Contribution ID: **43**   Type: **Oral**

Architecture for the implementation of the Fusion FAIR Data Framework

Tuesday 6 July 2021 15:50 (10 minutes)

Currently, largely for historical reasons, almost all fusion experiments are using their own tools to manage and store measured and processed data as well as their own ontology. Thus, very similar functionalities (data storage, data access, data model documentation, cataloguing and browsing of metadata) are often provided differently depending on experiment. The overall objective of the Fair4Fusion project is to demonstrate the impact of making experimental data from fusion devices more easily findable and accessible. The main focus towards achieving this goal is to improve FAIRness of the fusion data to make scientific analysis interoperable across multiple fusion experiments. Fair4Fusion is proposing a blueprint report that aims for a long term architecture for Fusion Open Data Framework implementation.

User stories about searching and accessing data and metadata, and from the perspective of data providers were collected. These use cases present the different perspectives of members of the general public, EUROfusion researchers and data providers that are the main target users of the analyzed scenarios. The basic requirements and user stories have been transformed into a list of functionalities to be fulfilled. These functionalities have been grouped in several general categories: search, visualisation and accessing outputs, reports, user annotation, metadata management, subscriptions and notifications, versioning and provenance, authentication, authorization, accounting, licensing and are related with different FAIR aspects. The collection of requirements and functionalities has been used as a basis for the iterative process of architecture design. We are assuming the use of the ITER Integrated Modelling & Analysis Suite (IMAS) Data Dictionary as a standard ontology for making data and metadata interoperable across the various EU experiments. The resulting architecture of the system consists of 3 main building blocks, namely Metadata Ingests, Central Fair4Fusion Services and Search and Access Services. In the figure we present a simplified version of this high level architecture.
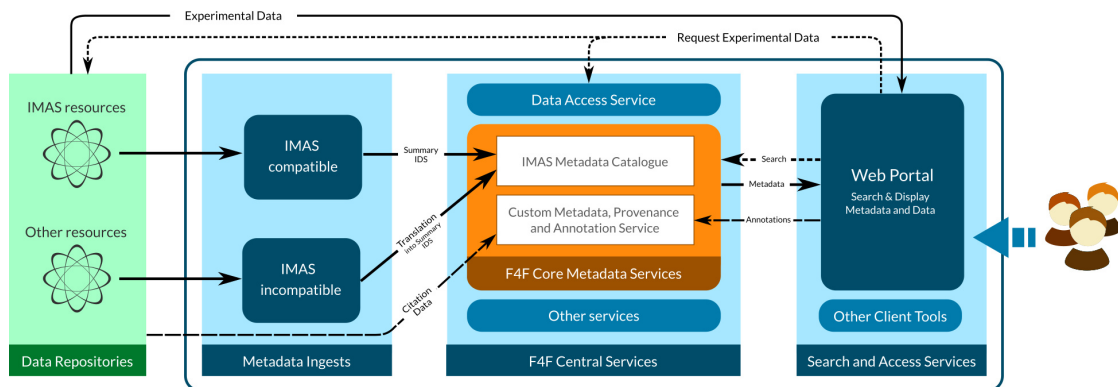

Figure 1: F4F Architecture

Metadata Ingests are the entry point to the system for the metadata produced by experiments. In the proposed design, Metadata Ingests stay within the administration of particular experiments, thus the experiments themselves can filter or amend data before they decide to expose it to the rest of the system. From Metadata Ingests the metadata is transferred to the next block of the system, i.e. Central Fair4Fusion Services. The Core Metadata Services, being the heart of this block and the entire system in general, operate on the IMAS data format, but thanks to the translation components can accept different formats of metadata as input. Central Fair4Fusion Services provide supplementary functionality for specification of data that is not strictly tied to experiments, such as user-level annotations or citations. The last main block of the system is a set of Search and Access Services. It contains all the user-oriented client tools that integrate with the Central Fair4Fusion Services. At this level of the system, key importance is given to the Web Portal that is expected to offer an extensive set of functionalities for searching, filtering or displaying metadata and data managed within the

system.

Speaker's Affiliation

PSNC IBCH PAS, Poznan

Member State or IGO

Poland

Primary author: PLOCIENNIK, Marcin (PSNC)

Co-authors: Mr BOSAK, Bartosz (PSNC IBCh PAS); Dr PALMA, Raul (PSNC IBCh PAS); Mr OWSIAK, Michal (PSNC IBCh PAS); Dr DE WITT, Shaun (UKAEA); Dr GIBBONS, George (UKAEA); Dr CUMMINGS, Nathan (UKAEA); Dr KLAMPANOS, Iraklis (NCSRD); Prof. IKONOMOPOULOS, Andreas (NCSRD); Dr KONSTAN-TOPOULOS, Stasinos (NCSRD); Prof. STRAND, Pär (Chalmers); Dr COSTER, David (MPIPP); Dr IMBEAUX, Frederic (CEA); Dr DECKER, Joan (EPFL); Dr MARTIN, Yves (EPFL); Dr SAUTER, Olivier (EPFL)

Presenter: PLOCIENNIK, Marcin (PSNC)

Session Classification: Database techniques for Information

Track Classification: Database Techniques for Information Storage and Retrieval