Complementary DM discoveries: How to be FAIR?

DM Summary plots and the Virtual Research Environment

Sukanya Sinha, L. Chazallet, C. Doglioni and Y. Ema (input from S. Kulkarni)
University of Manchester

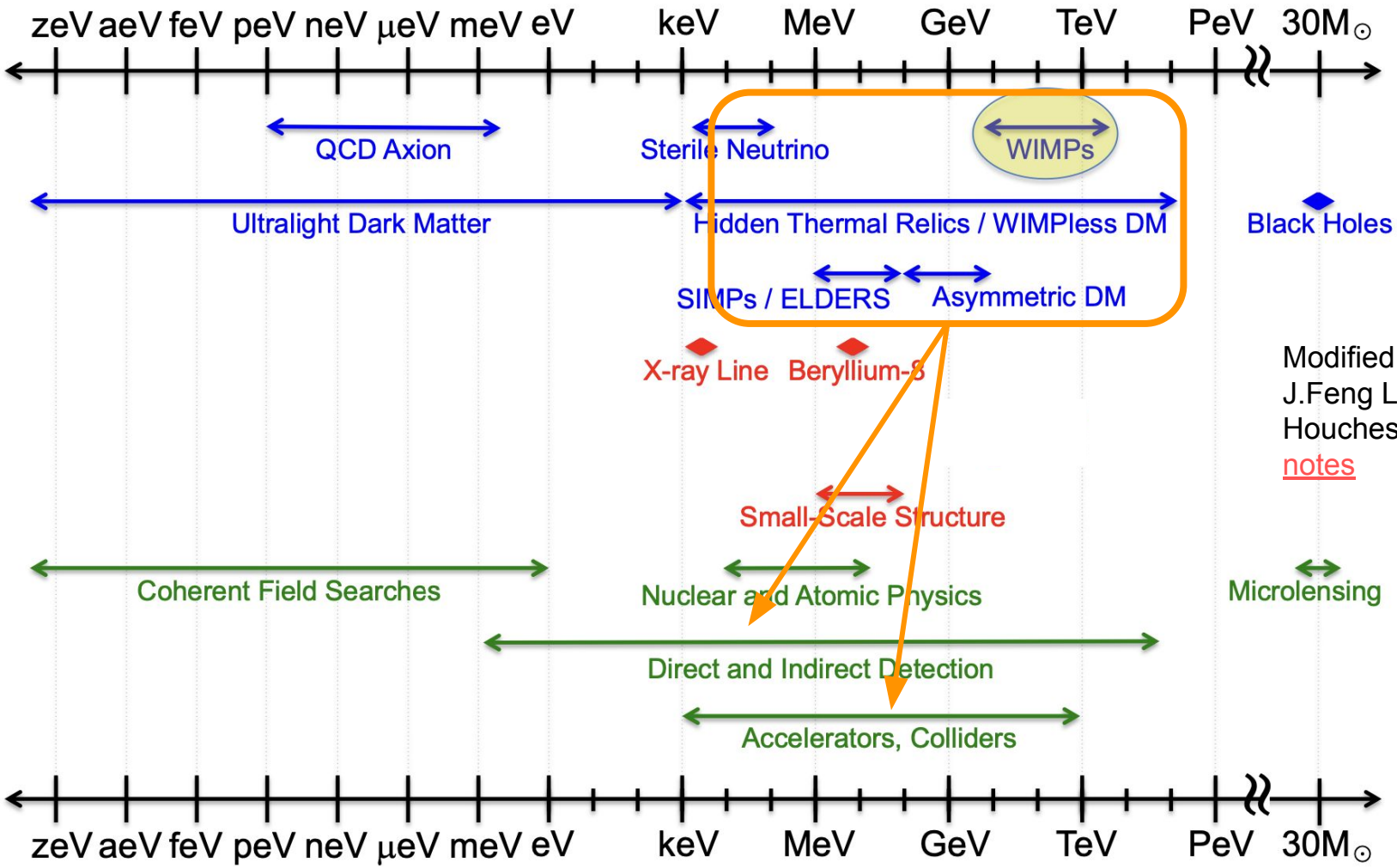
DM UK meeting
University of Birmingham, 1st December 2025

DM Complementarity going forward?

[Inspired by S. Kulkarni's Snowmass Community slides link]

Multiple observations, experiments and coherent theories are needed for DM discovery

- Observations motivating DM arise from astrophysics & cosmic probes
- Theoretical frameworks are crucial to put different observations into context
- **Direct Detection** ➜ can discover DM with cosmological origin
- **Indirect Detection** ➜ can probe decays of cosmological DM into SM particles
- **Colliders / accelerators** ➜ can produce DM and probe its dark interactions



Modified from J.Feng Les Houches school notes

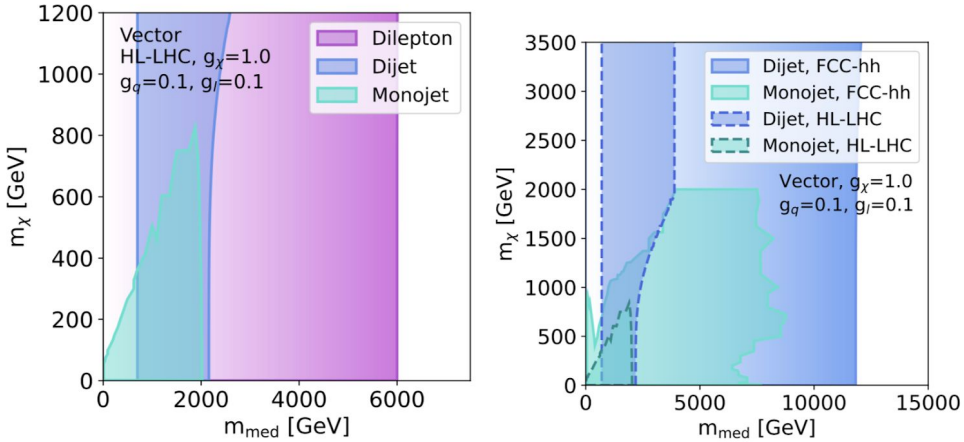
DM mediated by BSM particles

At collider energy scales, we can access the BSM particles that can mediate the SM-DM interactions

Mediator can decay into DM → points to an excess in missing energy

Mediator can also decay into SM particles → accessible using resonance searches

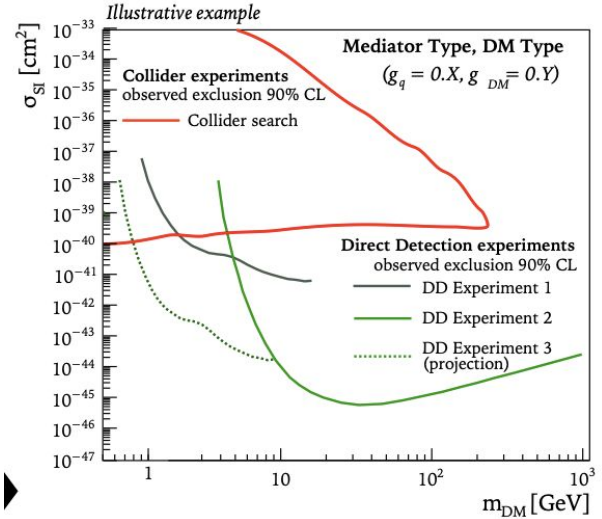
https://arxiv.org/abs/2203.12035



Example scenarios for reach of future colliders for vector mediator resonances

Overlaying results from collider simplified models searches to DD results in DM-nucleon plane:
Perfect region to discover DM would be where both colliders and direct detection have coverage
helps to determine both cosmological origin of discoveries as well as nature of interaction

https://arxiv.org/pdf/1912.12739

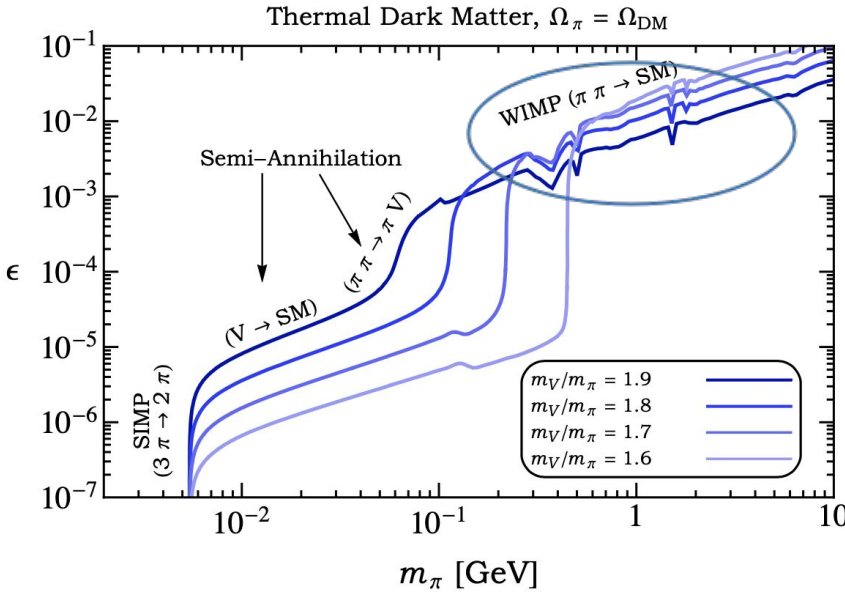


Potential common sketch
(lines should be made more realistic)

Extended dark sectors and dark showers

Dark QCD scenarios predict dark matter as a bound state originating due to confinement

- Early collider results exist
- First steps towards consistent collider analysis taken during Snowmass and current LHC dark showers initiative
- Overlap between colliders and direct detection needs more studies
- Excellent opportunity to exploit synergy with accelerator searches
- Some understanding via lattice QCD calculations for properties of SU(N) theories exists



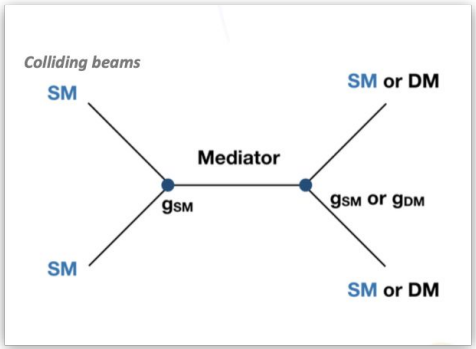
HV sector π abundance in m-ϵ plane
arXiv:1801.05805

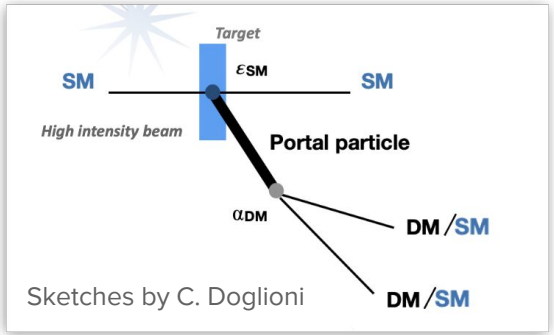
Contours are regions where abundance matches DM relic density

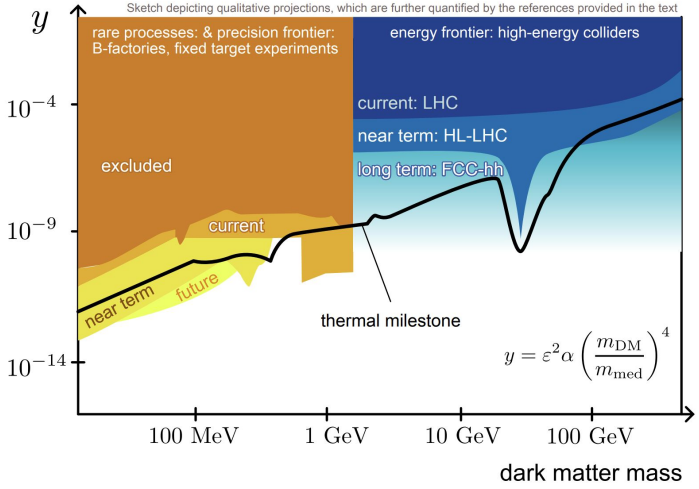
Light-ish dark matter at colliders

Some models of light dark matter (dark photon/dark Higgs) require new high-mass particles to be theoretically self-consistent

By using same benchmarks between colliders and accelerator experiments, colliders can discover directly the high-mass particles, complementing light DM discoveries at accelerator experiments







Joshua Greaves's Master's thesis
https://arxiv.org/abs/2209.13128
https://scipost.org/SciPostPhysCommRep.7/pdf

Potential for analysis preservation and reinterpretation

Multiple models predicting DM. ***Need to delve deep and search wide***!

Collider searches have fairly inclusive selections

→ **ALLOWS TO PROBE OTHER MODELS WITH SIMILAR FINAL STATES**

BUT several times it becomes difficult for even the ATLAS analysers to reproduce the analysis results after a few months à either code not publicly available or previous analyser left the field…

Code preservation of experimental analyses is essential!

Science 2061

From S. Kraml's talk

- Many members of the generation who may wish to re-analyze the LHC Run 3 data in 2061 have yet to be born.
- That generation may wish to re-analyze these data simultaneously with old LSST and LIGO data using BSM models that they invent from which (with the help of their AI assistants) they are able to compute testable predictions for LHC, LSST, and LIGO data.
- Now ask yourself: What are we doing wrong, or what are we forgetting, that risks thwarting such heroic efforts forty year hence?

Harrison Prosper, RiF 2021

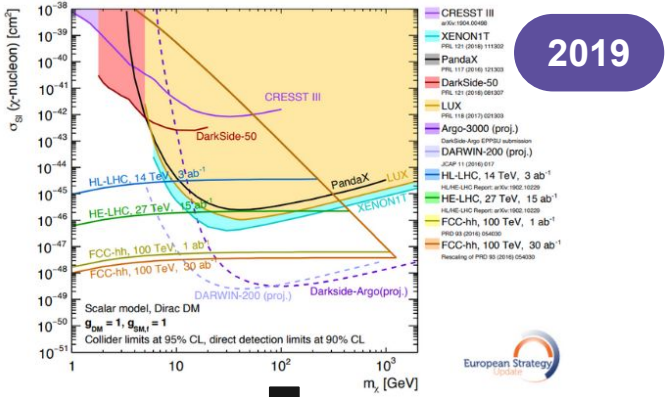
Make analyses live beyond the paper publication !

Example: DM Summary Plots reproducibility issue

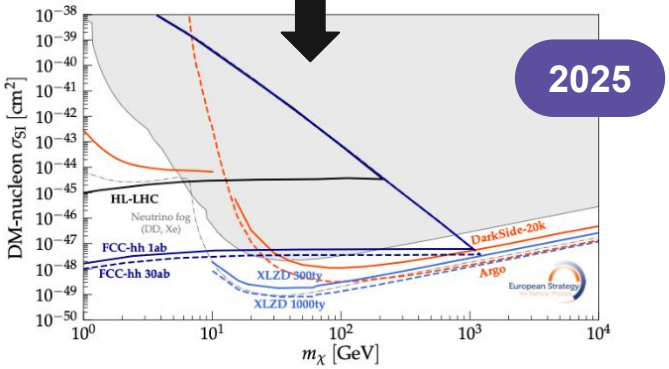
Difficulties with reproducing and updating summary plots from 2019 to 2025 ESPPU:

- **Git** repository to make the final plot available, but **private** (by student who worked on the plots in 2019)
- **Software** to translate HL-LHC results **only shared privately** via e-mail
- **References** on where to find the curves only available in figure caption, and **not descriptive enough** (HL-LHC report: 1418 pages, many plots...)
- **No README** or documentation

→ not an uncommon situation in the sciences



Contains input for



*Portraying the **complementarity of colliders and direct detection** for **heavy dark matter** (s-channel simplified model with scalar mediator, WIMP-like Dirac DM)*

Example: DM Summary Plots reproducibility issue

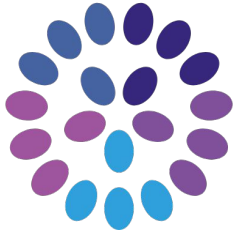
Think of an optimistic situation: if **one (or more) experiments discover hints of dark matter**, the community will want to **compare** and **scrutinise**:

- Experimental techniques
- Assumptions and uncertainties
- Datasets
- Data analysis software

Our solutions:

- Develop **notebooks** following **code quality** / **FAIR guidelines**
- Execute them on a **dedicated computing cluster** that **helps reproducibility**, the **ESCAPE Virtual Research Environment**

Aligns perfectly with the goals of the **OSCARS** project (details in backup) and with the **ESCAPE** Open Collaboration



OSCARS
Open Science Clusters' Action for Research & Society

ESCAPE: The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures

An **open collaboration** that aims to **address the Open Science challenges** shared by **ESFRI facilities** (e.g. HL-LHC) and pan-European **research infrastructures** (e.g. CERN) in astronomy and particle physics

ESCAPE and related initiatives have supported and **funded services** that are used in HEP and beyond (e.g. **REANA**, **RUCIO/data lake**, **authentication** services, a Zenodo **software catalogue**, the **VRE**, …)

reana

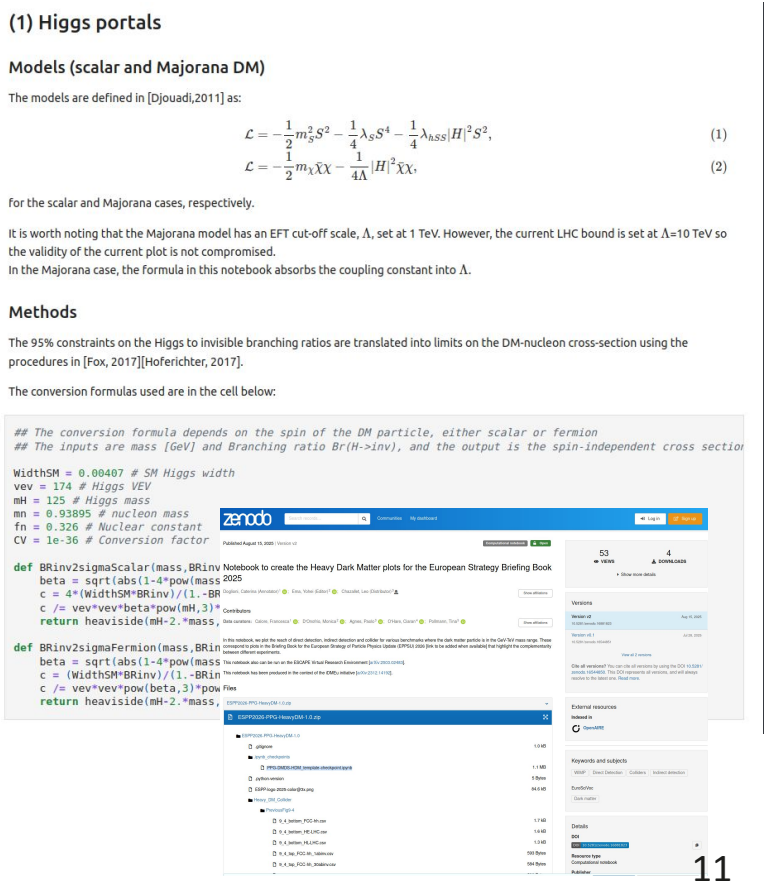
SCIENTIFIC DATA MANAGEMENT
RUCIO

Back to summary plots: Jupyter Notebook

The 2025 summary plots are now made using a **Jupyter notebook** with:

- tracking of **imported package** versions (using the uv package manager)
- installation **instructions** and **commented** code blocks
- light-touch **explanations** of model and parameters, with citations
- **Metadata tables** with provenance for the curves (location of digitized curves, citation with figure location)
- An **Open Source license** for conditions of reuse

The notebook is available on **Zenodo:**
DOI 10.5281/zenodo.16544850



(1) Higgs portals

Models (scalar and Majorana DM)

The models are defined in [Djouadi,2011] as:

$$\mathcal{L} = -\frac{1}{2}m_S^2 S^2 - \frac{1}{4}\lambda_S S^4 - \frac{1}{4}\lambda_{hSS}|H|^2 S^2, \quad (1)$$

$$\mathcal{L} = -\frac{1}{2}m_\chi \bar{\chi}\chi - \frac{1}{4\Lambda}|H|^2 \bar{\chi}\chi, \quad (2)$$

for the scalar and Majorana cases, respectively.

It is worth noting that the Majorana model has an EFT cut-off scale, Λ, set at 1 TeV. However, the current LHC bound is set at Λ=10 TeV so the validity of the current plot is not compromised.
In the Majorana case, the formula in this notebook absorbs the coupling constant into Λ.

Methods

The 95% constraints on the Higgs to invisible branching ratios are translated into limits on the DM-nucleon cross-section using the procedures in [Fox, 2017][Hoferichter, 2017].

The conversion formulas used are in the cell below:

In [15]:

```
## The conversion formula depends on the spin of the DM particle, either scalar or fermion
## The inputs are mass [GeV] and Branching ratio Br(H->inv), and the output is the spin-independent cross sectio

WidthSM = 0.00407 # SM Higgs width
vev = 174 # Higgs VEV
mH = 125 # Higgs mass
mn = 0.93895 # nucleon mass
fn = 0.326 # Nuclear constant
CV = 1e-36 # Conversion factor

def BRinv2sigmaScalar(mass,BRinv
    beta = sqrt(abs(1-4*pow(mass
    c = 4*(WidthSM*BRinv)/(1.-BR
    c /= vev*vev*beta*pow(mH,3)*
    return heaviside(mH-2.*mass,

def BRinv2sigmaFermion(mass,BRin
    beta = sqrt(abs(1-4*pow(mass
    c = (WidthSM*BRinv)/(1.-BRin
    c /= vev*vev*pow(beta,3)*pow
    return heaviside(mH-2.*mass,
```


This follows many of the guidelines of the ICFA Data Lifecycle Panel and of the EVERSE Project's RSQKit

Where does this Jupyter notebook run?

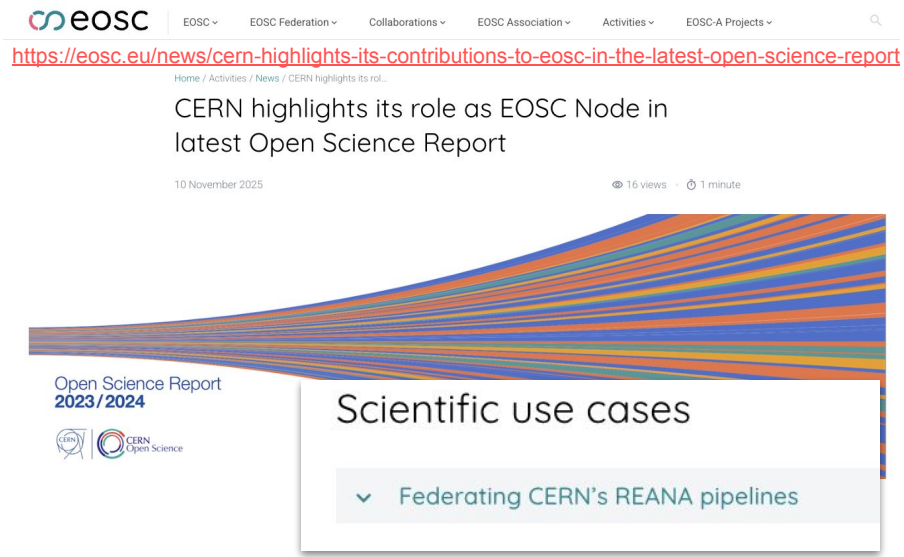
Anyone's laptop! Start from Zenodo 10.5281/zenodo.16544850, install uv, and you're ready to go

But what if I wanted an even **easier** way to **rerun** my workflow (especially applies for more **complex** ones, e.g. full reinterpretation, REANA-based…)

It would be nice to have a platform where *everything is already installed* and I can just *log in and run*!

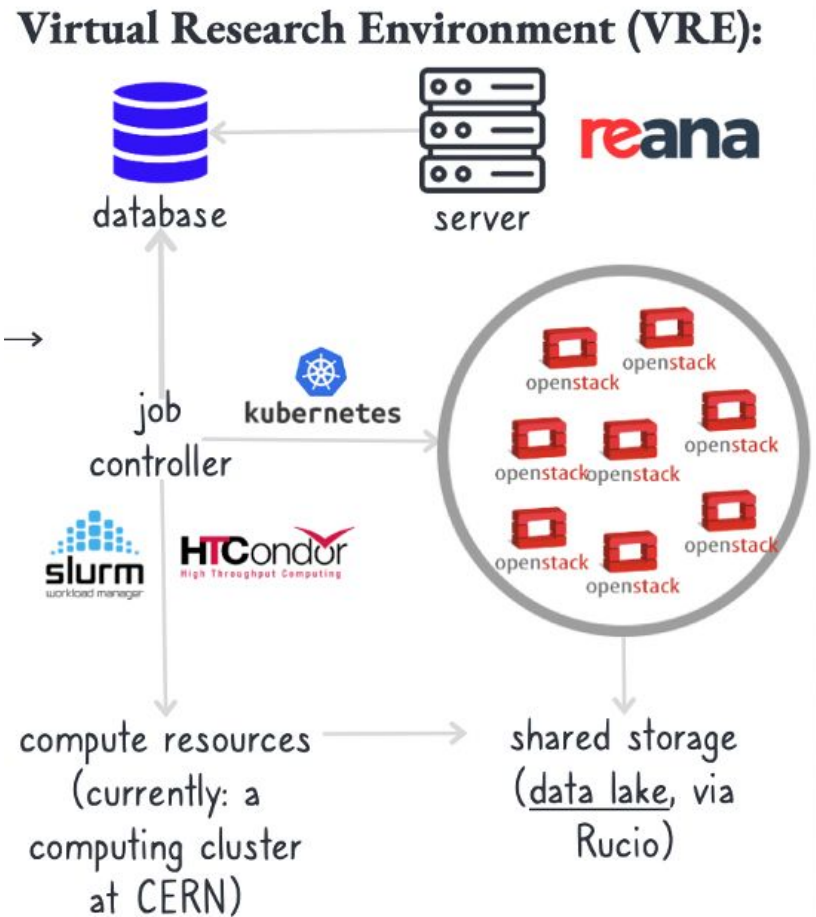
This is **where the Virtual Research Environment** comes in!

Related: CERN and the European Open Science Cloud (EOSC)

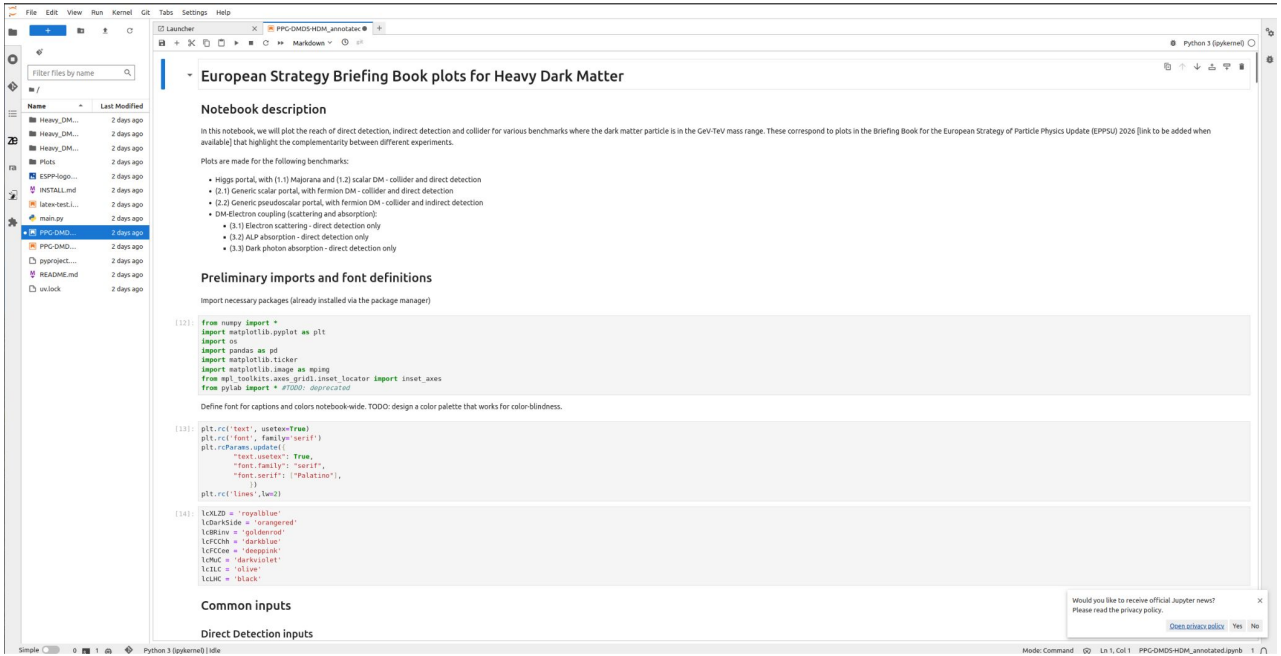


The ESCAPE Virtual Research Environment

1. AAI: A federated and reliable **Authentication and Authorization layer** for **logging in** and **protecting access** to private data
2. The Rucio Data Lake: A **federated distributed storage** solution, providing functionalities for **data injection and replication** through a Data Management framework (**Rucio**)
3. REANA: A **computing cluster** supplying the processing power to **run full analyses** with REANA, a re-analysis service (supports RECAST as well)
4. CVMFS: A **read-only file system** designed to **distribute software**, and more.
5. JupyterHub: A **notebook interface** with containerised environments to hide the infrastructure's complexity from the user.



Implementation: LAPP VRE



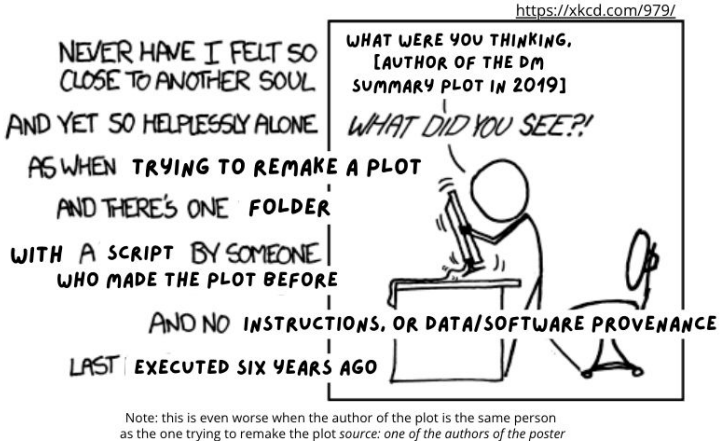
By executing the notebooks / our workflows on the VRE, we ensure that we have a **suitable computing infrastructure** that others can use

For examples of other workflows (including an ATLAS analysis using RECAST), see the ESCAPE Dark Matter Science Project

Conclusions and outlook

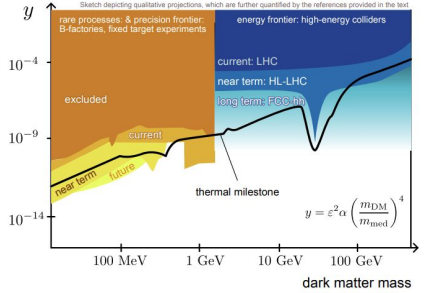
Solved the problem of "*not being able to reproduce the European Strategy DM plots*", at least for the next Strategy!

- New plots use a **Jupyter notebook** following FAIR guidelines, on Zenodo
- **Execution** is **enabled** on the ESCAPE Virtual Research Environment (**VRE**)
- The VRE provides an easy-to-use platform for rerunning workflows - and for **reinterpretation**? **TBD**!



Note: this is even worse when the author of the plot is the same person as the one trying to remake the plot *source: one of the authors of the poster*

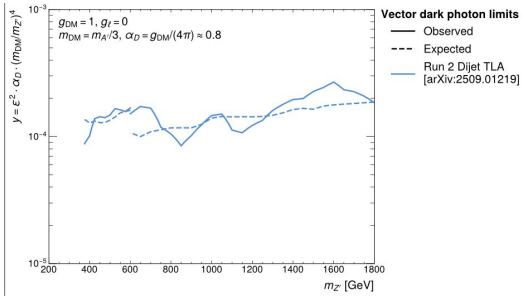
Coming soon: other LHC DM WG interpretations (e.g. dark photon)



https://scipost.org/SciPostPhysCommRep.7

https://github.com/mamerl/DMWG-dark-photon-tools, soon to be moved to the DMWG repository after cross-checks

(expert feedback welcome!)



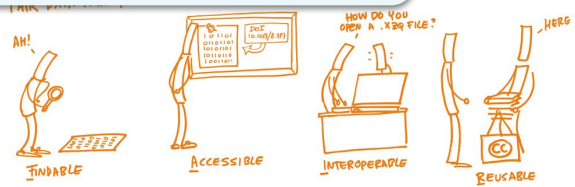
Test rescaling with published Run-2 ATLAS dijet (scouting dataset) vector mediator results. Courtesy: M. Amerl

BACKUP

OSCARS, Open Science and FAIR principles

Fostering the uptake of Open Science in Europe

EU grant that aims to enhance adoption of **Open Science** and supporting interdisciplinary **FAIR data services** and **practices**



€ 16 MILLIONS
IN OPEN CALLS FOR
OPEN SCIENCE
PROJECTS AND SERVICES

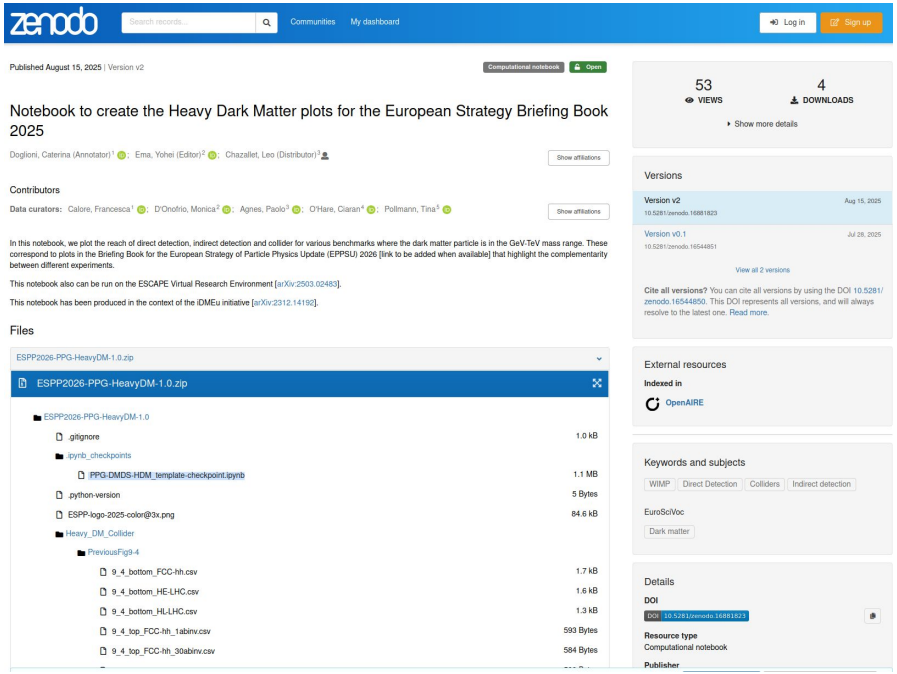
OBJECTIVES

Setup and implementation of **Clusters' Open Science Competence Centres (CLOCCs)**

Identify and provide **Composable Open Data and Analysis Services (CODAS)**

Contribute to a **Data space for science, research and innovation**

More on Zenodo

Notebooks are available on Zenodo

- Open-access research data repository operated by **CERN**
- Supports sharing of **datasets, papers, software, and more**
- Each upload gets a **DOI** for permanent citation
- Promotes **open science** and **data preservation**
- Compliant with **FAIR principles** (Findable, Accessible, Interoperable, Reusable)
- Ensures **long-term storage** and **global accessibility**