Georgios Kakaletris – Diamantis Tziotzios
Athena Research and Innovation Center

OpenDMP Service

Machine actionable open data management planning

CC BY

Overview & Status

OpenDMP: A service for machine actionable data management planning

- **Jointly designed and developed by OpenAIRE and EUDAT**
  - Core: OpenAIRE
  - EEStore: EUDAT
- **Launched implementation August 2017.**
- **Currently at RC stage development**
  - Updates rolled out on a monthly basis

Structural data model elements

- **Data Management Plan**
  - is a collection of *dataset descriptions*, associated with an activity (“project”)
  - may be versioned
  - is exportable to various forms, currently machine readable (xml) and human readable (pdf/openxml)
- **Dataset description**
  - is the description of a dataset according to a set of rules that is the *data set profile*
- **Dataset profile**
  - is the set of rules that describe what dataset descriptions contain and how they are handled
  - contains attributes/fields, behavioral rules (e.g. toggling visibility of fields) and validation rules
  - are linked to *DMP*s for focusing users
- **“Project”**
  - a logical entity that defines an activity that assembles one or more data management plans
  - may map to “grant”, “project”, “initiative” etc (will support attributes that track funders, programmes etc)
- **Users**
  - associated to DMPs and other entities for authorization
  - may map to researchers

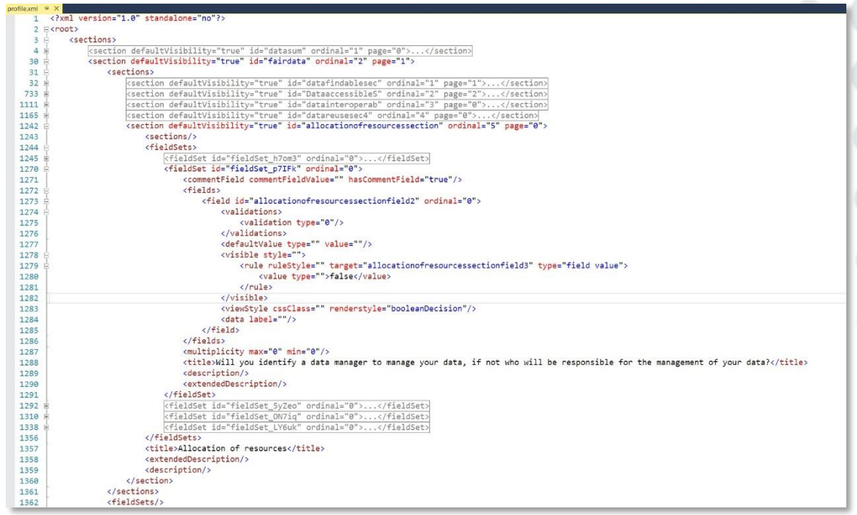
- **Secondary entities**
  - **Repository**: repositories where dataset resides on
  - **Dataset**: external linked datasets
  - **Registry**: catalogues listing the dataset
  - **Service**: services manipulating the dataset under various protocols.
  - **Researcher**: researchers associated with datasets
  - etc

Using OpenDMP Service: Workflow & Roles

- **Admins**
  - select and configure information sources
  - define dataset profiles
- **DMP managers launch DMPs**
  - select appropriate Dataset profiles
  - invite researchers/dataset managers
- **Researchers/Dataset managers**
  - describe datasets according to fixed (systemic) and custom (profile-defined) attributes
- **DMP Managers finalize DMPs**
  - close the editing process
  - publish the DMPs (url, pdf, xml etc)
- **System**
  - ...

Soft (Profile-based) data model elements

- **Dataset Profiles (templates)**
  - Dataset descriptions have little number of fixed structural properties.
  - Profiles define the data model, behavior and appearance of the description of a dataset
  - A dataset profile consists of:
    - Various field types (external lists, text fields/areas, value selectors, composite fields etc)
    - Labels/titles/descriptions
    - Field sets and Sections (grouping elements)
    - Visibility rules (hide/show fields depending on prior choises)
    - Validation rules
      - On per field basis (patterns, multiplicity, existence)
      - On entire dataset description
  - Dataset profiles are linked to DMPs
  - Each dataset description conforms to a dataset profile that is linked to the DMP that the Dataset description belongs to.

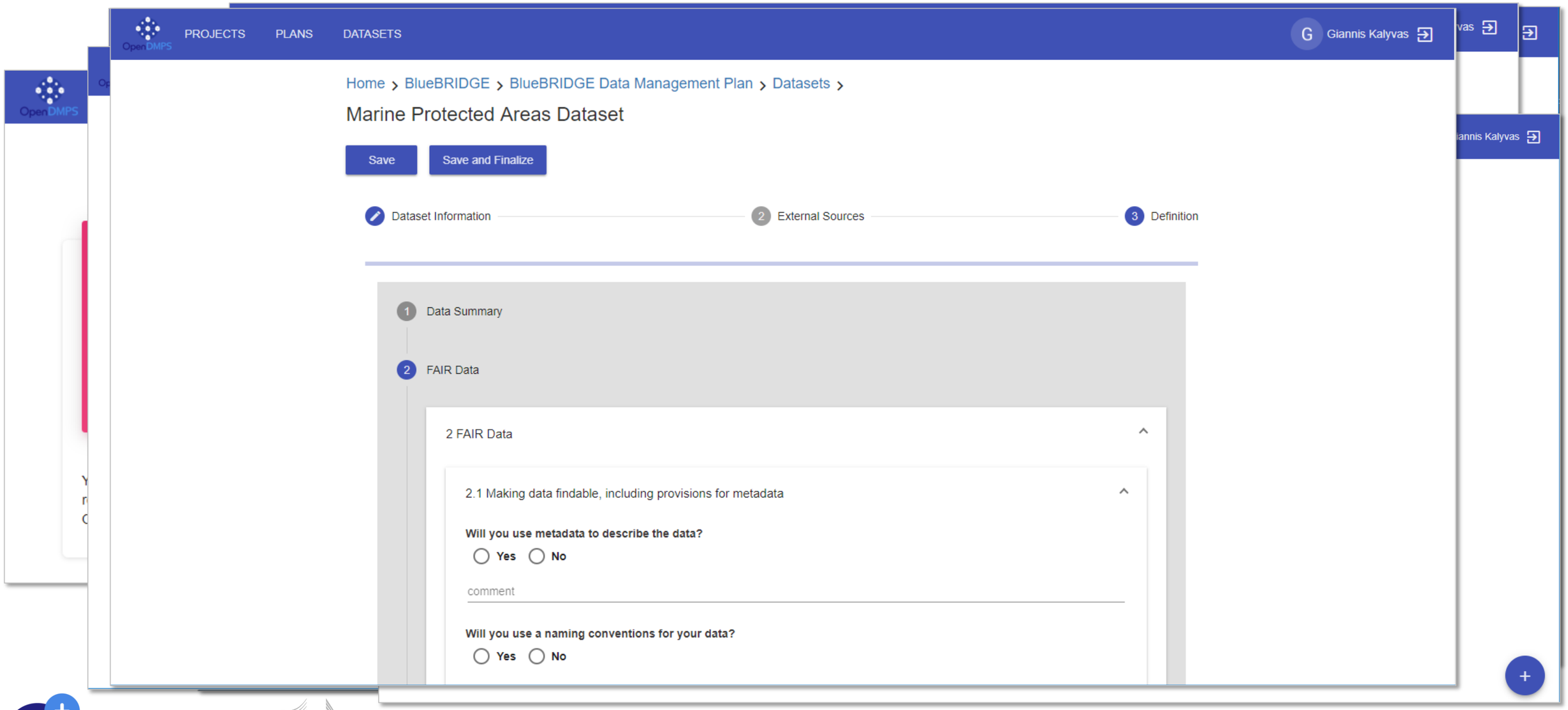


- **DMP Profiles**
  - DMPs have a set of structural properties (fixed data model elements)
  - DMPs may have properties of their own that are configurable via DMP profiles
    - Restricted to simple fields that may be : text, date, term lists (externally provided)

Powering interoperability via…

- **REST API for utilization of core components by external services**
- **Open architecture**
  - for integrating external information providers to the DMP entities
    - EEstore abstracts over the behavior and performance of external providers
  - M2M communication data repositories, registries and other dataset handling services
    - E.g. depositing and verification of dataset(s) and DMP(s)
- **Machine readable DMP export**
  - currently proprietary format
- **Flexible data model that can power different interoperability cases**
  - Currently verified for H2020 – template is provided.

... a few screenshots

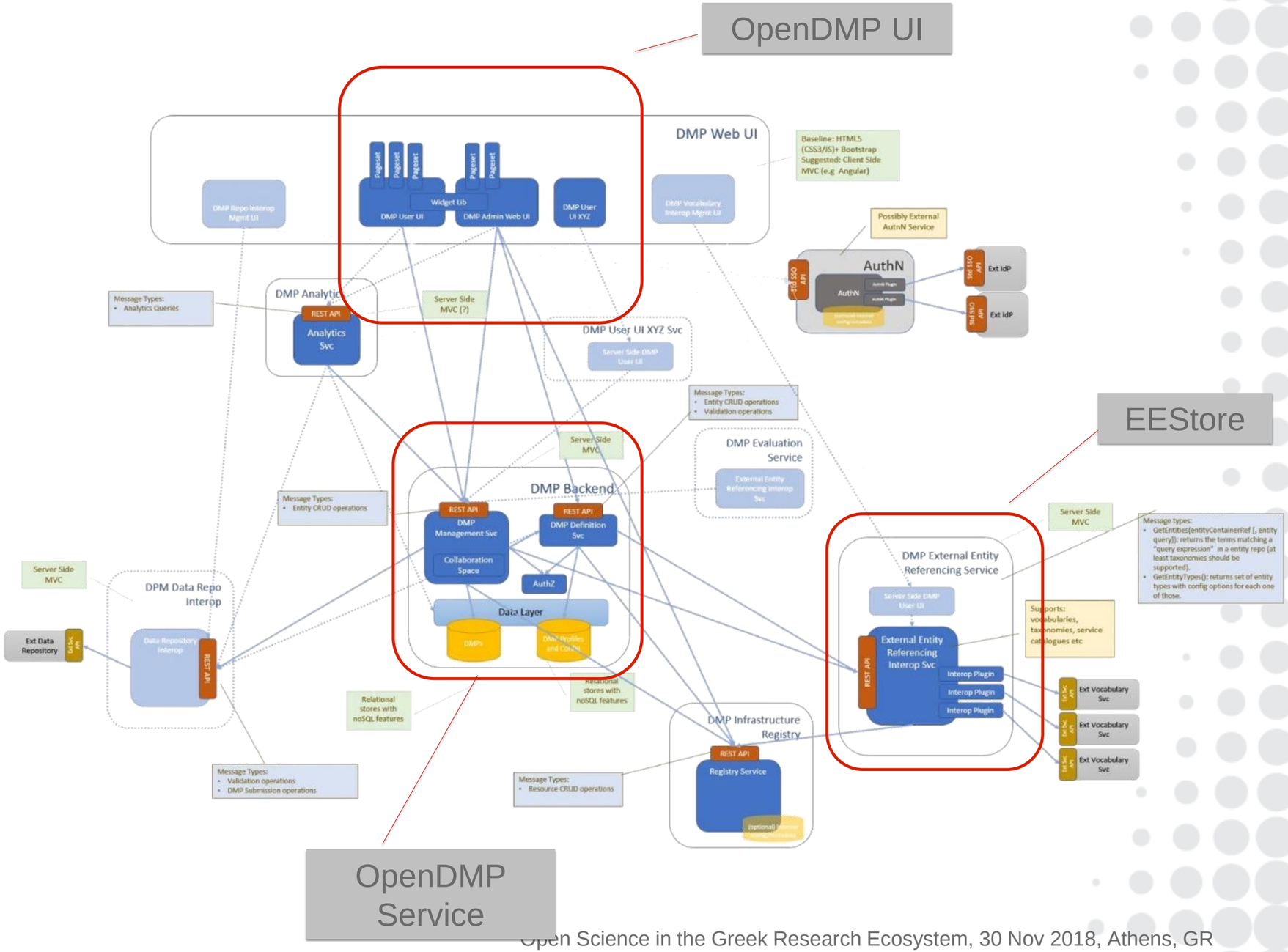


Technology & License

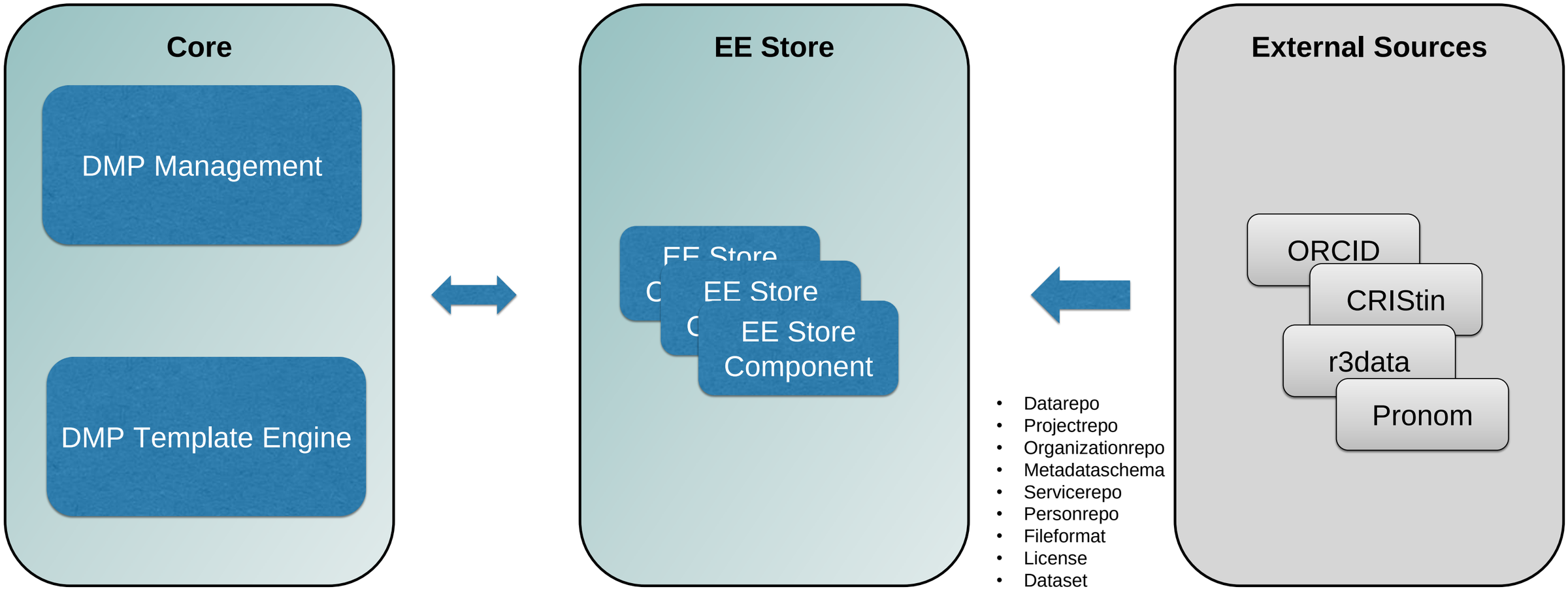
- **Latest technology web application**
- **Frontend:**
  - Angular / typescript
  - Material Design
  - Responsive
- **Backend:**
  - REST backend services
  - Java/Spring
  - PostgreSQL RDBMS
- **License:**
  - Unrestricted Open Source (to be announced soon)

Architecture

- **OpenDMP Core**
  - DMP Service: Core tools for editing, management and administration of DMPs/Dataset descriptions.
  - DMP UI: End-user UI for editing DMPs and Dataset descriptions
  - DMP Admin UI: Administrative UI and dataset profile definition UI
- **EEStore: bridge to various providers of data for the OpenDMP**
  - Core bridge component
  - Several adapters for external providers



Key components (currently implemented)



- Datarepo
- Projectrepo
- Organizationrepo
- Metadataschema
- Servicerepo
- Personrepo
- Fileformat
- License
- Dataset
- Mdrepo
- Taxonomy

Why OpenDMP

- Is backed by **collaboration** that activates on several areas of EU Open Science paradigm
- Can be exploited under **different models of service provisioning**
- Closely **follows standards** and well-known specifications
- Has been **implemented from scratch** taking into account trends on Actionable DMPs, Open Science and experiences from existing tools
- … and for several technical and functional reasons!

Is OpenDMP a questionnaire engine ?

The reply is “Definitely No”: The dataset profile (questionnaire) engine is a key, powerful, element of OpenDMP but it is only a small part of it. Yet , even its questionnaire engine, offers essential facilities for DMP handling.

- **OpedDMP streamlines the handling of a DMP, not generic questionnaires: From the initiation to publication time, the tools and the underlying data model facilitate navigation, discovery and reuse of dataset descriptive information.**
- **OpenDMP service is designed to handle workflows on actionable DMPs, where the questionnaire is a small step of the process. Metadata and tools that allow validation of information provided and interacting with external Data-oriented services are fundamental objectives behind the design of the system.**
- **OpenDMP enables collaborative DMP composition that will surpass the capabilities of any generic tooling .**

Plans - indicative list

- Extend authentication mechanism to support EGI AAI (finished)
- Finalize DMP exports (testing)
- Formalize API of various components (in progress)
- Implement a “mapping” mechanism / semantics
  - Essential for DMP analysis and interoperability
  - RDA DMP Common Data Model is required for launching the activity
- Implement statistics and basic reporting mechanism (in progress)
- Extend authorization capabilities (designed)
- Extend collaboration capabilities (designed)
- Implement basic services validation (planned)
  - E.g. check endpoint access and spec conformance
- Launch analytics subsystem implementation (planned)
  - Depends on mapping mechanism

Always working on:

- Maintaining beta public release and follow up with monthly updates (fixes & usability improvements)
- Refining DMP dataset profile language description
- Linking new external sources as data providers for dataset descriptors
- Improving various components look ad behavior
- Improving performance

Source code: https://gitlab.eudat.eu/dmp/

Beta testing site: http://devel.opendmp.eu/

Thank you!

George Kakaletris

gkakas@di.uoa.gr

Skype: gkakas