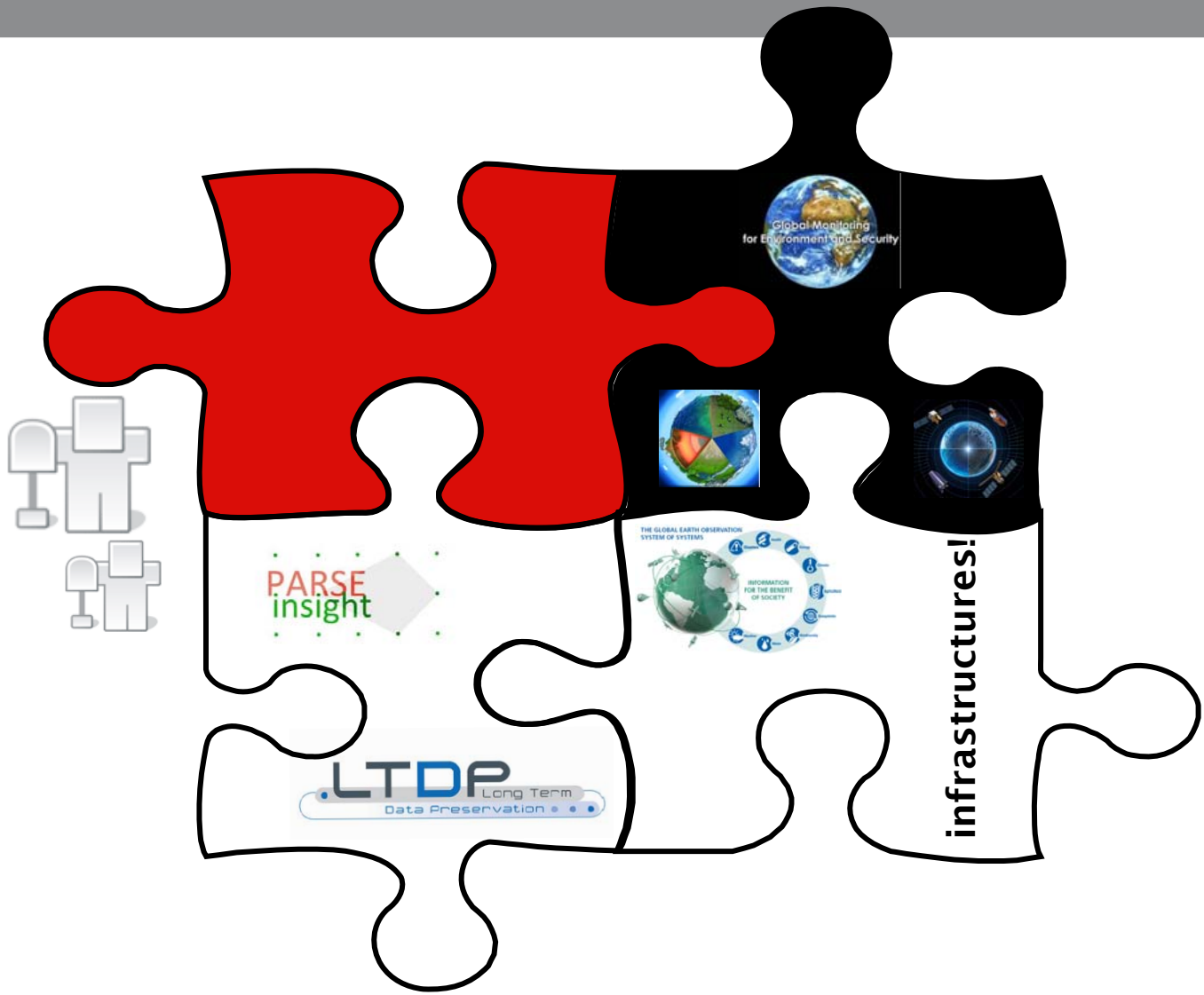
Potential barriers

- Lack of high level European policies
- Scattered initiatives at different levels, pilot, no real services
- Best efforts and personal initiatives
- Unclear long term commitment from organizations
- Lack of coordination and missing firm plans
- Absence of common models
- Difficulty in data access
- Resources limitation

Our table…


Global Monitoring for Environment and Security
PARSE insight
LTDP Long Term Data Preservation
THE GLOBAL EARTH OBSERVATION SYSTEM OF SYSTEMS
INFORMATION FOR THE BENEFIT OF SOCIETY
infrastructures!

And how that should be



One message to conclude

*"The added values of historical data **compiled in standard formats** and **publicly available** is much higher than at this time known."*

Thank you!

veronica.guidetti {*at*} esa.int