EOSC Association Board position paper on the EOSC Federation and the role of Nodes

EOSC Association Board of Directors

Draft dated 13th October 2023

This document provides a high-level description of EOSC: EOSC is an open, trusted, federation of infrastructures that enables European researchers to store, share, process, analyse, and reuse research digital objects (e.g. data, publications and software). EOSC accelerates Open Science, FAIR data management and use of digital methods and services in research by stimulating cooperation across European, national, and institutional levels, resulting in increased research productivity and reproducibility. To achieve these goals, EOSC brings together the European Commission, the MS/AC governments and the many stakeholders involved in the European Research Area, in particular Research Infrastructures, to collaborate in its co-creation process.

To fulfil this vision, EOSC intends to become a federation of distributed systems, combined into a system of systems, consisting of multiple "nodes" that are interconnected and can collaborate to share and manage information and resources (FAIR data & services) across thematic and geographical boundaries. The node concept adds value by grouping resources with common policies and ownership or responsibility.

Examples of existing federated, distributed systems where nodes play a crucial role include those established by many Research Infrastructures. These systems will continue to exist and remain the means for serving their primary user communities. The EOSC Federation permits such systems to share resources with other EOSC communities and leverage common services.

The precise definition and characteristics of a node vary depending on the specific context (national, regional, institutional or thematic), but they share some common characteristics, including:

- **Collaboration**: Nodes in the EOSC Federation enable collaboration to achieve common goals. This involves sharing information, coordinating activities, or combining resources to complete tasks. This benefits the end-users by making more service and resources available, and by increasing the number of ways an end-user can combine those resources to reach their own objectives. The concepts of node and EOSC Federation should be transparent to the end-users.

- **Governance**: A governance structure is established at the EOSC Federation level to determine federation policies including the inclusion/exclusion of nodes. There are basic policies, rules of participation as well as technical profiles and standards at the EOSC Federation level that apply to all nodes. These rules are intended to maximize the possibility of additional "systems" (or nodes) joining the EOSC Federation, while at
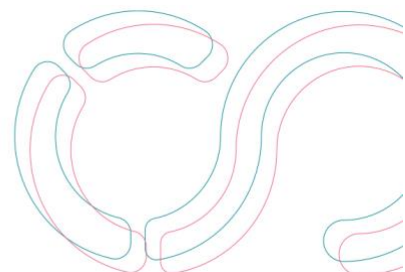
the same time encouraging minimum levels of quality, interoperability and consistency (of the user experience). Each node has an identified legal representative that takes responsibility for ownership and accountability.

- **Autonomy**: The legal entity responsible for a node has a degree of autonomy so it may operate independently to perform specific tasks or functions. There may be node specific policies that vary from one node to another (while remaining consistent with the policies at the federation level). For example, access policies may differ between nodes. The legal entity that is responsible for a node ensures that all applicable policies are enforced and monitored for the resources within the node. Each node has control over its own operations and the resources it manages. For example, nodes may have their own processing capabilities, storage, communication and training capabilities.

- **Interconnectivity**: Nodes offer interfaces that respect the EOSC Interoperability Framework. Nodes are connected to one another through a network, enabling them to communicate and exchange data or information. Information about a node and its resources should be visible to the whole federation.

- **Resources:** All nodes contain resources (FAIR data and/or services) that should add value to the EOSC Federation (as well as its users). The services may be of a technical nature (i.e. allow to perform actions on data) but may also include training services or expertise such as centres of competence.

- **Scalability**: The federation can be scaled by adding new nodes that adhere to the federation-level policies allowing for flexibility in accommodating changing workloads, requirements or capacity.

- **Heterogeneity**: Nodes in the EOSC Federation belonging to different categories may vary in the collection of resources to which they provide access and the infrastructure on which they are built.

The EOSC Federation model shall be enabled by an EOSC Interoperability Framework (EOSC IF), which defines the interfaces to interconnect EOSC Nodes and their resources.

Initiatives (European, national, regional, institutional or thematic) may join the EOSC Federation and establish EOSC Nodes when they are able to share resources (datasets, services, etc.) with other EOSC Nodes in compliance with the EOSC Interoperability Framework (technical compliance) and EOSC Rules of Participation (policy compliance, including acceptable use policies and resource access policies). Through the EOSC federation of nodes, end-users can access additional resources beyond their usual environment.