

SV Community Behaviours

A **behaviour** of a community is a composition of **processes** addressing specific scientific concerns. Behaviours are performed by **SV Community Roles**. In the ENVRI RM, the modelling of community behaviours is based on analysis of the common operations of research infrastructures which has resulted in [Appendix A Common Requirements of Environmental Research Infrastructures](#). The community behaviours model focuses on supporting a predefined **minimal set** of functionalities. A community behaviour can be either a single function or a composition of several functions from the function list.

A **process** is a collection of **steps** taking place in a prescribed manner and leading to an objective. A **step** is an abstraction of an action, used in a **process**, that leaves unspecified the objects that participate in that action. Steps are introduced in the definition because not all the action details need be specified in the composition.

Note

The Science viewpoint of the ENVRI RM only defines high level behaviours. A behaviour can be refined further by defining its constituting processes and steps, however, this task is left to the actual system designers.

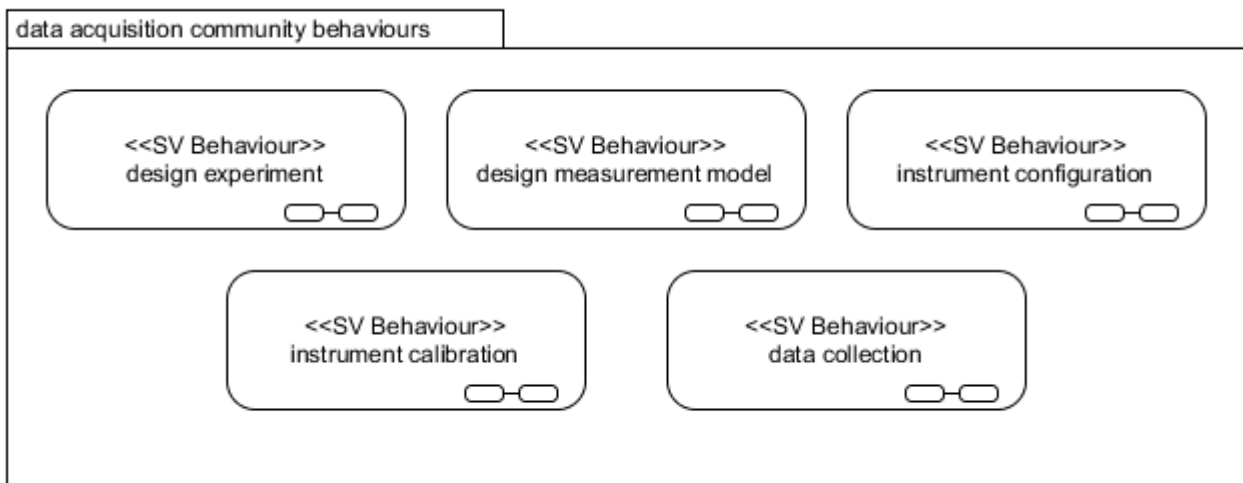
- [Behaviours of the Data Acquisition Community](#)
- [Behaviours of the Data Curation Community](#)
- [Behaviours of the Data Publishing Community](#)
- [Behaviours of the Data Processing Community](#)
- [Behaviours of the Data Use Community](#)

Behaviours of the Data Acquisition Community

The key behaviours of the data acquisition community through the interaction of the community roles include:

- **Design Experiment:** A behaviour performed by a *Environmental Scientist* that designs the scientific experiment which motivates the data acquisition activities.
- **Design Measurement Model:** A behaviour performed by a *Measurement Model Designer* that designs the measurement or monitoring model based on scientific requirements.
- **Instrument Configuration:** A behaviour performed by a *Technician* that sets up a *sensor* or a *sensor network*.
- **Instrument Calibration:** A behaviour performed by a *Technician* that controls and records the process of aligning or testing a *sensor* against dependable standards or specified verification processes.
- **Data Collection:** A behaviour performed by a *Data Collector* that control and monitor the collection of the digital values from a *sensor* instrument or a human sensor, such as a *Measurer* or a *Observer*, associating consistent time-stamps and necessary metadata.

Acquisition Community Behaviours



Notation

The roles of the data acquisition community are described in [Acquisition Community Roles](#).

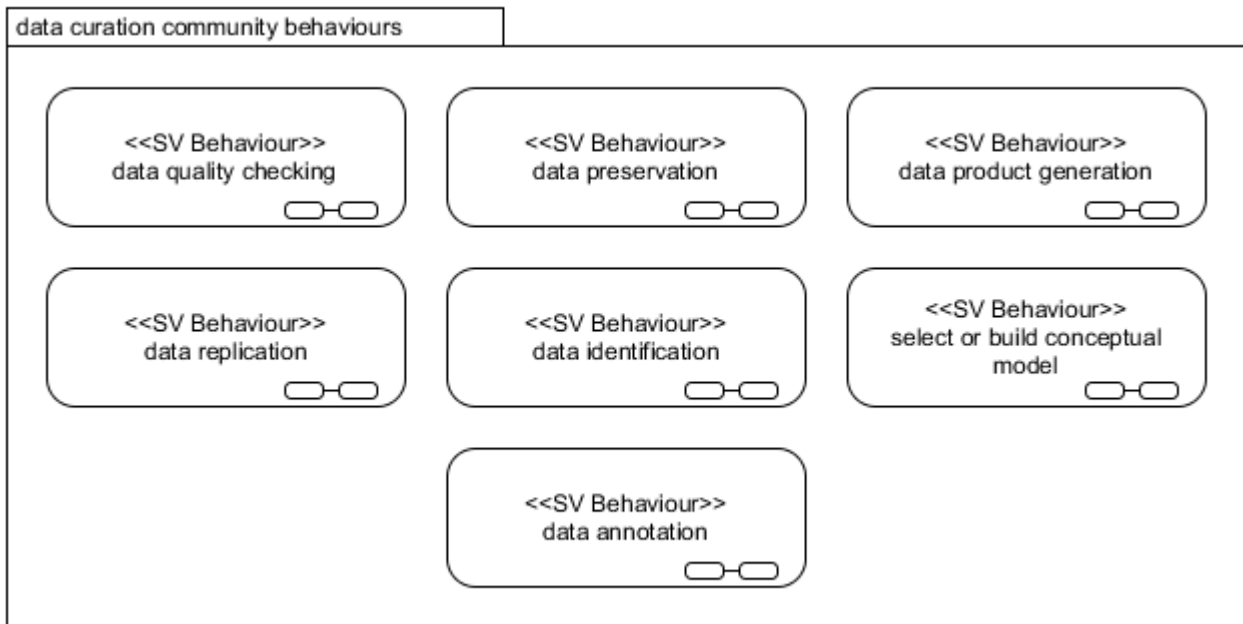
Behaviours of the Data Curation Community

The main behaviours of the data curation community include:

- **Data Quality Checking:** A behaviour performed by a *Data Curator* that detects and corrects (or removes) corrupt, inconsistent or inaccurate records from data sets.

- **Data Preservation:** A behaviour performed by a *Data Curator* that deposits (over long-term) the data and metadata or other supplementary data and methods according to specified policies, and makes them accessible on request.
- **Data Product Generation:** A behaviour performed by a *Data Curator* that processes data against requirement specifications and standardised formats and descriptions.
- **Data Replication:** A behaviour performed by a *Storage Administrator* that creates, deletes and maintains the consistency of copies of a data set on multiple storage devices.
- **Data Identification:** A behaviour performed by a *PID manager* which provides a unique PID for data and metadata being curated.
- **Select or Build Local Conceptual Model:** A behaviour performed by a *Semantic Curator* which supports the annotation of data and metadata.
- **Data Annotation:** A behaviour performed by a *Semantic Curator* which supports the linking of data and metadata with a local conceptual model.

Curation Community Behaviours



Notation

The roles of the data curation community which are described at [Curation Roles](#).

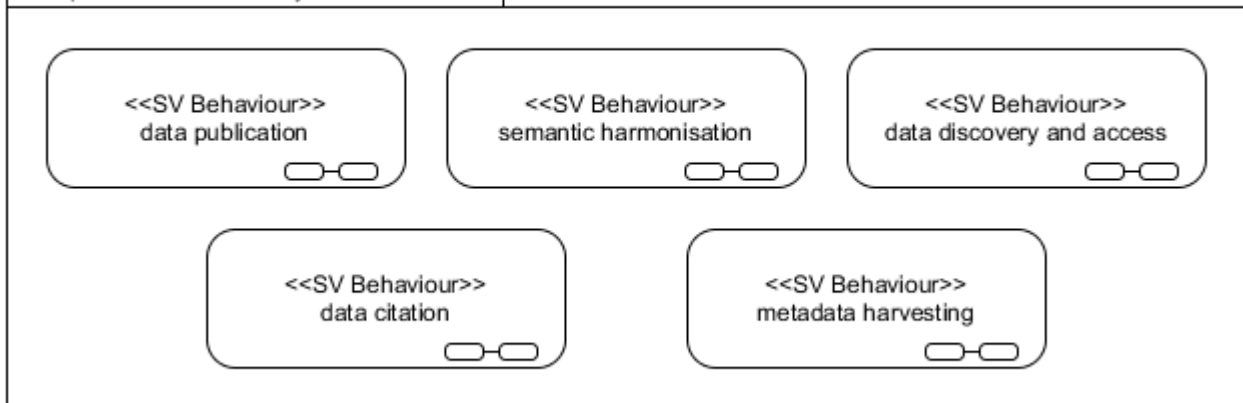
Behaviours of the Data Publishing Community

The data publishing community may perform the following behaviours:

- **Data Publication:** A behaviour that provides clean, well-annotated, anonymity-preserving datasets in a suitable format, and by following specified data-publication and sharing policies, to make the datasets accessible publicly or to those who agree to certain conditions of use, and to individuals who meet certain professional criteria.
- **Semantic Harmonisation:** A behaviour enabled by a *Semantic Mediator* that unifies similar data (knowledge) models based on the consensus of collaborative domain experts to achieve better data (knowledge) reuse and semantic interoperability.
- **Data Discovery and Access:** A behaviour enabled by a *Data Discovery and Access System* that retrieves requested data from a data resource by using suitable search technology.
- **Data Citation:** A behaviour performed by a *PID Manager* that assigns an accurate, consistent and standardised reference to a data object, in the same way as researchers routinely provide a bibliographic reference to printed resources. The RI publishing the data can define the citation contents such as authors, and dates for different citation styles.
- **Metadata Harvesting:** A behaviour performed by a metadata harvester to gather metadata from data objects in order to construct catalogues of the available information. A functionality that (regularly) collects metadata (in agreed formats) from different sources.

Publication Community Behaviours

data publication community behaviours



Notation

The roles of the data publication community are described at [Publication Roles](#).

Behaviours of the Data Processing Community

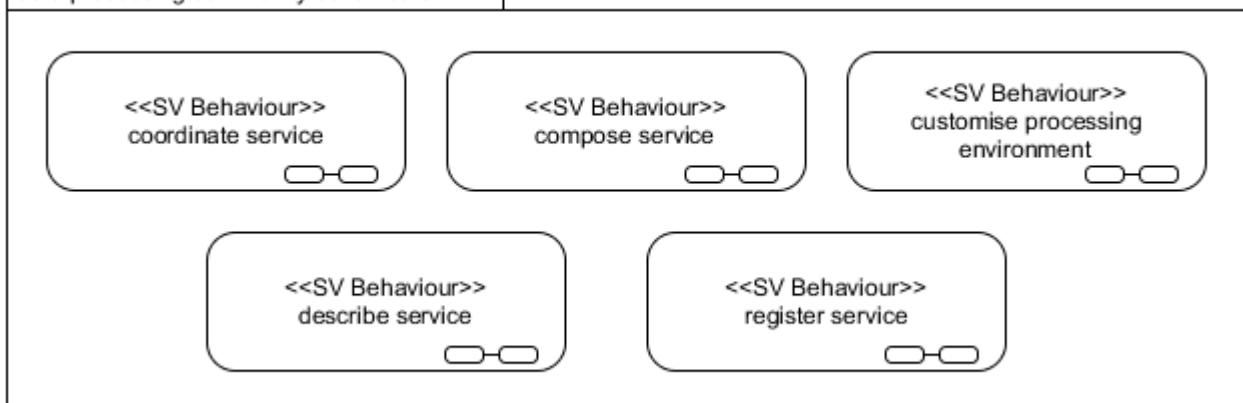
The following behaviours of the data processing community are modelled:

- **Coordinate Service:** A behaviour performed by a *Service Provider* to coordinate the actions of distributed applications in order to reach consistent agreement on the outcome of distributed transactions.
- **Compose Service:** A behaviour performed by a *Service Provider* to combine multiple services which can be achieved by either *Choreography* or *Orchestration*. **Service Choreography** is a collaboration between *Service Providers* and *Service Consumers*. **Service Orchestration** is the behaviour that a *Service Provider* performs internally to realise a service that it provides [[Bibliography#ref35](#)].
- **Customise Processing Environment:** A behaviour performed by a processing environment planner to enable a Data Processing Subsystem to prepare customised infrastructure and service platforms for managing specific data processing applications optimally, including the planning, provisioning and deployment sub-activities.
- **Describe Service:** A behaviour performed by a *Service Provider* to provide the information needed in order to use a service [[8](#)].
- **Register Service:** A behaviour performed by a *Service Provider* to make the service visible to *Service Consumers* by registering it in a service registry [[8](#)].

These are general behaviours of a service-oriented computing model. In the context of environmental science research infrastructures, a data processing community will focus on the implementation of domain special services, in particular those supporting **Data Assimilation, Data Analysis, Data Mining, Data Extraction, Scientific Modelling and Simulation, (Scientific) Workflow Enactment** (See [Terminology and Glossary](#) for the definitions of these functionalities).

Processing Community Behaviours

data processing community behaviours



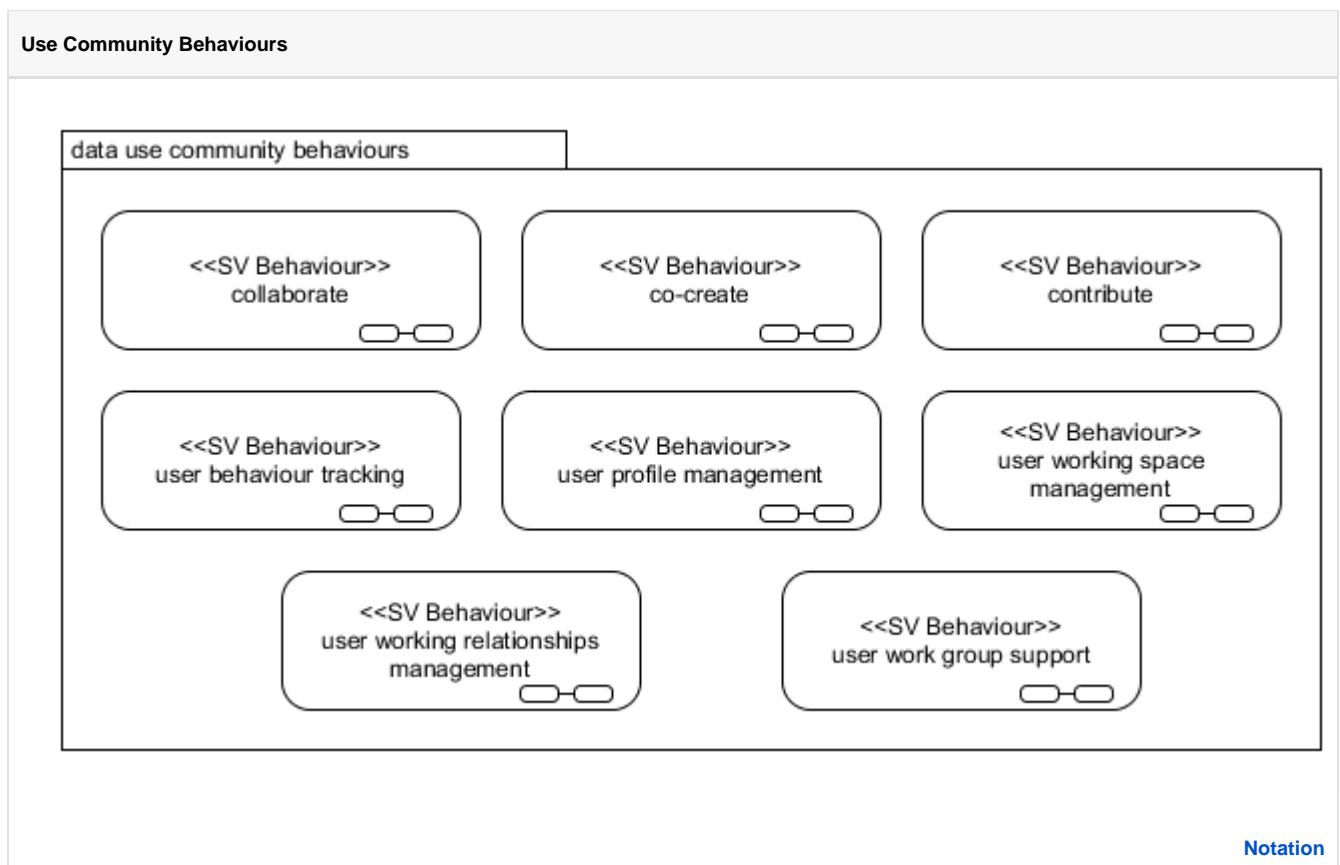
Notation

The roles of the data processing community are described at [Processing Roles](#).

Behaviours of the Data Use Community

The data use community can be divided in two main groups: (1) the behaviours performed by active roles (human actors) and (2) the behaviours performed by passive roles (computer resources). The first group encompasses the activities performed by human actors using the RI to interact with the different components of the RI. This can extend to all the actors in all the communities defined in the SV, in addition to the ones in the use community, these behaviours form a specialisation of the use community which can be called community support behaviours (or user support behaviours). The second group corresponds to the behaviours that enable the authorisation, authentication, and accounting of the activities of users, conforming a second specialisation which can be called AAAI behaviours.

- **Co-create:** A behaviour in which active roles from all communities participate in the design and planning of activities for the collection, preservation, analysis or publishing of research data.
- **Collaborate:** A behaviour performed by active roles which entails assisting/participating in some of the phases of the collection, preservation, analysis or publishing of research data.
- **Contribute:** A behaviour performed by active roles which entails directly collecting, preserving, analysing, or publishing research data held by the RI, according to a predefined protocol.
- **User Behaviour Tracking (synonym: real user monitoring):** A behaviour enabled by a Data Use Subsystem that to track the Users. User Behaviour Tracking is the analysis of visitor behaviour on a website. The analysis of an individual visitor's behaviour may be used to provide options or content that relates to their implied preferences; either during a visit or in the future visits. Additionally, it can be used to track content use and performance.
- **User Profile Management:** A behaviour enabled by a *Data Use Subsystem* to support persistent and mobile profiles, where profiles will include preferred interaction settings, preferred computational resource settings, and so on.
- **User Working Space Management:** A behaviour enabled by a *Data Use Subsystem* to support work spaces that allow data, document and code continuity between connection sessions and accessible from multiple sites or mobile smart devices.
- **User Working Relationships Management:** A behaviour enabled by a *Data Use Subsystem* to support a record of working relationships, (virtual) group memberships and friends.
- **User Work Group Support:** A behaviour enabled by a *Data Use Subsystem* to support controlled sharing, collaborative work and publication of results, with persistent and externally citable PIDs.



The roles of the data use community are described at [Use Roles](#).