nffa-di

Piano Nazionale di Ripresa e Resilienza

WP3 third meeting:

OFED & pipelines update & next roadmap

Federica Bazzocchi

RIT-Area Science Park

Trieste, 28.05.2024

CC BY

AREA SCIENCE PARK

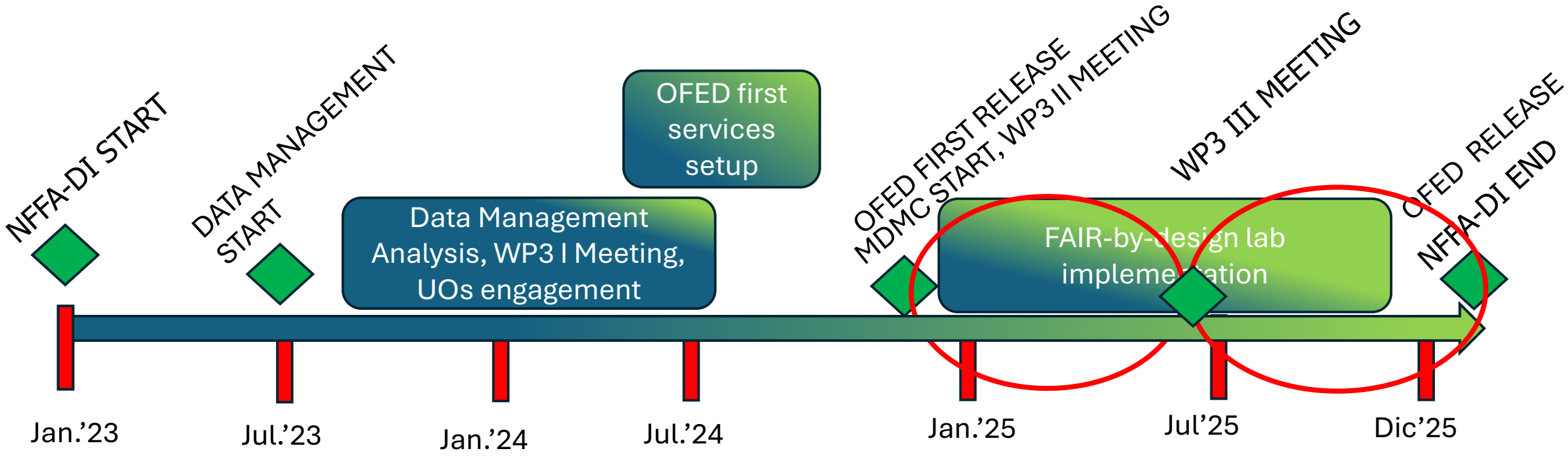
AGENDA

- WP3 state of art:
  - ❑ OFED:
    - Breif recap
    - Services in production
    - Services in test/develpment
  - ❑ FAIR-by-design pipelines:
    - Some numbers and remarks
    - What has been done so far
    - What has to be done yet
  - ❑ Possible Strategies for Next Months:
    - Look at your «similar»»
    - II MDMC edition
    - New tool

Foto di Brian Penny da Pixabay

AREA SCIENCE PARK

Timeline


NFFA-DI START
DATA MANAGEMENT START
Data Management Analysis, WP3 I Meeting, UOs engagement
OFED first services setup
OFED FIRST RELEASE MDMC START, WP3 II MEETING
FAIR-by-design lab implementation
WP3 III MEETING
OFED RELEASE NFFA-DI END
Jan.'23
Jul.'23
Jan.'24
Jul.'24
Jan.'25
Jul'25
Dic'25

OFED: Overarching Fair Ecosystem for Data

… more than a datalake

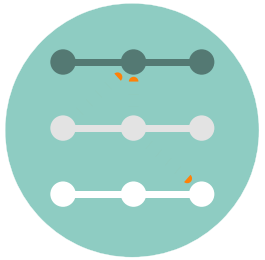
Overarching

Fair

Ecosystem

Data

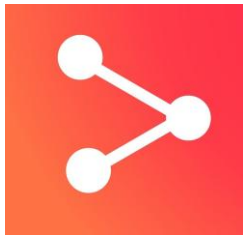
access



plan

store & collect

- Tools and interoperable services
- Find and access metadata engine
- Innovative software through virtual access

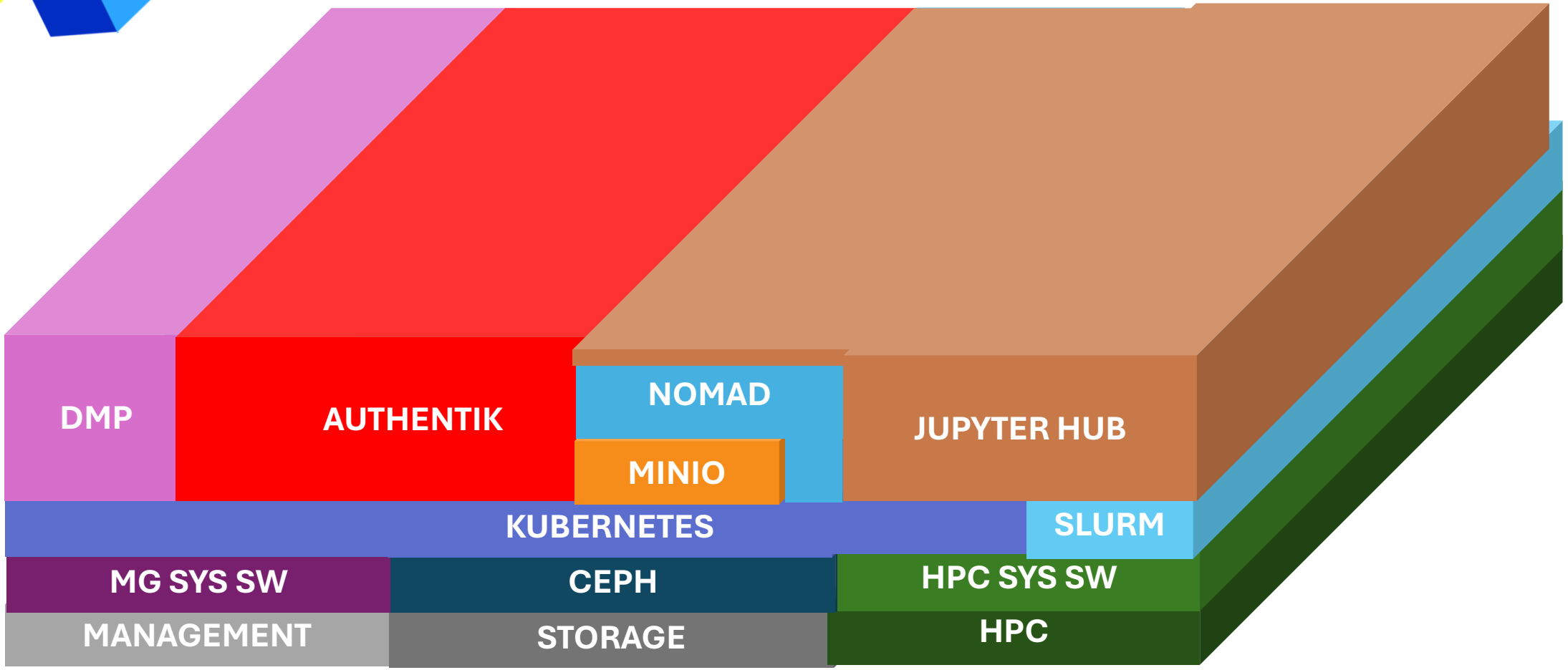


share

analyze

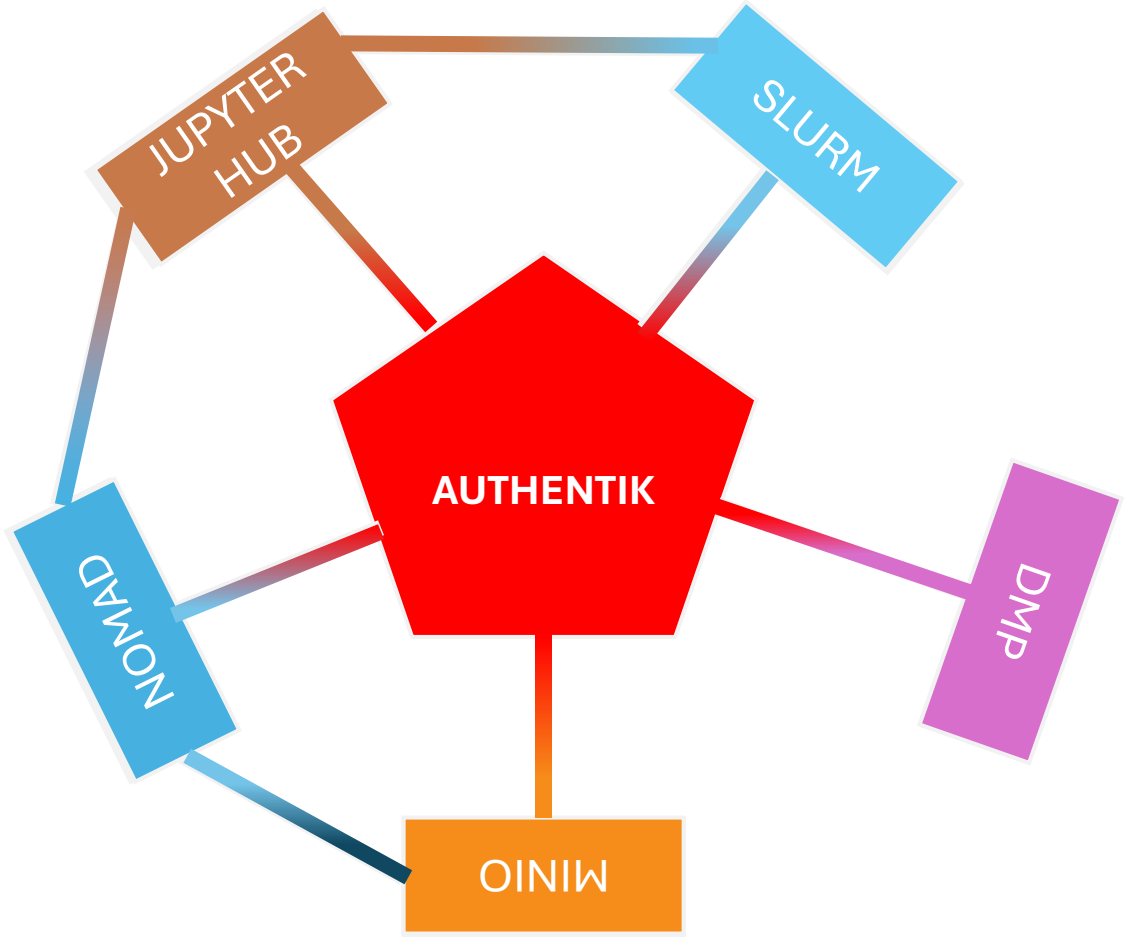
AREA SCIENCE PARK

OFED Specific SW Services



OFED Specific SW Services

Authentik is a CORE SW of OFED Structure because AUTHENTIKATION is required by each service



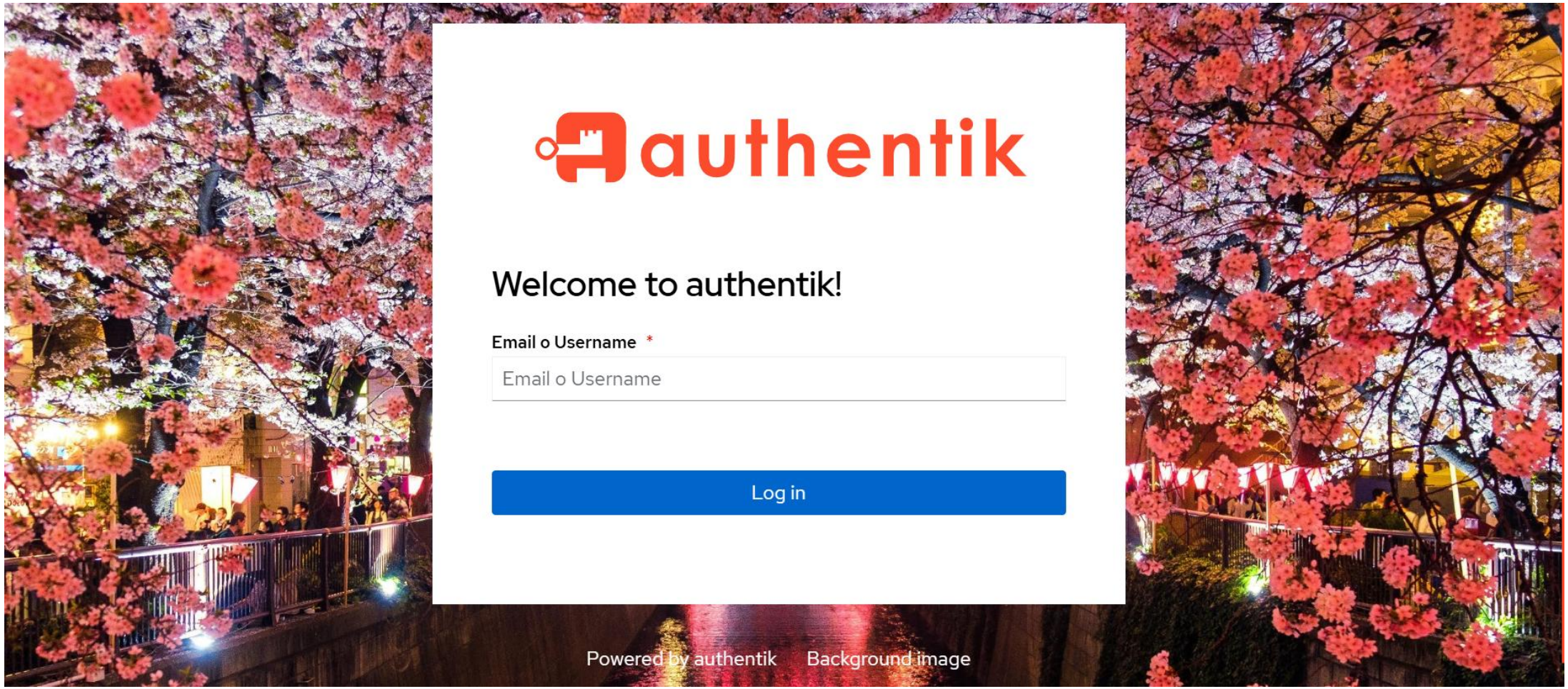
Data Policy

- **Introduction of the FAIR Data Manager** (not a fixed-term revenge)
- **Defined standards** (NeXus or file compatible with OFED)
- **Defined fixed acronym for the UO**
- **Defined buckets naming convention both for users and research-in-house (*nffa-di_[proposal_id] / nffa-di_[UO]_[project_id]*)**
- **Defined file naming convention (*nffa-di_[proposal_id]_[UO]_[UO_internal_id]/ nffa-di_[UO]_[project_id]_key*)**

AREA SCIENCE PARK

Authentik

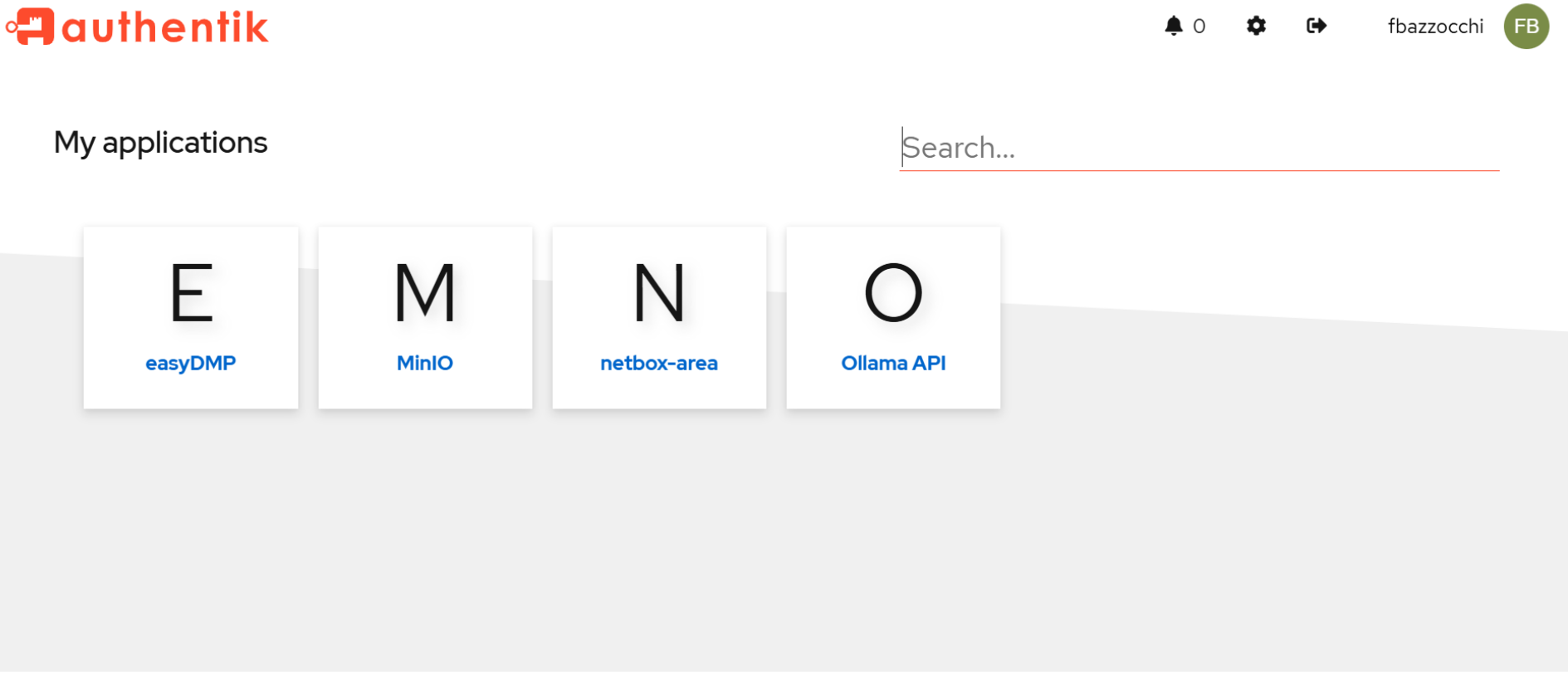
access



Service in production

Authentik

access



AREA SCIENCE PARK

Service in production

Authentik

access

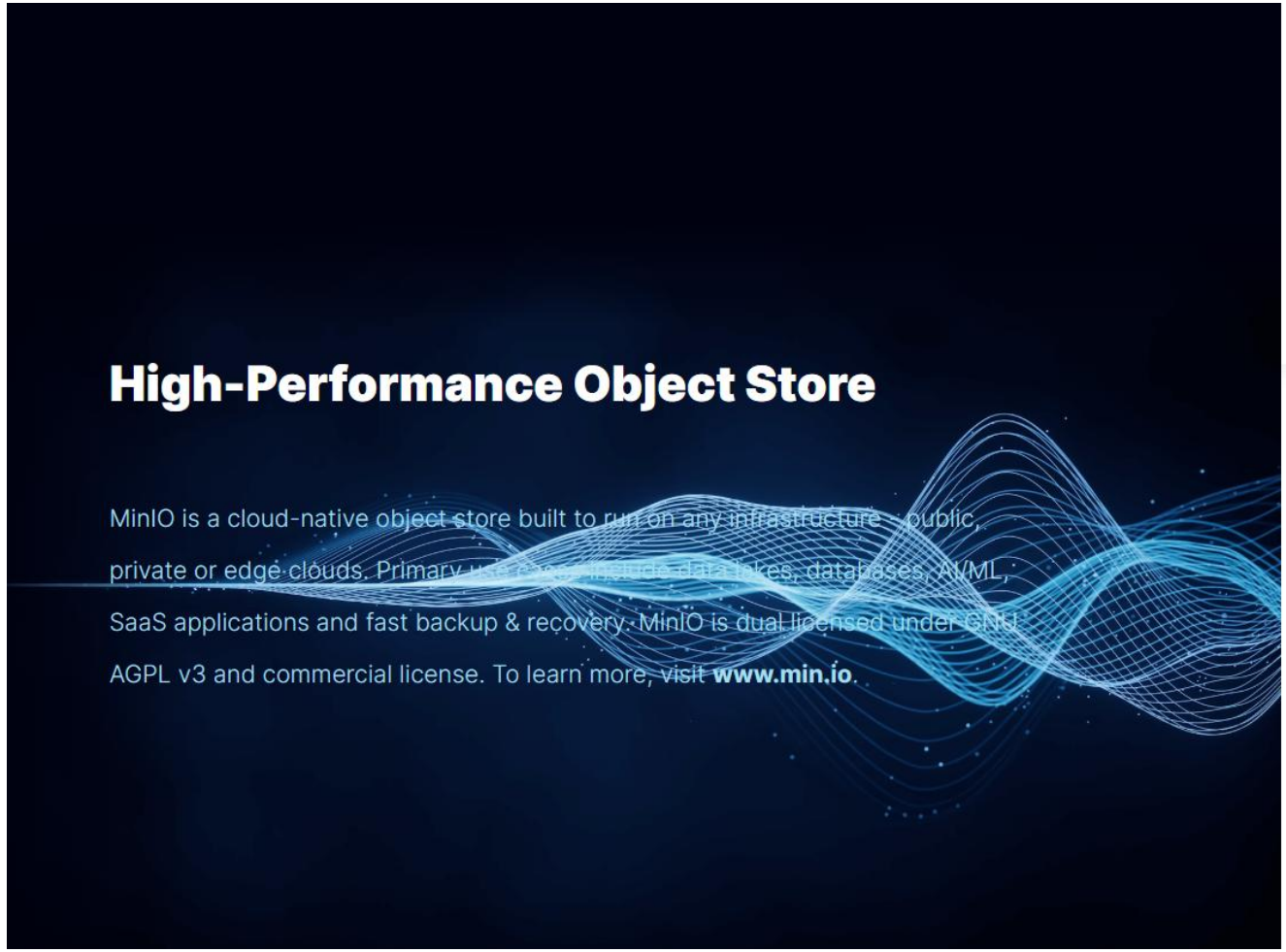
- **At the momentit requires a VPN**
- **write an email to *support@areasciencepark.it* to activate a VPN** (with in cc *Gianfranco.Gallizia@areasciencepark.it*)
- **write an email to *Gianfranco.gallizia@areasciencepark.it* or *federica.bazzocchi@areasciencepark.it* to create an account** (soon the registration page will be active)

AREA SCIENCE PARK

Service in production

MinIO

store & collect



MINIO
OBJECT STORE
AGPL LICENSE

Login with SSO (Authentik)

Other Authentication Methods

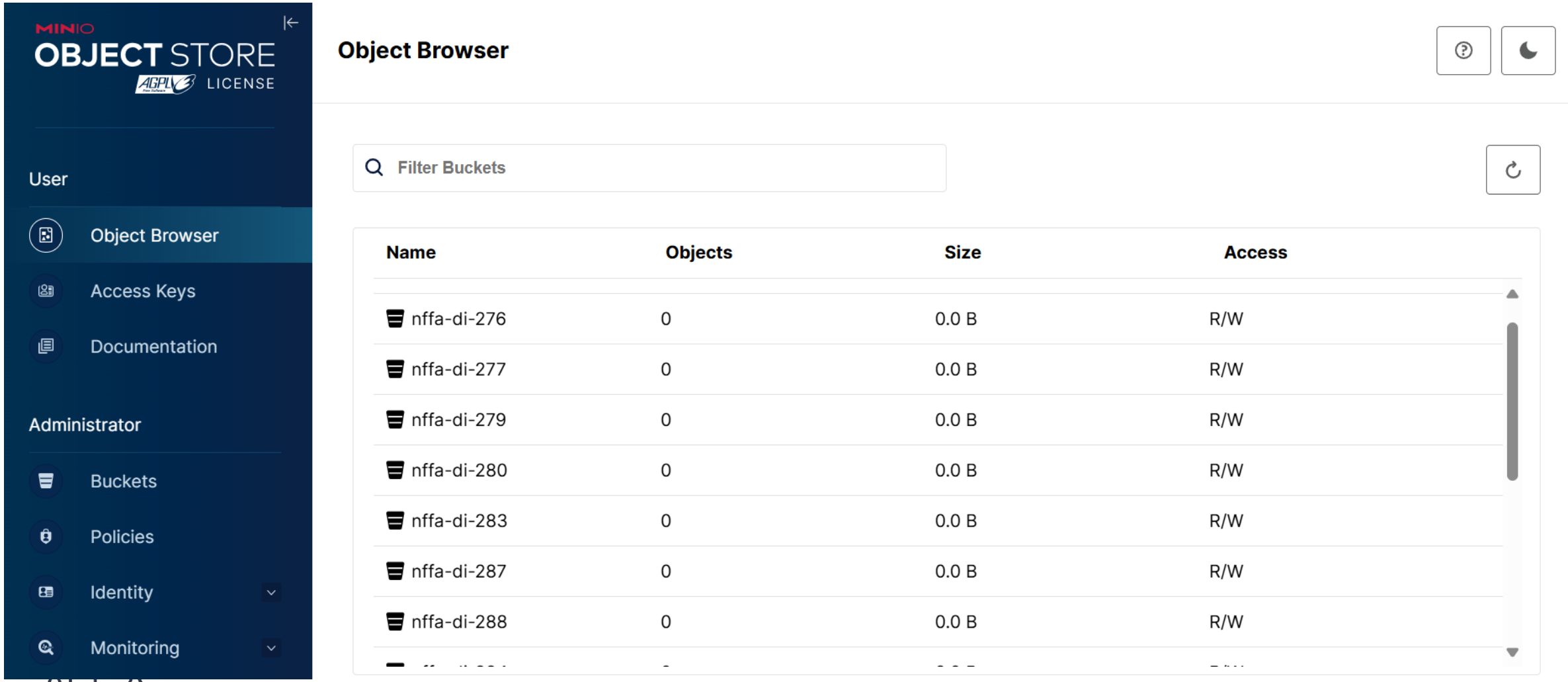
Documentation | GitHub | Support | Download

AREA SCIENCE PARK

Service in production

MinIO

store & collect

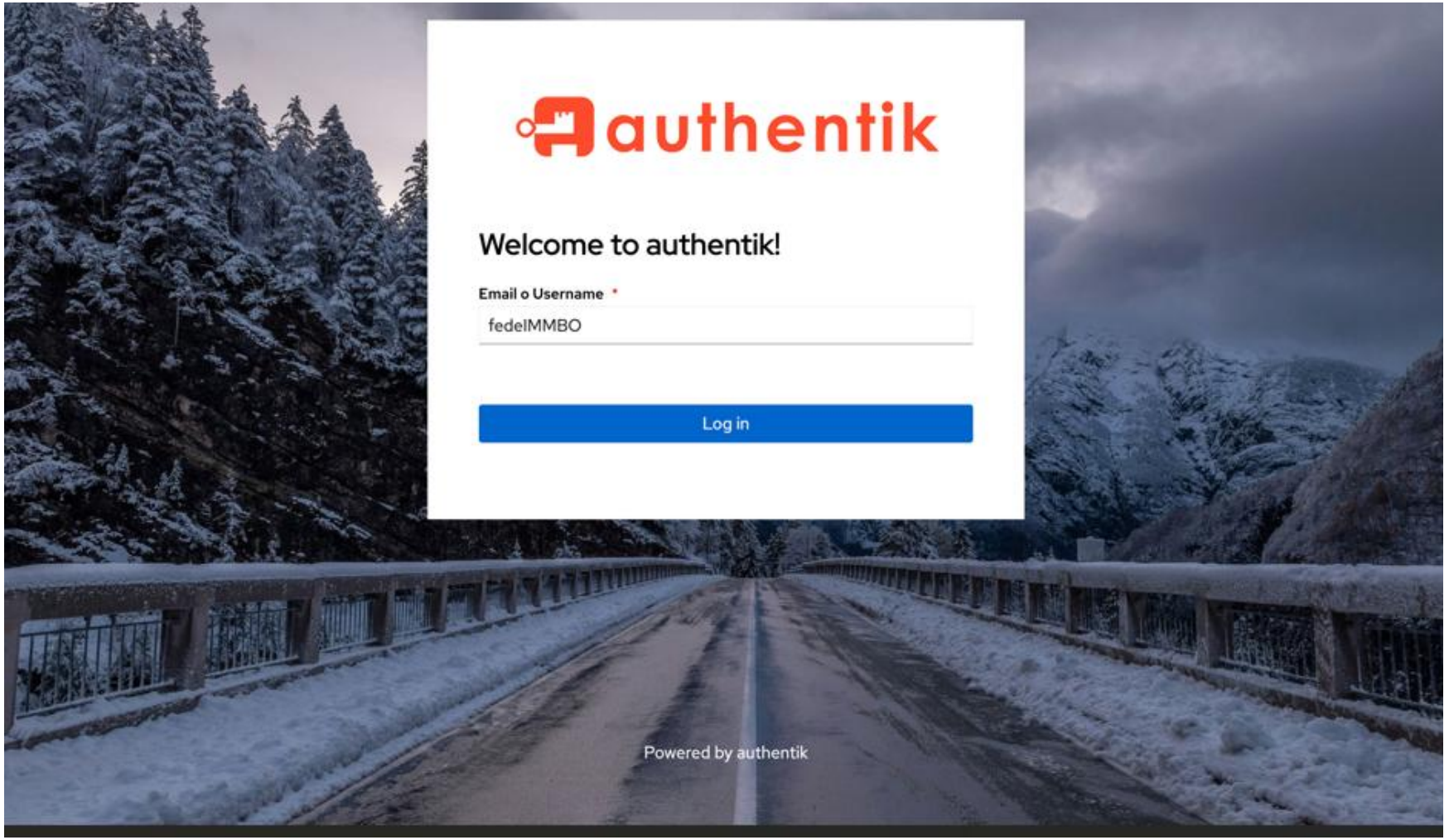


AREA SCIENCE PARK

Service in production

easyDMP

plan



Service in test

easyDMP

plan

NFFA-DI EASY DMP

Your UO is Consiglio Nazionale delle Ricerche -Istituto per lo studio dei materiali nanostrutturati- Bologna

Experimental technique

- This field is required.

Select your Technique: Chemical Vapour Deposition

Select

Instrument

Select

intstrument

Save

NFFA-DI EASY DMP

Your UO is Consiglio Nazionale delle Ricerche -Istituto per lo studio dei materiali nanostrutturati- Bologna

Experimental technique

Select your Technique: Chemical Vapour Deposition

Select

Instrument

Select your Instrument: ISMNBO - Furnace LPCVD for TEOS

Select

I Administrative Information

1. Installation Information

A) Institution Name

Consiglio Nazionale delle Ricerche -Istituto per lo studio dei materiali nanostrutturati- Bologna

B) Institution Site

Bologna

C) Facility Type

Laboratory source

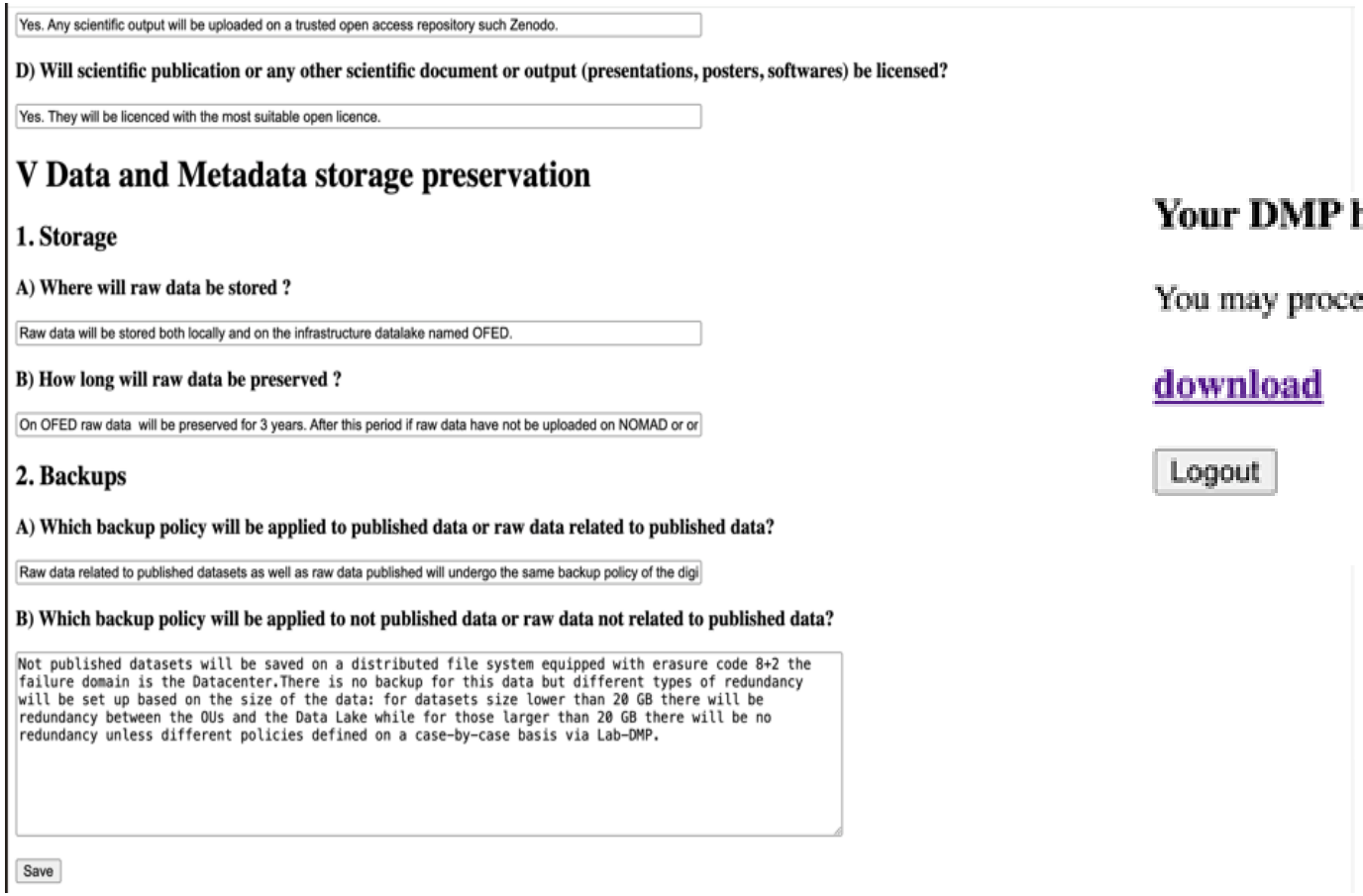
D) Kind of Installation

Growth & Synthesis

Service in test

easyDMP

plan

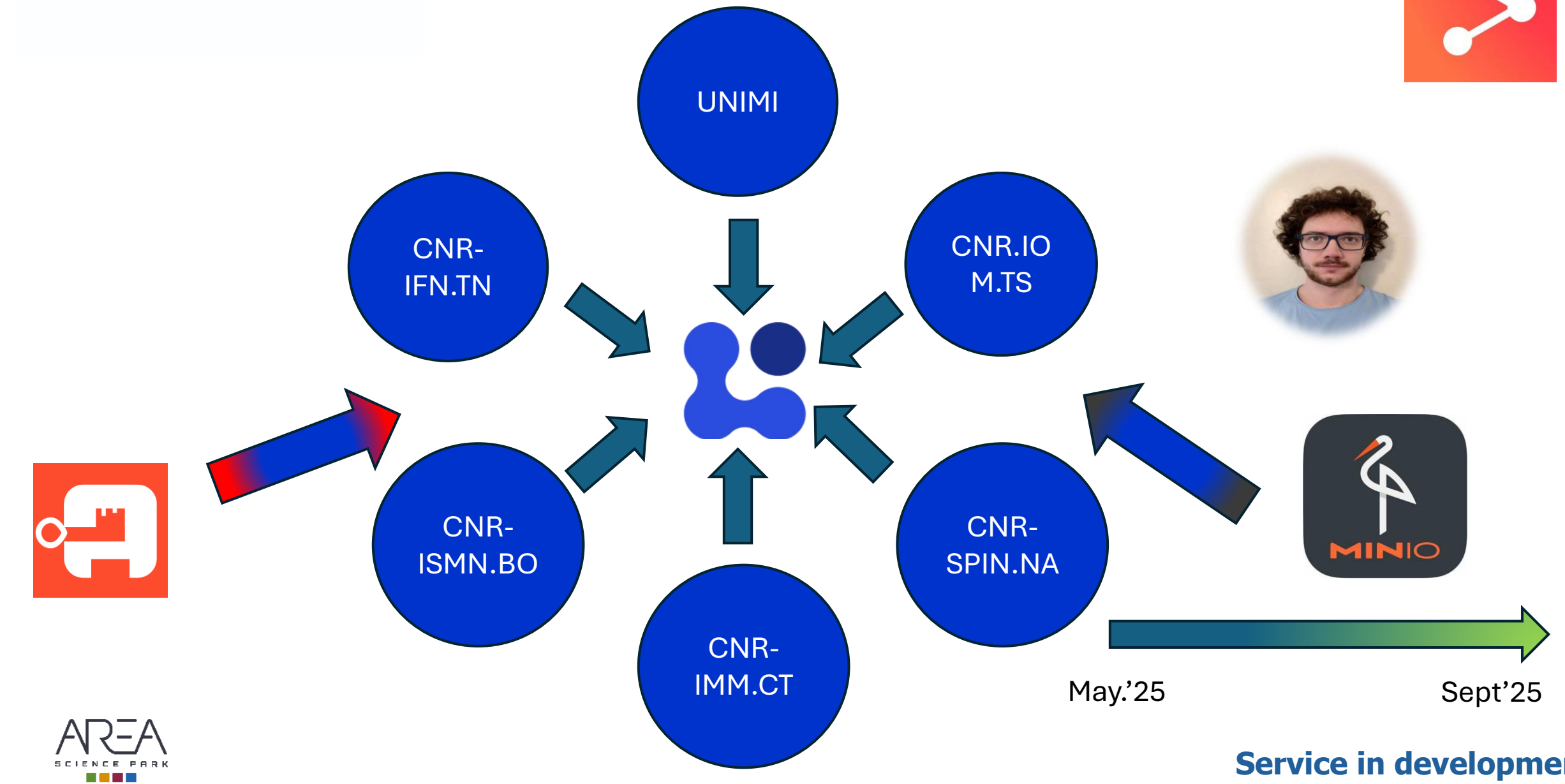


AREA SCIENCE PARK

Service in test

NOMAD

share



AREA SCIENCE PARK

Service in development

jupyterhub

analyze

July’25

Dec’25

FAIR-by-design pipelines

Some numbers regarding NFFA-DI offer

UO	# Techniques	# Instruments
UNIMI-Fisica	15	20
CNR-IOM@TS	14	22
POLIMI-POLIFAB	14	17
CNR-IMM@CT	11	15
CNR-ISMN@BO	8	18
CNR-NANOTEC@LE	8	9
CNR-SPIN@NA	7	8
CNR-IFN@TN	5	6
CNR-ISM@RM	2	2
AREA-RIT	2	2
CNR-IFN@MI	1	5

124!!

I suppose you know ALL the LABS need their FAIR-BY-DESIGN pipelines towards OFED as described in the project

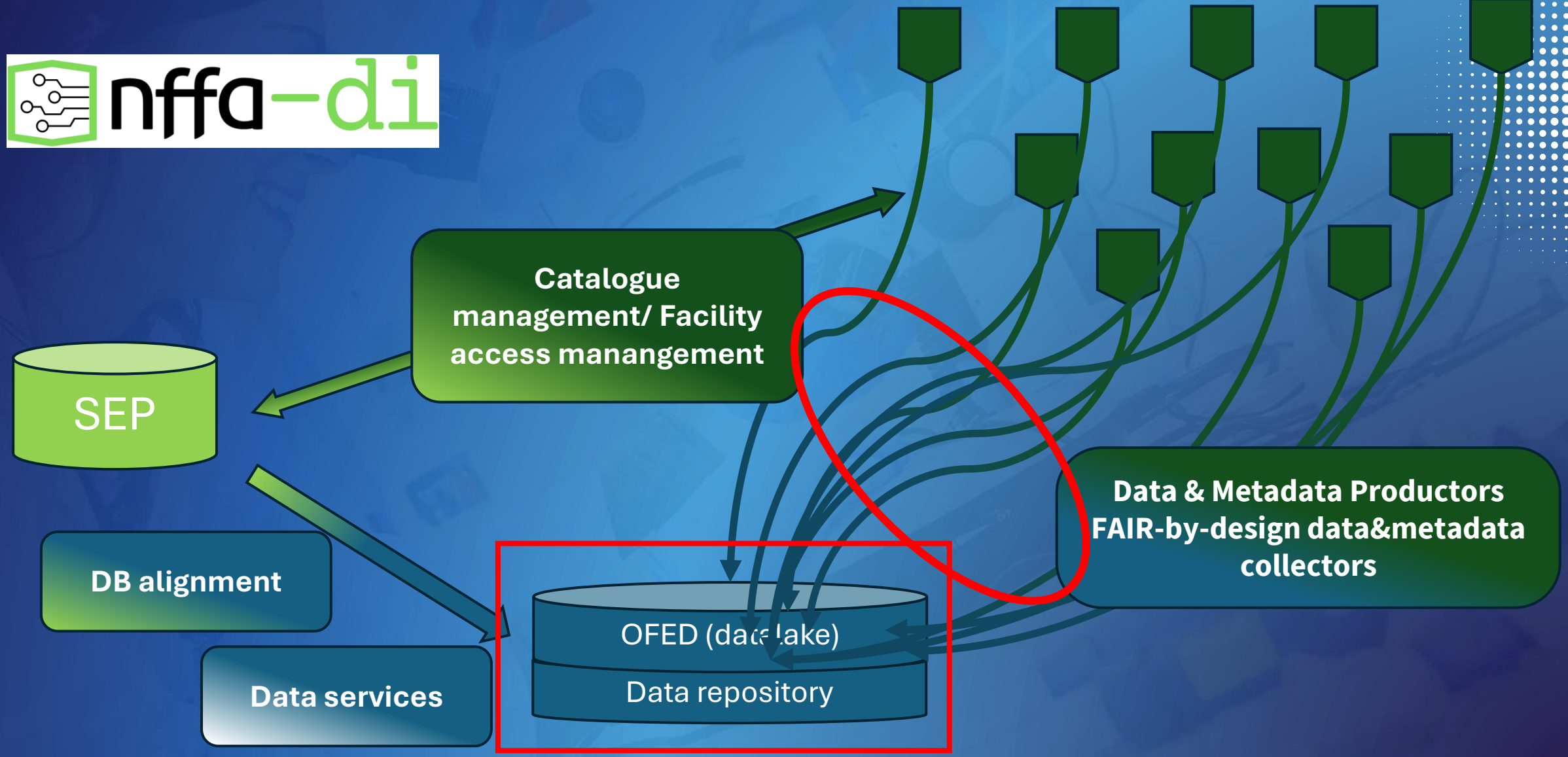
AREA SCIENCE PARK

Most popular techniques ….

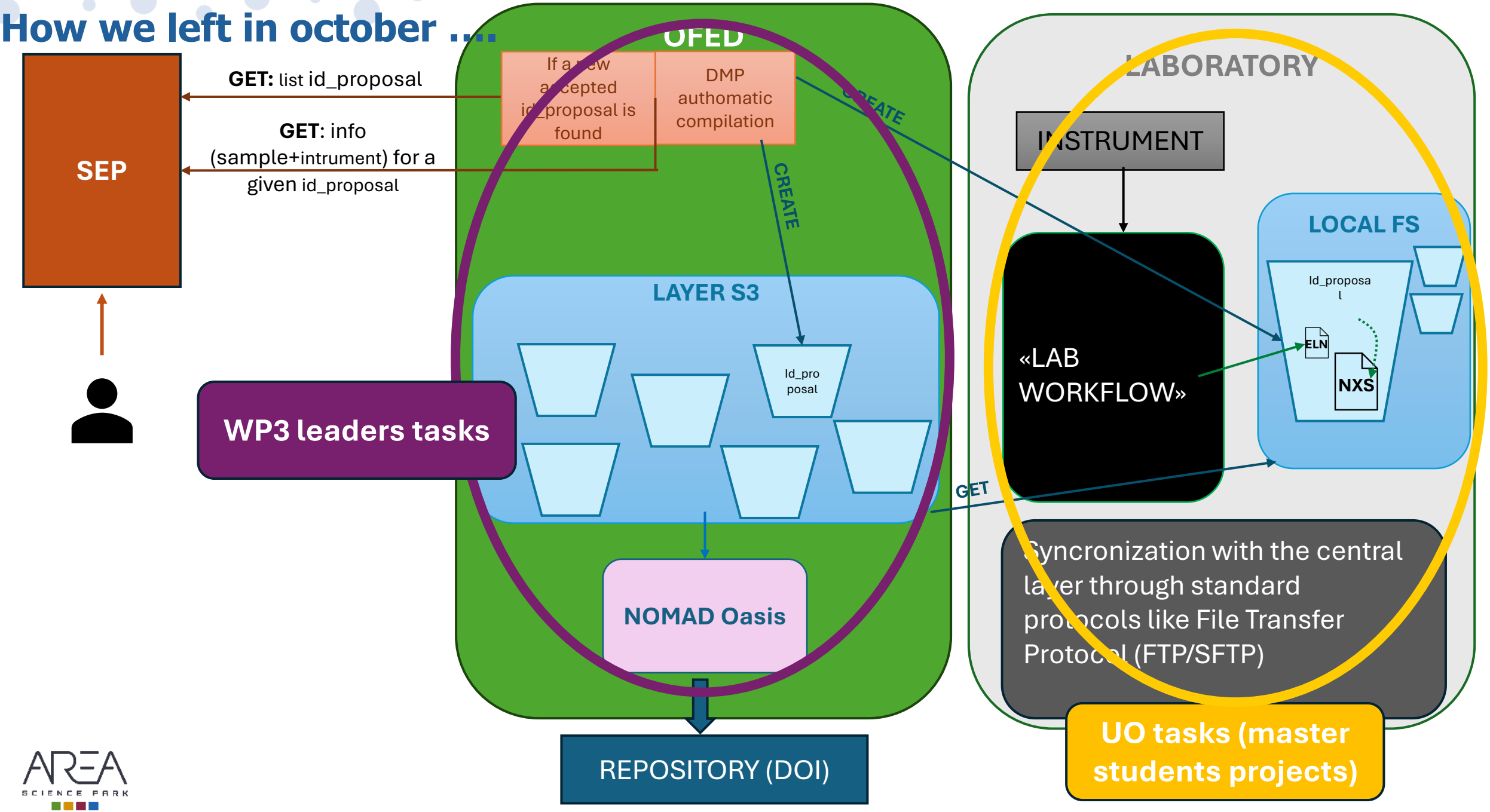
Technique	# UO
Scanning Electron Microscopy	5
Chemical Vapour Deposition	4
Photo-Electron Spectroscopy	4
Focused Ion Beam	3
Pulsed Laser Deposition	3
PVD sputtering	3
Scanning Tunneling Microscopy	3
Ultrafast-Spectroscopy	3
X-Ray Diffraction	3
Angle Resolved Photoelectron Spectroscopy	2
Atomic Force Microscopy	2
Atomic Layer Deposition	2
Dicing-saw	2
Electron Beam Lithography	2
FIB-Scanning Electron Microscopy	2
Magneto-Optic Kerr Effect	2
Molecular Beam Epitaxy	2
Optical-Lithography	2
Reactive Ion Etching	2
Transmission Electron Microscopy	2

AREA SCIENCE PARK

To remember ... elements of the Research Digital Infrastructure



How we left in october ....



WHAT WE HAVE SEEN IN THE LAST TWO DAYS

Maybe one of the largest number of FAIR-by-design (prototype) pipelines in experimental labs belonging to the same project …

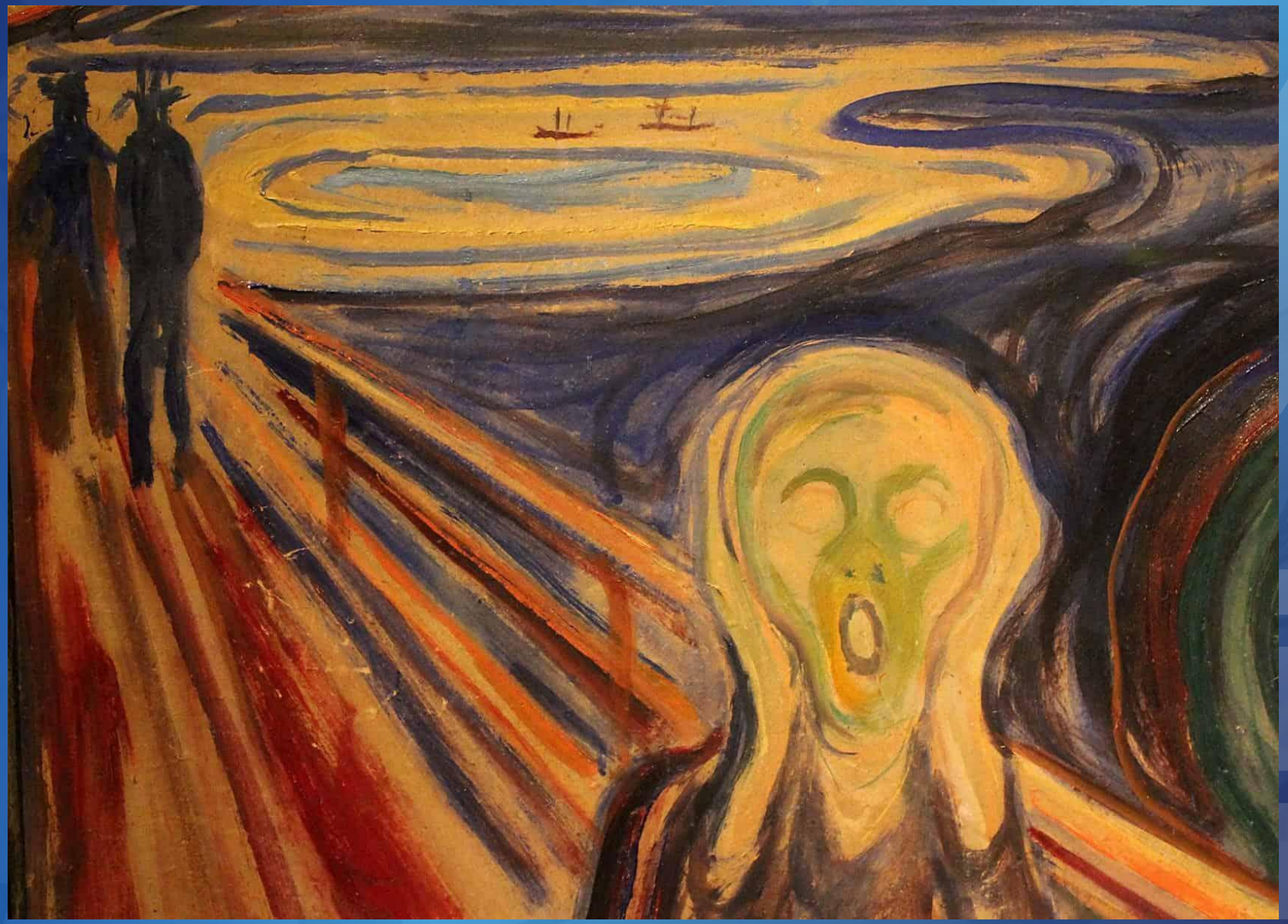
I think the largest so far in ITALY

AREA SCIENCE PARK

- Collaboration among:
  - CNR-IFN.TN&CNR-ISMN.BO
  - CNR-IFN.MI&CNR-ISM.RM
- Dialogue wih FAIRMAT (UNIMI, CNR-IMM.CT, CNR-IOM.TS)
- Reuse of the tool developed (CNR-NANOTEC.LE vs AREA, CNR-SPIN.NA vs CNR-IFN.TN&CNR-ISMN.BO)
- The EM NFFA-DI community agreed on NXem

- Same technique approached in different ways (simulation and synthesis)
- No complete semantic harmonization
- Many different approaches to the sample
- If it is difficult finding an agreement between a subset of the community, how we may thing to reach standards? (this is a wider discussion)

Now….



Questa foto di Autore sconosciuto è concesso in licenza da CC BY-NC-ND

Where we are now …

UO	STATUS (DMP/ PIPELINE)						
	YES&DONE	YES&ALMOST DONE	YES&IN PROGRESS	NO&ALMOST DONE	YES&NOT DONE	NO&NOT DONE	TOTAL
CNR-IOM@TS		21				1	**22**
UNIMI-Fisica			4		2	14	**20**
CNR-ISMN@BO			18				**18**
POLIMI-POLIFAB			1			16	**17**
CNR-IMM@CT		1	1			13	**15**
CNR-NANOTEC@LE			1			8	**9**
CNR-SPIN@NA		1		1	1	5	**8**
CNR-IFN@TN			6				**6**
CNR-IFN@MI		3			2		**5**
CNR-ISM@RM	2						**2**
AREA-RIT			2				**2**
TOTAL	**2**	**26**	**33**	**1**	**5**	**57**	**124**

Possible Strategies

I.
MDMC NEW EDITION 2025-2026

Shaping Data, Shaping the Future!

The MDMC is a full-time program, which consists of 1500 hours of training activities, corresponding to 60 ECTS credits:

Program Component	Hours	ECTS Credits
Classroom lectures	250	10
Internship	500	20
Final Project (Thesis)	125	5
Individual Study	625	25
Total	1500	60

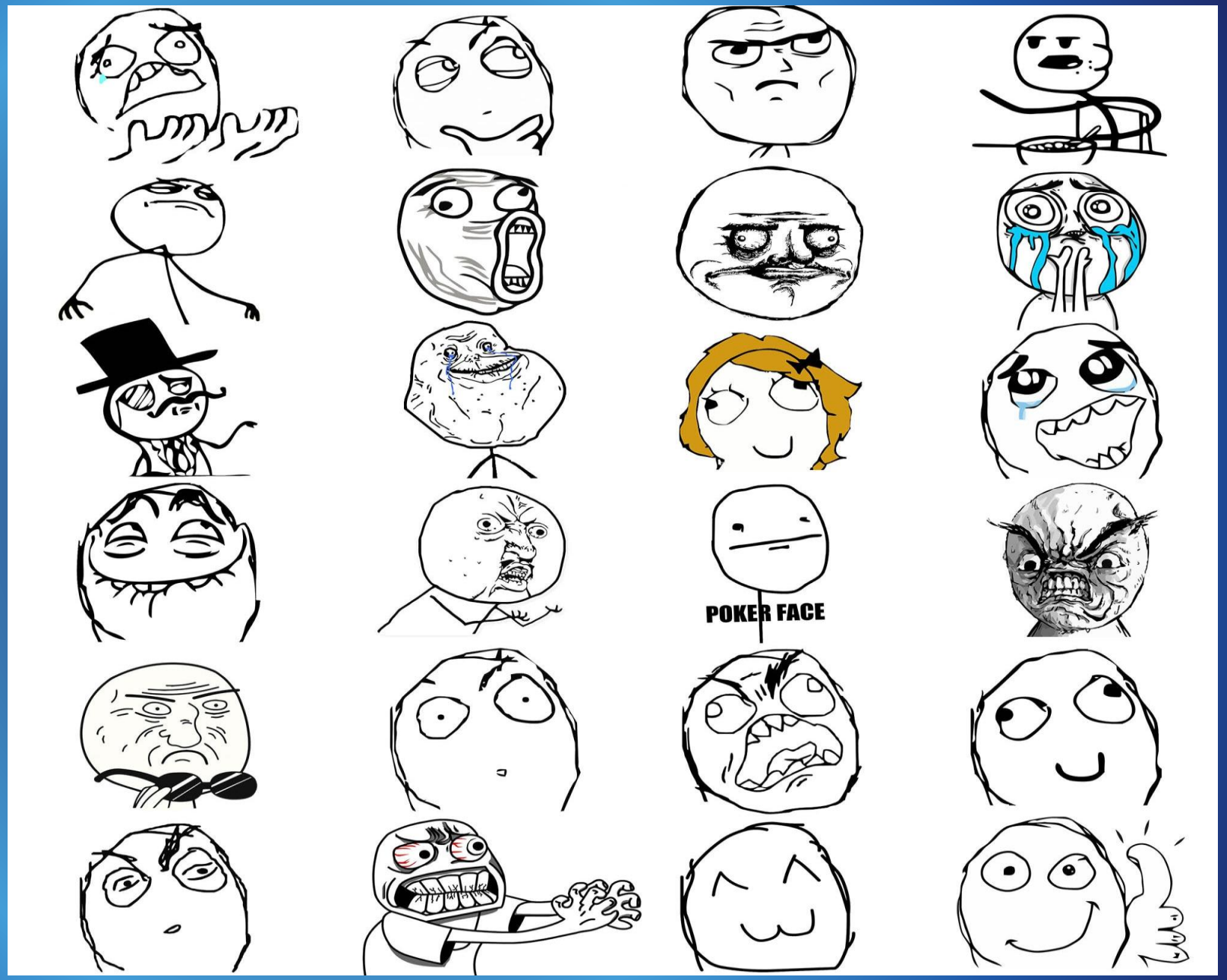
- **Total Duration:** 10 months, from September 2025 to June 2026
- **Teaching Language:** English
- **Teaching Mode:** In-person, with some remote sessions during the internship period
- **Location of in-presence lectures:** Trieste – Area Science Park & SISSA
- **Internship:** 6 months of hands-on experience in research laboratories and data-intensive environments

Full tuition fee € 4.500,00
Scholarships available!!!

AREA SCIENCE PARK

SISSA

II.
LOOK FOR YOUR SIMILAR!

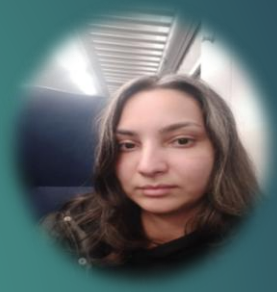


TECHNIQUE	AREA-RIT	CNR-IFN.MI	CNR-IFN.TN	CNR-IMM.CT	CNR-IOM.TS	CNR-ISM.RM	CNR-ISMN.BO	CNR-NANOTEC.LE	CNR-SPIN.NA	POLIFAB	UNIMI	TOTAL
Chemical Vapour Deposition			1	2	2		6					11
Ultrafast-Spectroscopy		5			3	1						9
Photo-electron Spectroscopy					3				1	1	1	6
Reactive Ion Etching			1				5					6
Scanning Electron Microscopy			1	1				1		1	1	5
Molecular Beam Epitaxy					3					1		4
Pulsed Laser Deposition					1				2	1		4
PVD sputtering				1			2		1			4
Scanning Tunneling Microscopy					2				1		1	4
Focused Ion Beam			1	1				1				3
Atomic Force Microscopy										1	2	3
X-Ray Diffraction				1				1	1			3
Transmission Electron Microscopy	1			1								2
Atomic Layer Deposition								1		1		2
Dicing-saw							1			1		2
Electron Beam Lithography			1							1		2
Optical-Lithography							1			1		2
RAMAN											2	2
X-Ray Reflectivity				1				1				2
Angle-resolved photoelectron spectroscopy					1					1		2
FIB - Scanning Electron Microscopy	1						1					2
Magneto-Optic Kerr Effect					1					1		2
Structural and ground-state electronic properties											2	2

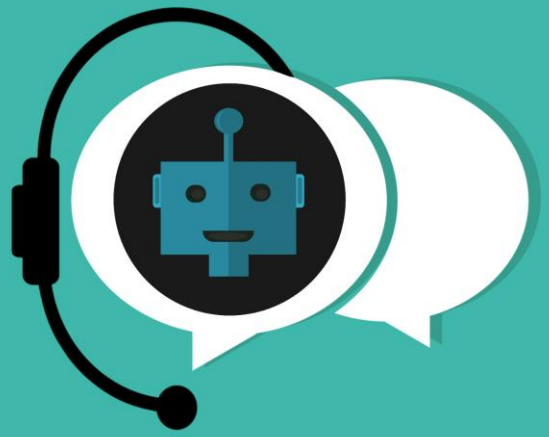
LOOK FOR YOUR SIMILAR!

- ❑ **All the codes developed are uploaded on github** (https://github.com/Master-Data-Management-and-Curation/Thesis_Code.git)
- ❑ **All the thesis will be at disposal**
- ❑ **UO network have been fruitful during the master so let's learn from the youngest!**

III.
New tool : Data Curator Assistant



- Creating an AI agent that could be deployed as additional OFED service to help FAIRify data
- Bottom up approach from the techniques better implemented so far
- Do not substitute a FAIR-by-design pipeline



Deploy
~~Submit~~ and … do not relax yet

Final Remarks

- **NFFA-DI is still far from reach the goal**
- **OFED** and the **UO** have now a solid base
- Many **deadlines** are coming (services, deliverables,..) so a deep analysis of the resources at disposal is needed by all of us

AREA SCIENCE PARK

THANKS FOR YOUR ATTENTION!