

OpenNP: Premier pas vers EOSC

A. Matta,¹ E. Clement,² A. Lemasson²

¹LPC Caen, CNRS/IN2P3 ²GANIL, CNRS/IN2P3



GANIL

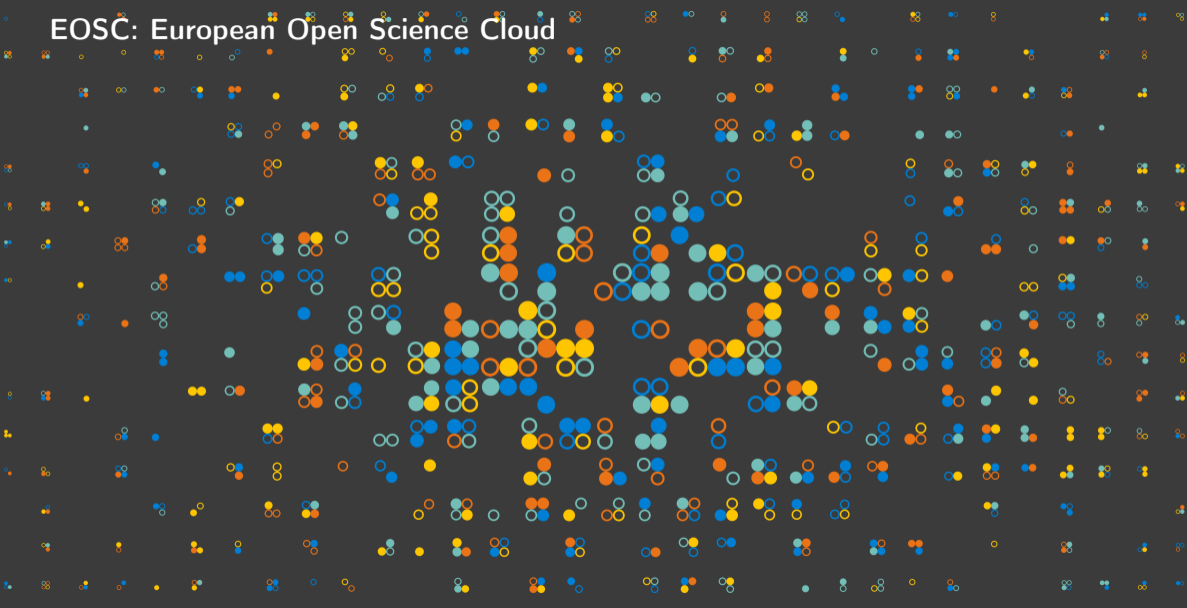


UNIVERSITÉ
CAEN
NORMANDIE



Normandie Université

EOSC: European Open Science Cloud



ooooo

oo

EOSC: European Open Science Cloud

Open Science in a Nut Shell

- Access to all researcher to every data-set, expertise, services and documentation
 - Reuse of data-set in a new context with optimal resources
- Open Data + Open Source + Open Access + Open Infrastructure FAIR principles

EOSC: European Open Science Cloud

Open Science in a Nut Shell

- Access to all researcher to every data-set, expertise, services and documentation
 - Reuse of data-set in a new context with optimal resources
- Open Data + Open Source + Open Access + Open Infrastructure FAIR principles

Should we do Open Science ?

- ✓ Open Science is a legal obligation for public founded research
 - ✓ Loi pour une Republique Numerique (2016)
 - ✓ Plan National pour la science ouverte (21-24)
- ✓ Maximise scientific output and projects sustainability
- ✓ Transparency, Accountability and good practices

EOSC: European Open Science Cloud

Open Science in a Nut Shell

- Access to all researcher to every data-set, expertise, services and documentation
 - Reuse of data-set in a new context with optimal resources
- Open Data + Open Source + Open Access + Open Infrastructure FAIR principles

Should we do Open Science ?

- ✓ Open Science is a legal obligation for public founded research
 - ✓ Loi pour une Republique Numerique (2016)
 - ✓ Plan National pour la science ouverte (21-24)
- ✓ Maximise scientific output and projects sustainability
- ✓ Transparency, Accountability and good practices

EOSC: eosc-portal.eu

- Web of FAIR research data and services
- Central access of all European research data and tools
- Free at point of use

OpenNP : a novel initiative within EURO-LABS

Context: EURO-LABS

- European scale project: GANIL, LPC Caen, IJCLab, GSI/FAIR, INFN, Jyvaskyla
- Organise and facilitate access to European facilities

OpenNP : a novel initiative within EURO-LABS

Context: EURO-LABS

- European scale project: GANIL, LPC Caen, IJCLab, GSI/FAIR, INFN, Jyvaskyla
- Organise and facilitate access to European facilities

OpenNP: In a nutshell

- Open science initiative dedicated to nuclear physics
- Catalogue of data-set, related information and tools
- ESCAPE-like, for future integration

OpenNP : a novel initiative within EURO-LABS

Context: EURO-LABS

- European scale project: GANIL, LPC Caen, IJCLab, GSI/FAIR, INFN, Jyvaskyla
- Organise and facilitate access to European facilities

OpenNP: In a nutshell

- Open science initiative dedicated to nuclear physics
- Catalogue of data-set, related information and tools
- ESCAPE-like, for future integration

OpenNP: Work packages

- 1 Open science desk (GANIL/LPC Caen): promote good practice: DMPs, source repo., ...
- 2 OpenNP catalogue (GANIL/LPC Caen): the product itself
- 3 AAI (IJCLab): Provide necessary infrastructure to access and manage the catalogue
- 4 AAI (GSI/FAIR): Prototype for data lake interface

OpenNP : a novel initiative

Short-term goals (i.e. within EUROLABS)

OpenNP : a novel initiative

Short-term goals (i.e. within EUROLABS)

- E Overview of existing raw-data sets
- E Overview of existing apparatus: ion sources, accelerator, separator and detector

Difficulty scale

- E Technically easy and little work required

OpenNP : a novel initiative

Short-term goals (i.e. within EUROLABS)

- E Overview of existing raw-data sets
- E Overview of existing apparatus: ion sources, accelerator, separator and detector
 - I Associated aux-data (i.e. log book)
 - I Associated software to exploit raw-data and aux-data
 - I Overview of existing analysed and simulated data set

Difficulty scale

- E Technically easy and little work required
 - I Technically easy but lot of implication by all actors

OpenNP : a novel initiative

Short-term goals (i.e. within EUROLABS)

- E Overview of existing raw-data sets
- E Overview of existing apparatus: ion sources, accelerator, separator and detector
 - I Associated aux-data (i.e. log book)
 - I Associated software to exploit raw-data and aux-data
 - I Overview of existing analysed and simulated data set
- D Associated software to exploit and produce analysed and simulated data

Difficulty scale

- E Technically easy and little work required
 - I Technically easy but lot of implication by all actors
- D Technically difficult, and require a lot of implication by all actors

OpenNP : a novel initiative

Getting the community on board

OpenNP : a novel initiative

Getting the community on board

✓ Automatic aggregation: only small, easy action required

- Create habits, integrate OpenNP in everyday workflow
- Everybody involved: ITA, researcher, collaboration, direction
- Elementary task: better definition of each actor responsibilities

OpenNP : a novel initiative

Getting the community on board

✓ Automatic aggregation: only small, easy action required

- Create habits, integrate OpenNP in everyday workflow
- Everybody involved: ITA, researcher, collaboration, direction
- Elementary task: better definition of each actor responsibilities

✓ Easier bibliography, experiment planning, and re-use of existing data set

- Does a dataset exist where ^{54}Ca was populated?
- Which facility provide the most intense ^{18}C beam?
- Was ^{10}He ever measured by missing mass?
- Mixing information from different data-set (e.g. Fission with VAMOS/LICORNE/ILL)

OpenNP : a novel initiative

Getting the community on board

✓ Automatic aggregation: only small, easy action required

- Create habits, integrate OpenNP in everyday workflow
- Everybody involved: ITA, researcher, collaboration, direction
- Elementary task: better definition of each actor responsibilities

✓ Easier bibliography, experiment planning, and re-use of existing data set

- Does a dataset exist where ^{54}Ca was populated?
- Which facility provide the most intense ^{18}C beam?
- Was ^{10}He ever measured by missing mass?
- Mixing information from different data-set (e.g. Fission with VAMOS/LICORNE/ILL)

✓ Provide metric: increase visibility, facilitate evaluation

- 25% of publication on Coulex reaction used the AGATA array
- FASTER has been used in 54 exp. over the last 5 years
- VAMOS has been used in 16 different focal plane configurations

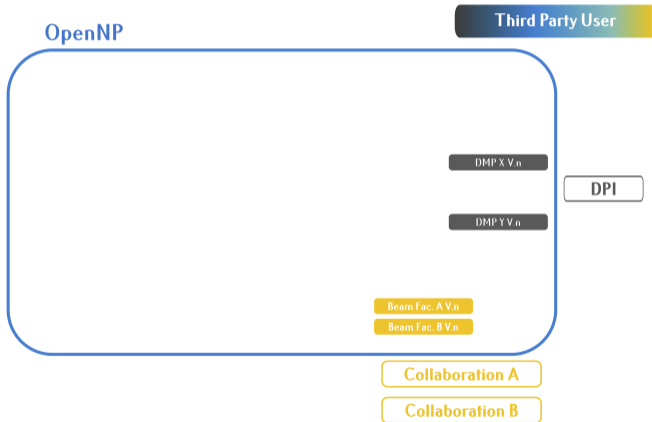
OpenNP : a novel initiative

Schematic interaction map



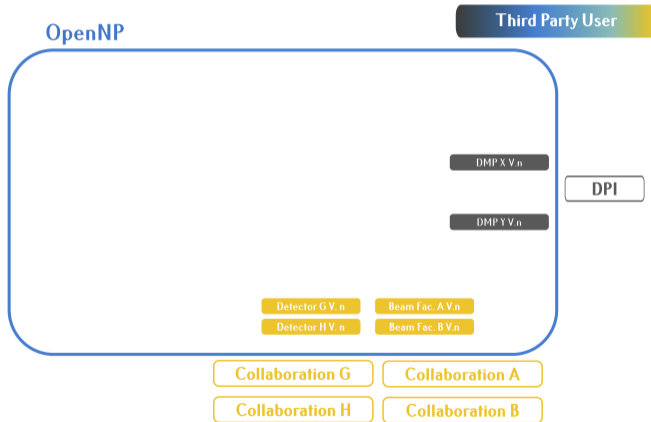
OpenNP : a novel initiative

Schematic interaction map



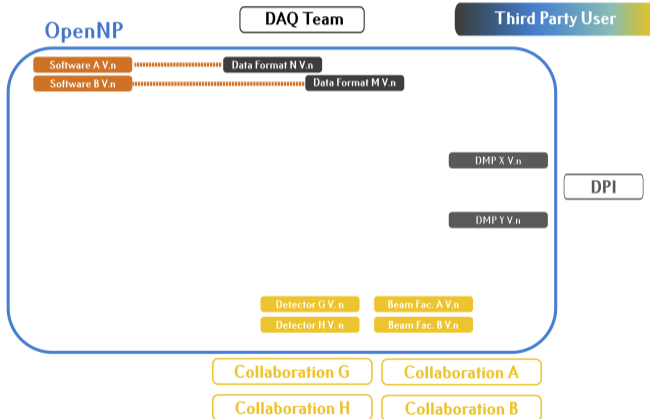
OpenNP : a novel initiative

Schematic interaction map



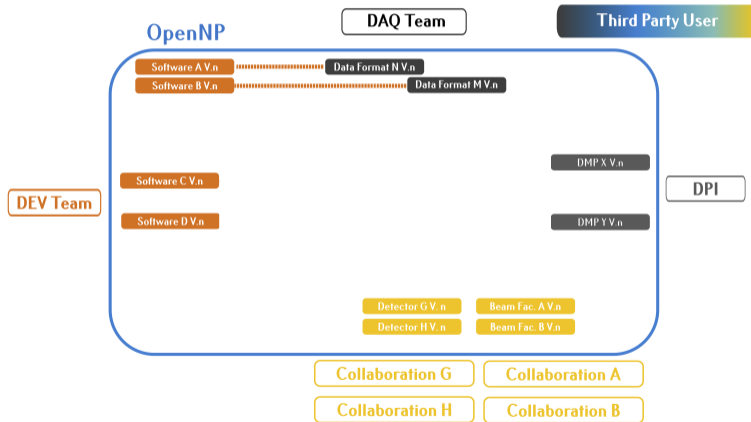
OpenNP : a novel initiative

Schematic interaction map



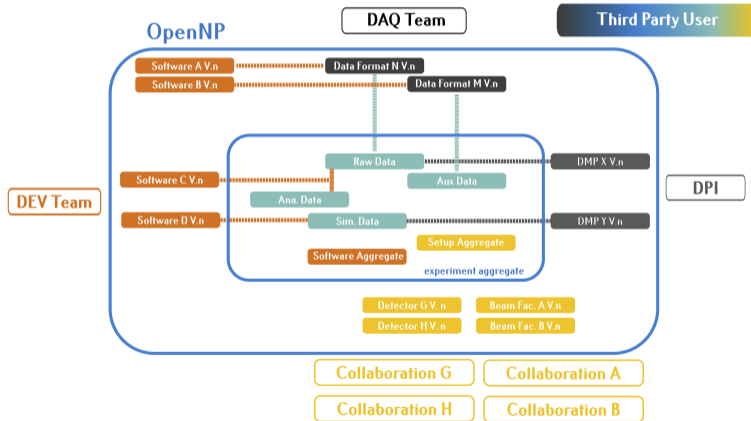
OpenNP : a novel initiative

Schematic interaction map



OpenNP : a novel initiative

Schematic interaction map



OpenNP EURO-LABS and Beyond

A complete road-map

Short-Term:

- A fully working catalogue
 - ⇒ CDD 3 years (GSI) → starting this fall
 - ⇒ CDD 1.5 years (IJCLab) → starting 2023
 - ⇒ CDD 2+1 years (GANIL) → starting 2023
 - ⇒ GDR Resanet kick-off meeting → fall 2022 or winter 2023
 - ⇒ Mock-up (LPCC/GANIL) → early 2023 for international evaluation

OpenNP EURO-LABS and Beyond

A complete road-map

Short-Term:

- A fully working catalogue
 - ⇒ CDD 3 years (GSI) → starting this fall
 - ⇒ CDD 1.5 years (IJCLab) → starting 2023
 - ⇒ CDD 2+1 years (GANIL) → starting 2023
 - ⇒ GDR Resanet kick-off meeting → fall 2022 or winter 2023
 - ⇒ Mock-up (LPCC/GANIL) → early 2023 for international evaluation

Mid-Term:

- OpenNP synergy and integration with other initiative (ESCAPE/PANOSC/...)
 - ⇒ Learning from other fields
 - ⇒ Bridging gap at domain interfaces
 - ⇒ Seamless integration to data lake

OpenNP EURO-LABS and Beyond

A complete road-map

Short-Term:

- A fully working catalogue
 - ⇒ CDD 3 years (GSI) → starting this fall
 - ⇒ CDD 1.5 years (IJCLab) → starting 2023
 - ⇒ CDD 2+1 years (GANIL) → starting 2023
 - ⇒ GDR Resanet kick-off meeting → fall 2022 or winter 2023
 - ⇒ Mock-up (LPCC/GANIL) → early 2023 for international evaluation

Mid-Term:

- OpenNP synergy and integration with other initiative (ESCAPE/PANOSC/...)
 - ⇒ Learning from other fields
 - ⇒ Bridging gap at domain interfaces
 - ⇒ Seamless integration to data lake

Long-Term:

- Analysis and simulation as a service
 - ⇒ Single entry point data lake
 - ⇒ Shared computing platform
 - ⇒ Analysis and simulation tool at the ready

Future needs

Challenges ahead

Triggerless DAQ era:

- From ~ 1 Go/week to ~ 10 To/week the last 10 years
 - ⇒ Increase in storage needs

Future needs

Challenges ahead

Triggerless DAQ era:

- From ~ 1 Go/week to ~ 10 To/week the last 10 years
 - ⇒ Increase in storage needs

Complexity of analysis:

- Heterogeneous systems
 - ⇒ Need for inter-operable analysis/simulation tools
 - ⇒ Need for common computing infrastructure

Future needs

Challenges ahead

Triggerless DAQ era:

- From ~ 1 Go/week to ~ 10 To/week the last 10 years
 - ⇒ Increase in storage needs

Complexity of analysis:

- Heterogeneous systems
 - ⇒ Need for inter-operable analysis/simulation tools
 - ⇒ Need for common computing infrastructure

Multi-site / Multi-Collaboration research:

- Heterogeneous policies ⇒ Usually no provision on data stewardship in MoU
 - ⇒ Need for clear policies
 - ⇒ Need for data officer

Future needs

Identified needs

Researcher:

- Open Science compliant practices (Open Software, CI/CD, Documentation ...)
- Develop awareness of importance of software development and technical skills

From Institutions

- The community needs a strong framework:
 - ⇒ **Policies** : Clear and constraining policies for nuclear physics
 - ⇒ **Relay** : Identified person with clear responsibilities
 - ⇒ **Visibility** : Include Open Science in evaluation
- Standardisation of practices around the diversity of experiments:
 - ⇒ **Coordination** : Virtual Organisations
 - ⇒ **Open Access** : Data Storage
 - ⇒ **Processing** : Analysis & Simulation Platform
 - ⇒ Define the correct granularity
 - ⇒ Quantify the projected needs
 - ⇒ Support for deploying services

Future needs

Identified needs

Researcher:

- Open Science compliant practices (Open Software, CI/CD, Documentation ...)
- Develop awareness of importance of software development and technical skills

From Institutions

- The community needs a strong framework:
 - ⇒ **Policies** : Clear and constraining policies for nuclear physics
 - ⇒ **Relay** : Identified person with clear responsibilities
 - ⇒ **Visibility** : Include Open Science in evaluation
- Standardisation of practices around the diversity of experiments:
 - ⇒ **Coordination** : Virtual Organisations
 - ⇒ **Open Access** : Data Storage
 - ⇒ **Processing** : Analysis & Simulation Platform
 - ⇒ Define the correct granularity
 - ⇒ Quantify the projected needs
 - ⇒ Support for deploying services

IN2P3 is the right actor to drive the community to its goals