# nffa.eu PILOT 2021 2026

## DELIVERABLE REPORT

WP16 JA6 – Implementing FAIR data approach within NEP

D16.1 Design of the service platform

Due date M9



This initiative has received funding from the EU's H2020 framework program for research and innovation under grant agreement n. 101007417, NFFA-Europe Pilot Project



#### **PROJECT DETAILS**

PROJECT ACRONYM	PROJECT TITLE
NEP	Nanoscience Foundries and Fine Analysis - Europe PILOT
GRANT AGREEMENT NO:	FUNDING SCHEME
101007417	RIA - Research and Innovation action
START DATE	_
01/03/2021	

WORK PACKAGE DETAILS			
WORK PACKAGE ID	WORK PACKAGE TITLE		
WP16	JA6 – Implementing FAIR data approach within NEP		
WORK PACKAGE LEADER			
Giuseppe Piero Brandino (eXact lab)			

DELIVERABLE DETAILS			
DELIVERABLE ID	DELIVERABLE TITLE		
D – D16.1	Design of the service platform		

#### DELIVERABLE DESCRIPTION

This deliverable presents the initial design of the service platform and a description of the different elements that compose it, together with an explanation of how they are planned to be connected to each other.

DUE DATE	ACTUAL SUBMISSION DATE
M9 30/11/2021	26/11/2021

#### AUTHORS

Rossella Aversa (KIT) Thomas Jejkal (KIT) Mirco Panighel (CNR) Irene Modolo (Elettra) Elda Osmenaj (CNR) Giuseppe Piero Brandino (eXact lab)





#### PERSON RESPONSIBLE FOR THE DELIVERABLE

Dr. Rossella Aversa (KIT)

#### NATURE

- $\times$ R - Report
- P Prototype
- DEC Websites, Patent filing, Press & media actions, Videos, etc
- 0 Other

#### DISSEMINATION LEVEL

$\boxtimes$	P - Public	
	PP - Restricted to other programme participants & EC:	(Specify)
	RE - Restricted to a group	(Specify)
_		

CO - Confidential, only for members of the consortium 





#### REPORT DETAILS

ACTUAL SUBMISSION DATE 26/11/2021

#### NUMBER OF PAGES

15 (right-click and select "update the field")

#### FOR MORE INFO PLEASE CONTACT

Dr. Rossella Aversa, Karlsruhe Institute of Technology Hermann-von-Helmholtz-Platz 1 76344 Eggenstein-Leopoldshafen Germany email: rossella.aversa@kit.edu

VERSION	DATE	AUTHOR(S)	DESCRIPTION / REASON FOR MODIFICATION	STATUS
1	18/10/2021	R.Aversa	First draft	Draft
2	02/11/2021	M.Panighel	Services/interfaces	Draft
3	09/11/2021	E.Osmenaj, I.Modolo	Data Management, SDW	Draft
4	16/11/2021	P.Brandino	SSO, Database	Draft
5	23/11/2021	R.Aversa, I.Modolo, M.Panighel, E.Osmenaj	Restructuring of the document	Final version

CONTENTS		
Introduction	6	
Elements of the infrastructure		
1. Basic Infrastructure	7	
1.1 Single-Sign-On (SSO)	7	
1.2 Database (DB)	7	
1.3 Data Collaboration Platform	8	
1.4 Data Repository	8	
1.5 MetaStore	9	
1.7 Data Stewardship Wizard (DSW)	10	
1.8 Electronic Laboratory Notebook (ELN)	10	
1.9 Virtual Access (VA) Services	10	





2		Additional elements of the infrastructure	11
	2.1	Custom Metadata Ingest Interface	11
	2.2	Mapping Tool	11
	2.3	Mapping Tool Interface	12
	2.4	Data Converter	12
Use Cases			12
	Use	e Case 1: Raw Data generation with Metadata	12
	Use	e Case 2: Data Analysis with Metadata	13
	Use	e Case 3: Data Publication	14
References		15	



# INTRODUCTION

This deliverable presents the initial design of the infrastructure for the NFFA-Europe Pilot (NEP). The infrastructure is planned to consist of diverse elements for the Data and the Metadata Management, as well as different services (in the frontend, in the backend, and for Virtual Access) which will be gradually developed and integrated in a seamless way.

We distinguish between the *basic* elements, which are essential parts of the infrastructure planned in the NEP proposal, and *additional* elements which were not initially planned but might improve the interconnections and facilitate the Research Users, in case they will be developed as output of the scouting activities of the Task 16.4 of the Joint Activity 6 (Work Package 16).

The elements of the infrastructure will be connected to each other and will be accessed by users or by other services thanks to interfaces. Figure 1 shows a global view of the elements of the infrastructure, including interfaces, and how they will be interconnected. The additional elements are shadowed and connected by dotted lines. We note that this is the initial design, and that the final infrastructure may be subject to variations.

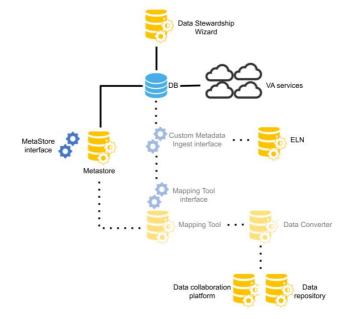


Figure 1: Overview of the initial design of infrastructure for NEP. The additional elements, which will be optionally developed and integrated, are shadowed and connected by dotted lines.

The document is structured as follows: Section 1 describes the elements planned to be included in the basic infrastructure, while Section 2 presents the additional elements which will be optionally developed and integrated. Finally, three explicative use cases are described at the end of the document. Throughout the document, the terms included and defined in the NFFA-Europe glossary [12] are written with capital letters.



 $\mathbb{PILOT}^{\frac{2021}{2026}}$ 

# ELEMENTS OF THE INFRASTRUCTURE

## 1. Basic Infrastructure

## 1.1 Single-Sign-On (SSO)

The authentication and authorization will be in charge of the Identity Provider service, based on the open source package *keycloak* [1]. This service will provide all the necessary parts to manage registration, authentication and roles of all Research Users, and it is one of the central services of the NFFA-Europe IT infrastructure, providing the Single-Sign-On (SSO) system for almost all the services that will require authentication. This means that the services interfaced with the Identity Provider will be accessed using the same credentials, while keeping the possibility for a user to have different roles in different services.

The minimal list of services that will be interfaced to the Identity Provider are:

- Data Collaboration Platform
- MetaStore
- Data Stewardship Wizard
- Virtual Access services

and other services in the frontend and backend (e.g., Proposal Management system, Monitoring and accounting service, NFFA-Europe Portal) whose description is not covered in this document. The SSO will cover the additional elements of the infrastructure in case they will be developed, for example the Custom Metadata Ingest interface [Section 2.1].

## 1.2 Database (DB)

The NEP Database (DB) is one of the central elements of the NEP IT infrastructure. It is a collection of relational and non relational databases that can be accessed by means of endpoints in the backend service. The user database is relational and logically separated by the other databases, and all the interactions with these databases are secured behind authenticated endpoints. The database containing the NEP technique catalog, the instrument catalog and the proposal data are non relational databases, with authenticated endpoints for all sensitive data. Sensitive data, such as the content of a proposal, are accessible only by the owner or by the evaluation committee with the sole purpose of erogating NEP services.

Other databases or extensions of the existing ones are planned to support advanced management and data services.

The minimal list of services planned to have access to the DB are:

- MetaStore
- Data Stewardship Wizard
- Virtual Access Services

and other services in the frontend and backend (Proposal Management system, Monitoring and accounting service, NFFA-Europe Portal) whose description is not covered in this document. The NEP DB will integrate the additional elements of the infrastructure in case they will be developed, for example the Custom Metadata Ingest Interface [Section 2.1].





#### 1.3 Data Collaboration Platform

With their NFFA credentials, Research Users will have automated access via the SSO to their personal account on NFFA Datashare [2], a file share and collaboration platform based on the Nextcloud [3] open-source software and hosted on servers under NFFA-Europe custody. This platform offers Research Users a secure means to store, retrieve, and access Raw Data and Analysed Data in every file format, collaborating in real time with NFFA-Europe Instrument Scientists and other Team Members. Research Users can store their Research Data on the platform to have a secure backup, access them from every device, invite Team Members to view, download and collaborate in real time on files and folders.

Laboratories offering user access are invited to use NFFA Datashare or their institutional cloud service to store and share Research Data in a secure way, without the risk of losing data by placing them on portable data storage devices. Heads of Laboratories can require the creation of an ad hoc Laboratory account on NFFA Datashare. From the Laboratory account, the Head of Laboratory can share the Raw Data to the NFFA Datashare personal account of the Team Members.

The folders or the partition containing the Raw Data of the NEP Research Users on the Laboratory data acquisition local machines can be automatically synchronized on NFFA Datashare thanks to the Nextcloud desktop client [4], available for Windows, macOS, FreeBSD and Linux.

#### 1.4 Data Repository

To deposit and ensure Open Access to peer-reviewed scientific publications and to the underlying Research Data (Publication Data) and Metadata, complying with the regulations of Horizon 2020, repositories that meet EOSC requirements must be used. A suggested Data Repository is Zenodo [5], hosted at CERN in Geneva (Switzerland). Alternatively, an overview of compliant subject-specific repositories can be found at Re3data [6]. Deposited items must provide at least the following descriptive metadata:

- Author(s)
- Funding: Horizon 2020
- Grant project name, acronym and number: *Nanoscience Foundries and Fine Analysis* - *Europe/PILOT, NEP, 101007417*
- Proposal ID
- Licensing terms

The Data Repository of choice must provide Persistent Identifiers (PID) to the deposited items, so that they can be cited, linked and tracked. Open Access via the repository must be ensured under the latest available version of the Creative Commons Attribution International Public License (CC BY) or Creative Commons Public Domain Dedication (CC 0) or a licence with equivalent rights.

Files may also be deposited under restricted, closed, or embargoed access. *Restricted* access allows Research Users to set conditions under which they grant access to and approve sharing of the files. Files deposited under *closed* access (for archival purposes only) are protected against unauthorized access at all levels [7]. Content under *embargoed* access is restricted until the end of an embargo period set by the Research User; then it will automatically become publicly available. In any case, a DOI will be assigned to the deposited files.





#### 1.5 MetaStore

The MetaStore [8] will offer a generic metadata repository and metadata schema registry, providing Metadata versioning, search, and visualisation. From the technological perspective, it will be accessible via *RESTful* service interfaces. It will enable users to register Metadata Schemas in any supported format, e.g., XSD, JSON Schema and at a later time RDF, and to store Metadata Documents which will be automatically validated at upload time against the corresponding registered Metadata Schema.

The MetaStore will be populated with Metadata Documents linked to scientific Datasets resulting from transnational access Experiments, as well as with the administrative Metadata Documents regarding the proposal and the Research User information.

As stated in the NEP Data Management Plan (DMP) [9], Research Users will be asked to provide access to the data produced for their scientific publications, including descriptive Metadata. At publication time, the Research User will not need to collect and check the Metadata: using MetaStore at each step of the Experiment (e.g., Sample Preparation, Measurement, Data Analysis) will ensure to already have complete validated Metadata, linked to the relevant Raw and/or Analysed Data.

The different Metadata Documents will be co-referenced using internal IDs. In this way, the MetaStore will give the possibility to match proposals with publications and the corresponding Data Provenance, offering a clearer insight on the scientific process from the project idea to the research output.

The MetaStore will be connected to the SSO [Section 1.1]. Every access to the Metastore that requires to read, write or modify metadata documents, as well as to retrieve relevant information, such as Research User and Proposal Data, will be authenticated.

In the basic scenario, the Research Users will access the MetaStore directly, using a Graphical User Interface [Section 1.6] and, according to their authorization level, they will be able to list all the available Metadata Documents or to search, retrieve, add and edit the Metadata Document describing a specific Measurement. It is worth mentioning that the registration of Metadata Schemas to the MetaStore will not be allowed to the Research Users, but only to the Data Curators.

In the ideal scenario, an additional indirect access to the MetaStore will be performed by the Mapping Tool [Section 2.2], which will trigger the registration of a Metadata Document describing a performed Measurement, or by the Custom Metadata Ingest Interface [Section 2.1], which will query the MetaStore to retrieve a specific Metadata Document and all its versions. This scenario can be very specific depending on the involved Instruments and scientific domain. The Metadata Schemas will be developed (or adopted from existing standards, if any) for each specific case in an incremental process.

#### 1.6 MetaStore interface

This Graphical User Interface to the MetaStore allows Research Users and Data Curators to visualise (according to the access permissions) the registered Metadata Schemas and Metadata Documents together with all their previous versions (if any), to search for specific Metadata Schemas or Metadata Documents according to keywords, and to manually register, edit or import new Metadata Documents or Metadata Schemas. It is important to note that only the Data Curators are allowed to register and edit Metadata Schemas.

The MetaStore interface is accessed via the SSO [Section 1.1]; in this way, Research Users can see and edit only the Metadata Documents according to their access permissions, which are set by the authorization group they belong to and by the sharing options assigned to each Metadata Document (provided by the Access Control document). In the basic case in





which nothing is shared to "public", Research Users can select only the Metadata Documents related to the proposals they belong to.

In case of manual registration of a new Metadata Document, the Research User information and the Proposal details are automatically taken from the DB, once the Proposal ID has been selected from a dropdown menu.

#### 1.7 Data Stewardship Wizard (DSW)

The Data Stewardship Wizard (DSW) is an open source tool that supports the creation of DMPs in a smart way suitable for everyone (from Research Users to Data Stewards). It consists of an open source web questionnaire tool [10], based on a so-called Knowledge Model, and a system to maintain Knowledge Models. Knowledge Models refer to a tree-like structure of chapters, questions, answers and references that allow to collect the relevant information related to data management within a Laboratory. Knowledge Models can be adjusted by Data Stewards and then used by Research Users. By filling the questionnaire into a formatted document (DMP). It also provides context-dependent guidance through text, links, external resources and integrations to important services such as FAIRsharing.org and the central DB to get the list of Institutions and Techniques. For the NEP community, a self-managed DSW instance integrated with the SSO and the DB has been installed, and a personalized Knowledge Model has been defined for all the Laboratories who have to write a DMP.

#### 1.8 Electronic Laboratory Notebook (ELN)

An Electronic Laboratory Notebook (ELN) is a computer program designed to replace paper Laboratory notebooks. It is used by Instrument Scientists and Research Users to document research, procedures, workflows performed during an Experiment and typically related to a particular Instrument. Many ELNs are available online. Among them, the free open source ELN *eLabFTW* [11] is the recommended one. A dedicated self-managed instance will be installed on the servers managed by NFFA-Europe, at the disposal of the Laboratories wishing to use the service. Heads of Laboratories can require the creation of an ad hoc Laboratory account. From the Laboratory account, the Head of Laboratory can grant access to Instrument Scientists and Research Users. The integration of other ELNs of widespread use could be taken into account at a later stage, if needed.

In the basic scenario, Instrument Scientists and Research Users will manually insert in the ELN any information related to the Experiment. In the ideal scenario, the Metadata related to the Experiment could be sent to the ELN by the Research User using the Custom Metadata Ingest Interface [Section 2.1].

### 1.9 Virtual Access (VA) Services

Virtual Access (VA) services will be innovative simulation services, databases, machine learning services, Data and Metadata services, offered as cloud services and seamlessly integrated in the infrastructure for NEP. Each partner will operate its own infrastructure, made of one or more installations. VA will be granted only to identified Research Users. All involved Access Providers will offer identified User Access to the VA services under their own schemes based on institutional policies.

The initial list of VA services which will be offered includes: the MetaStore, the Materials Cloud, the Trieste Advance Data Services (TriDAS), a user-friendly interface for Magnetic





Resonance image reconstruction and contrast prediction, a Nuclear Magnetic Resonance spectral database, a Machine Learning data analytics and simulation platform to predict complete data maps for nano-morphological features produced on various types of materials upon pulsed laser manufacturing, and a Materials Modeling Toolbox Service to perform simulations.

## 2. Additional elements of the infrastructure

### 2.1 Custom Metadata Ingest Interface

This Graphical User Interface allows to collect all Metadata related to a specific Measurement (namely information on Sample Preparation, Instrument configuration and Proposal information) and to send them to the Mapping Tool [Section 2.2] for their mapping into an XML document and registration to the MetaStore. Within this interface the Research Users can select the Metadata Schema from the ones available in the MetaStore and fill the required Metadata from scratch, or retrieve a Metadata Document previously registered in the MetaStore and update it by adding new Metadata (e.g., retrieve a Raw Metadata Document and add Data Analysis Metadata). The Metadata fields shown in the interface are determined by selecting a specific Metadata Schema available in the MetaStore. In case the schema is not available for a specific Technique, the Head of the Laboratory or the Research Users can submit a request for a new one using a template (e.g. Excel table) in which list and structure the desired fields they need.

Metadata can be inserted manually, through the interface itself, or programmatically, through APIs. The interface checks that the mandatory Metadata is inserted, in particular the unique ID identifying the proposal.

The Custom Metadata Ingest Interface will not be covered by the SSO, but it will be authenticated: Heads of Laboratories can require the creation of an ad hoc Laboratory account, which will be used by all the Research Users. Thus, when accessing the interface, the Research User can select the proposal ID from a dropdown menu containing only the Proposals involving Instruments related to the Laboratory. Optionally, this interface will allow Research Users to send and insert these Metadata (or a selection of them) to a supported Electronic Laboratory Notebook [Section 1.7].

## 2.2 Mapping Tool

The Mapping Tool is planned to be a local tool to map metadata into XML documents, according to specific Metadata Schemas. The mapping will be triggered by the Research User in the Custom Metadata Ingest Interface, with the assistance of the Head of Laboratory. The tool will take as input the Metadata related to the Research User and to the Data from the Custom Metadata Ingest Interface [Section 2.1] and the corresponding XSD Metadata Schema, and will create two documents:

- an XML Metadata Document, by mapping the Metadata related to the Data following the XSD structure
- an "access control" document, by mapping the Research User information. This document will contain the authorization rights assigned to the Metadata Document in a structured way needed by the MetaStore.

The Mapping Tool will trigger the registration of a Metadata Document to the MetaStore. The generated outputs will be a XML metadata document, which will be registered into the MetaStore [Section 1.5], and the related "access control" document used by the MetaStore



 $\mathbb{PILOT}^{2021}_{2026}$ 

interface [Section 1.6]. The XML Metadata Document will be also provided as input for the Data Converter Tool [Section 2.4] in order to generate a NeXus/HDF5 file.

### 2.3 Mapping Tool Interface

This interface will act primarily as a validation of the inputs needed for the Mapping Tool [Section 2.2]. Specifically, it will check that a valid Metadata Document is provided, that the corresponding Metadata Schema is selected among the ones already registered in the MetaStore, and that the mandatory Metadata are present, in particular the Research User information and the unique ID identifying the Proposal.

### 2.4 Data Converter

The Data Converter will be an offline tool to generate a NeXus/HDF5 file from some data and their related metadata, according to a specific NeXus application definition. It will accept as input the Data (in a suitable file format), the XML Metadata Document of the related Metadata, as generated by the Mapping Tool [Section 2.2], and the NeXus application definition defining the structure of the output NeXus/HDF5 file.

## **USE CASES**

#### Use Case 1: Raw Data generation with Metadata

**Situation:** The Research User has submitted a Proposal via the NFFA-Europe Portal. This Proposal is accepted and the involved Team Members access the Laboratory to perform some Measurements.

#### Workflow:

- 1. The Research User travels to the Laboratory, performs the Measurement(s) as planned, and generates Raw Data.
- 2. If the folder or partition containing the Raw Data of the Research User on the Laboratory data acquisition computer is synchronized on a Data Collaboration Platform [Section 1.3] (via the Nextcloud desktop client if NFFA Datashare is used), this folder is shared by the Head of Laboratory to the Research User on the first day of access. The Research User logs in the Data Collaboration Platform via the SSO and finds the Raw Data there. Alternatively, the Research User uploads and stores the Raw Data (in any preferred format) on a Data Collaboration Platform to have a secure backup and collaborate easily.
- 3. Raw data that are mature enough for permanent storage, can be deposited in a Data Repository with their preferred level of access (closed/restricted/open). The data are not necessarily published, just safely archived.
- 4. BASIC SCENARIO:
  - a. Team Members retrieve the empty Metadata Document describing a specific Measurement according to the Metadata Schema. They edit the Metadata Document filling the required Metadata, then register it into the MetaStore. At registration time, the validation against the Metadata Schema is performed. The Metadata Document must contain the location or the PID associated with the deposited Raw Data in the Data Repository.
- 5. IDEAL SCENARIO:





- a. Through the Custom Metadata Ingest Interface [Section 2.1], the Research User fills the required Metadata: the user information, the Proposal ID (selected from a dropdown menu), the Instrument information, the settings during the Measurement, and the Raw Data description, including the location where the Raw Data is stored in the Data Collaboration Platform or, preferably, the PID associated to the deposited Data on the Data Repository. In the course of that, information on authorization are collected defining who is able to later on access the Metadata and with which permissions. The Research User, with the assistance of the Head of Laboratory, can then trigger the mapping of the Metadata into a XML document, thanks to the Mapping Tool [Section 2.2]. In addition to that, the information on authorization is mapped into a structured "access control" document.
- b. The Raw data and the corresponding XML metadata document can be given as input to the Data Converter [Section 2.4], in order to generate the Raw Data NeXus file, which will be stored in the Data Collaboration Platform or deposited in a Data Repository [Section 1.4] with the preferred level of access (closed/restricted/open) and associated with a PID. In this case, the XML Metadata document must be updated, adding the location where the Raw Data NeXus file is stored or, preferably, the PID associated with the deposited Data in the Data Repository. In both cases, at this point the XML Metadata Document as well as the "access control" document will be registered in the MetaStore [Section 1.5] using the Mapping Tool [Section 2.2].

#### Use Case 2: Data Analysis with Metadata

**Situation:** The use case "Raw Data generation with Metadata" has been successfully completed. The Research User wants to perform the Data Analysis on the Raw Data generated.

#### Workflow:

- 1. The Research Users log in the Data Collaboration Platform [Section 1.3] via SSO [Section 1.1] and download the selected Raw Data on which they want to perform the Analysis. The Research Users perform the Data Analysis using the appropriate Software, and generate Analysed Data.
- 2. Analysed Data and other Data that are mature enough for permanent storage, can be deposited in a Data Repository with their preferred level of access (closed/restricted/open). The data are not necessarily published, just safely archived.
- 3. BASIC SCENARIO:
  - a. Team Members retrieve the Raw Metadata Document previously filled. They edit the Metadata Document filling the required Metadata, then register it into the MetaStore. At registration time, the validation against the Metadata Schema is performed. The Metadata Document must contain the location or the PID associated with the deposited Analysed Data in the Data Repository.
- 4. IDEAL SCENARIO:
  - a. Through the Custom Metadata Ingest Interface [Section 2.1], the Research User selects the XML Metadata Document (previously registered in the MetaStore) describing the Raw Data which was analysed, and appends the Data Analysis Metadata. In the course of that, information on authorization are collected, defining who is able to later on access the Metadata and with which permissions.





- b. The Research User can then trigger the mapping of the Metadata into a XML document, thanks to the Mapping Tool [Section 2.2]. In addition to that, the information on authorization is mapped into a structured "access control" document.
- c. The Analysed Data and the corresponding XML Metadata Document could be given as input to the Data Converter [Section 2.4], in order to generate the Analysed Data NeXus file, which will be stored in the Data Collaboration Platform. If desired, the Analysed Data and/or the Analysed Data NeXus file can be deposited in a Data Repository [Section 1.4] with the preferred level of access (closed/restricted/open) and associated with a PID. At this point, the Analysed Data NeXus file is stored, or preferably the PID associated with the deposited Data in the Data Repository, and can be registered in the MetaStore [Section 1.5], together with the "access control" document, using the Mapping Tool [Section 2.2]. Internal identifiers in the MetaStore will link the Analysis Metadata Document to the Raw Metadata Document, in order to keep the provenance chain.

#### Use Case 3: Data Publication

**Situation:** The use case "Raw Data generation with Metadata" and/or the use case "Data Analysis with Metadata" have been successfully completed. The Experiment ends up to a peer-review publication and the Research User wants to publish the Research Data needed to validate the published results (Publication Data).

#### Workflow:

- 1. If not done yet, the Research User logs in the Data Collaboration Platform [Section 2.1] via SSO [Section 1.1], downloads the selected Raw Data and/or Analysed Data (strongly recommended in NeXus or any other open format) needed to validate the published results, and deposit it in a Data Repository with open access.
- 2. BASIC SCENARIO: Team Members retrieve the empty Metadata Document related to the Metadata Schema for Publication Data (for example, DataCite). They edit the Metadata Document filling the required Metadata, then register it into the MetaStore. At registration time, the validation against the Metadata Schema is performed. The Metadata Document must contain the DOI of the scientific paper and the PID associated with the deposited Publication Data in the Data Repository.
- 3. IDEAL SCENARIO: Through the Custom Metadata Ingest Interface [Section 2.1], the Research User selects the XML Metadata documents describing the relevant data. The Research User can then trigger the mapping of the metadata to the Metadata Schema for Publication Data (for example, DataCite), including the DOI of the scientific paper and the PID of the Publication Data, thanks to the Mapping Tool [Section 2.2]. At this point, the Publication Metadata Document can be registered in the MetaStore [Section 1.5]. Internal identifiers in the MetaStore will link the different Metadata Documents, in order to keep the provenance chain.





## REFERENCES

- [1] www.keycloak.org
- [2] https://datashare.nffa.eu
- [3] <u>https://nextcloud.com</u>
- [4] https://nextcloud.com/clients
- [5] <u>https://zenodo.org</u>
- [6] <u>https://www.re3data.org</u>
- [7] https://about.zenodo.org/policies

[8] Prabhune, A., Stotzka, R., Sakharkar, V. *et al.* MetaStore: an adaptive metadata management framework for heterogeneous metadata models. *Distrib Parallel Databases* **36**, 153–194 (2018). <u>https://doi.org/10.1007/s10619-017-7210-4</u>

[9] NEP Deliverable D1.1: Data Management Plan (DMP), Official current version, https://datashare.nffa.eu/index.php/s/SrcdWyCYsSgGDb3

[10] <u>https://ds-wizard.org/</u>

[11] https://www.elabftw.net/

[12] https://www.nffa.eu/apply/data-policy/glossary

