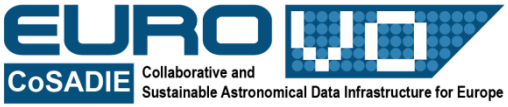
The user point of view

Françoise Genova



- At present, in same cases ESFRI users use the facility and data sharing is not the point
- There can be a gap to bridge between the facility users and the potentially wider community of data users

- Data users want to discover, access, use and compare/integrate data
- This requires that
  1. Data is in the system
  2. Data is available
  3. Data is understandable and reusable
  4. Data is discoverable
  5. Data can be compared to other data
- Some may want to perform computation

Data is in the system

- The ESFRIs have to set up data archives
- This requires to earmark part of their budget for this purpose
- This is in competition with infrastructure updates, new instruments, …
- The funding agencies have a role to play by establishing rules
- But it also requires to convince committees and the ESFRI user community - > usable and useful data

Data is available

- The ESFRIs should discuss their data policy with in mind that openness is a goal and exceptions should be justified
- Again the agencies have a role to play but the ESFRI committees and communities have to agree
- Are there ways to make openness accepted by communities?
- In astronomy for instance, in general a one year *proprietary period* for the scientists who got observation time in Calls for Proposals tough competition before opening data
  - Legitimate interests are protected
  - The community uses data obtained by others
  - Virtuous circle

Data is understandable and reusable

- Agree on data formats and data descriptions
- At least at the level of the ESFRI that provides the data, with in mind user needs
- Better if formats and descriptions are set up in common by similar ESFRIs or their communities – allows the usage of common tools

Data is discoverable

- There is a registry of resources and a mechanism to discover data and select data of interest
- The registry includes enough information to be able to select the data of interest
- At least ESFRI archive log
- More generally discovery from higher level registry(ies)
- Dublin core is not sufficient
- Work on the semantics at ESFRI(s)/disciplinary level

Data can be compared to other data

- ESFRI data can be compared between themselves and with other data of interest (data of other similar instruments – European and international, modelling data, …)
- Interoperability
  - Common formats or format translators
  - Semantics
  - Interoperability standards and tools

- ESFRIs are not isolated islands
- Discussions at community level at least on formats, semantics, registries, etc
- Compare data from one ESFRI with data from other ESFRIs, from other similar infrastructures, with modelling results, etc
- Communities are not at the same level of preparadeness
- They need support to get the discussions started and then to maintain the community framework
- The ESFRIs have a role to play as major data providers, to initiate discussions with the communities or to be major actors in the discussion organised at community level

- Likely not a single system in the end
- Use of generic building blocks when possible
- Check what projects such as EUDAT or the RDA are proposing or better participate in the discussions in those projects to pass the specific requirements and feedback
- The interface between these generic elements and well established community frameworks has to be assessed – but do not break what is working
- A role for e-Infrastructure also linked to cases in which users want to preform computation with data