




ESCAPE

European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures

Towards a Dark Matter Test Science Project Caterina Doglioni - Lund University

Input from: Antonio Boveia, Francesca Calore, Elena Cuoco, Lukas Heinrich,
Samuel Meehan, Graeme Stewart, Pasquale Serpico, Vincent Poireau, Florian
Reindl, Federica Petricca, [iDMEu](#) proponents

 [@CatDogLund, she/her](#)

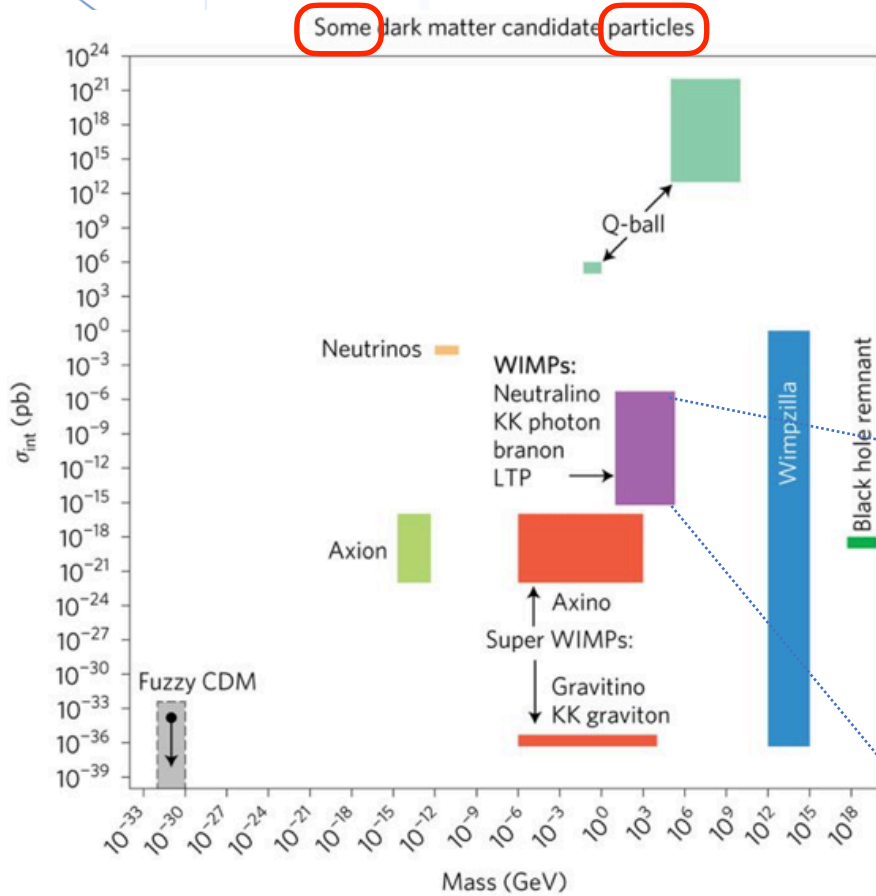
<http://www.hep.lu.se/staff/doglioni/>



ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.



Different kinds of dark matter



Many hypotheses for dark matter

→ many ways to detect it

→ many different experiments

→ many **different data / workflow needs**

→ many **different data / result sharing policies**

Starting point: **Weakly Interacting Massive Particles**

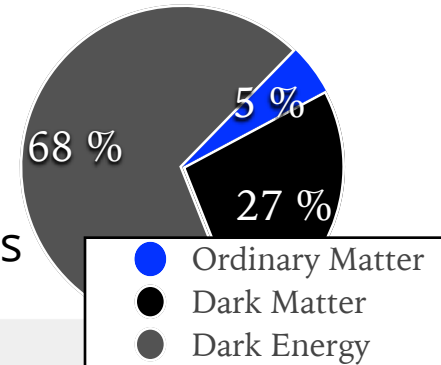
Well studied models with possibilities of extensions

Not yet completely excluded (could be *more rare*)

Complementary experiments could find it

Could make up all the DM in the universe

Searches for WIMPs can be extended/connected to many other models

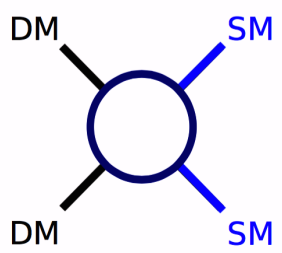


<https://www.nature.com/articles/nphys4049>

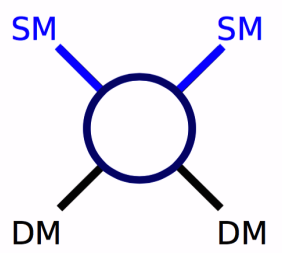
adapted from [The Dark Matter Scientific Assessment Group](#)

Dark matter complementarity

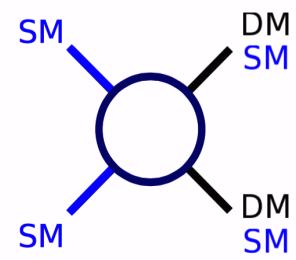
- Observations, experiments and theories all needed for DM discovery
 - DD/ID can discover DM with cosmological origin
 - Colliders / accelerators can produce DM and probe the dark interaction
 - Observations motivating DM come from astrophysics / gravitational interactions
 - Theoretical frameworks are necessary to put different observations in context



Indirect
Detection

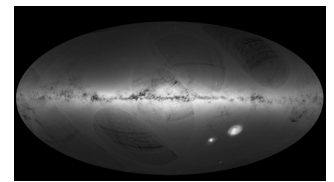


Direct
Detection

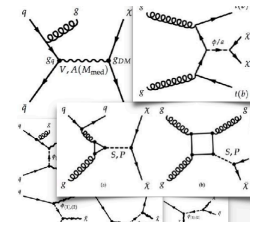


Collider &
Accelerator
(beam dump, fixed target...)

Credit: ESA/Gaia/DPAC.



Astrophysics



Theory



searches & interpretation

JENAS EoI: Initiative for Dark Matter in Europe and beyond: Towards facilitating communication and result sharing in the Dark Matter community (iDMEu)

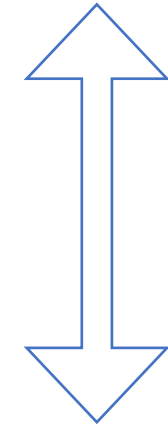
helps **define/compare common DM interpretations**
(& more, see [this ICHEP talk on Thu](#))



Common theory ground

build a discussion platform to facilitate collaboration of existing groups/efforts on **dark matter searches and interpretation**

<https://indico.cern.ch/event/869195/>
[ESCAPE newsletter](#) [APPEC newsletter](#)



instrumentation
(accelerators, beams, detectors,
vacuum & cryogenics,
control & automation...)

data acquisition,
software, computing,
data sharing
& open science



Towards a Dark Matter Test Science Project

[ESCAPE Progress Meeting, 2020](#)

compare **end-to-end analysis workflows** for WIMP searches, towards their implementation in a common **Software Catalogue** and as input to the design of the **European Open Science Cloud**

allows to **create experimental curves** in common DM plots with analysis pipelines that use *ESCAPE / EOSC tools*

software & data

More initiatives and links in backup slides

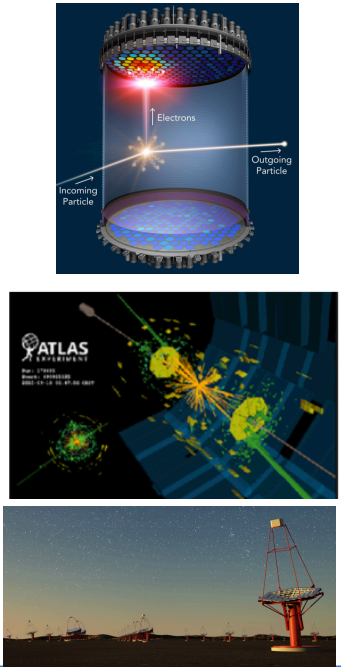


(Different) end-to-end WIMP analysis workflows

- Simplified abstraction of workflows to fit in this slide
- Different collaborations may implement software at different stages (e.g. not all boxes will come at the same time)

Generation & simulation of events

Experimental data



Data processing
(including reconstruction & calibration)

Analysis of events/
distributions
(including background subtraction, background estimation, statistical analysis)

Interpretation of results

Combination of results
with other searches/
experiments

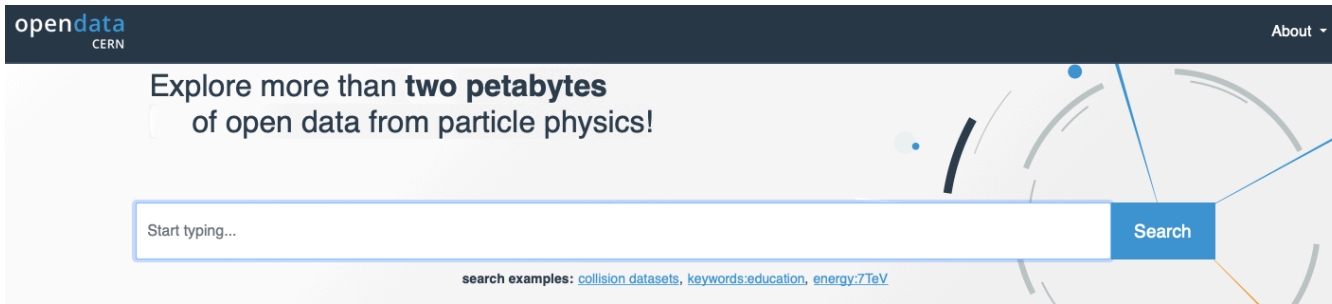
Comparison of results
with other searches /
experiments



Open pipelines @ ATLAS: HOWTO?

"Making data FAIR is relatively straightforward, but making FAIR data useful is difficult"

from a conversation with Stephen Serjeant, Hugh Dickinson



ATLAS: released partial dataset for educational purposes, may do more (many "sociological" questions in a collaboration of 3000 people...)

Many different possible use cases, e.g.:

- **preserve** data & pipelines (RECAST/REANA - [see HSF talk by L. Heinrich](#))
- **re-analyse** data with different pipelines
- **combine** data of different experiments
- **reinterpret** data for different DM models (pyhf / CONTUR / ... - [see this document](#))

How to deploy *diverse* software on EOSC?

How to facilitate user participation (documentation, documentation, documentation)?

How to ensure lifetime of tools beyond "postdoc contracts"?





ESCAPE

European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures

Backup slides

 [@CatDogLund](https://twitter.com/CatDogLund), she/her

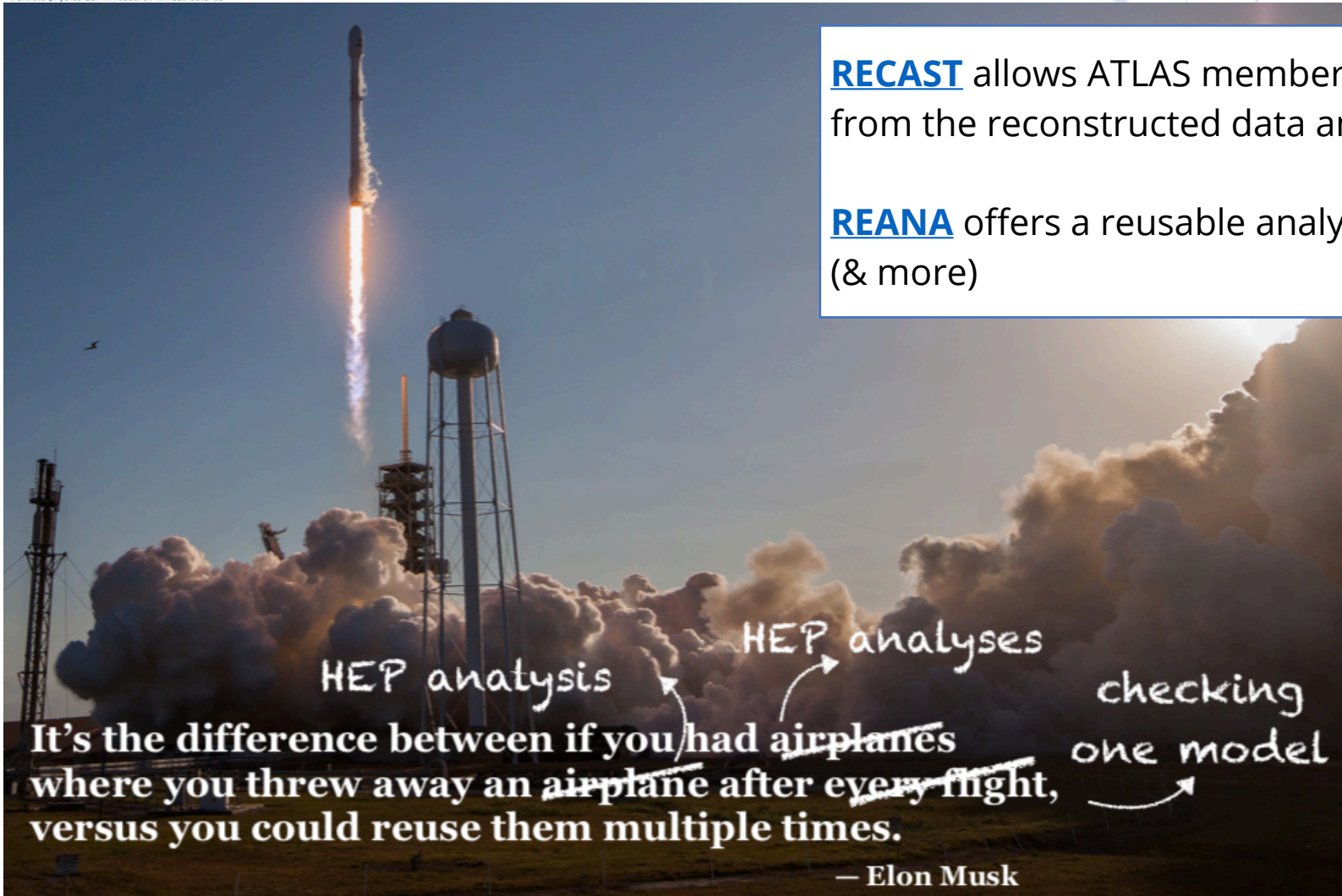
<http://www.hep.lu.se/staff/doglioni/>



ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.



The ideas behind RECAST / REANA



RECAST allows ATLAS members to preserve the analysis pipeline starting from the reconstructed data and rerun it easily using different (DM) models

REANA offers a reusable analysis platform to run the RECAST pipelines (& more)

All HEP/CERN software is Open Source

DM TSP for HEP:

1. implement a dark matter analysis pipeline in RECAST/REANA
2. adapt RECAST/REANA to work within ESCAPE/EOSC ecosystem (Data Lake, Software Catalogue, Science Analysis Platform)
3. produce results for a variety of DM interpretation, as input of iDMEu

[Lukas Heinrich's lightning talk @ S212](#)



TSP in more detail

for now: L. Heinrich, S. Meehan, K. Cranmer, C. Doglioni
open to others if interested!

Data sharing & processing

Follow updates to CERN-wide data sharing policies (<http://opendata.cern.ch>)

Benefit from HEP Software Foundation as platform to understand shared solutions for data processing challenges & interactions with ESCAPE software catalogue



Data analysis & interpretation

Start working on test "generic DM search":

data analysis & data preservation

- Preserve workflow & analysis code with **RECAST**
[ATLAS Note](#), [docs](#) and [REANA](#)
 - Built around the idea of containerized workflows
 - Interface containers to ESCAPE "Data lake"
- Preserve likelihood with **pyhf** [Zenodo](#), [docs](#)
 - Discussions ongoing with **Fermi** data analysers

interpretation of results

- Deposit digitized data & likelihoods in [HEPData](#)
- Include LHC measurements with [CONTUR](#)
- Could use [GAMBIT](#) (& [DDCalc](#)) for combinations

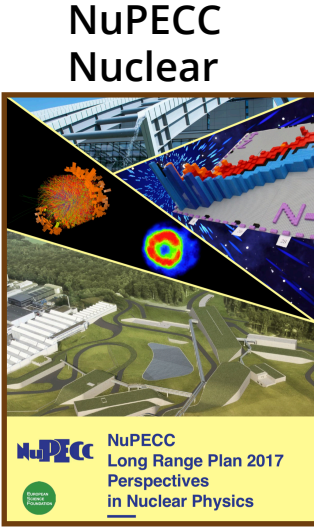
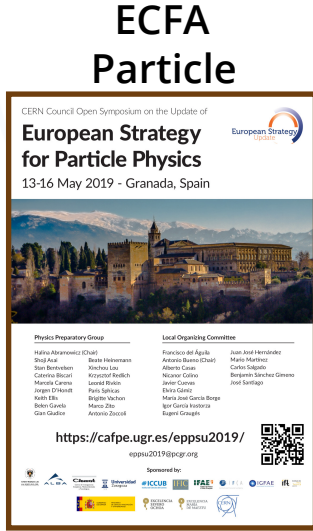
Concluding thoughts (from WP3)

Collected from chats with members of ESCAPE/DM community

- DM Test Science Project: ESCAPE/EOSC reaching out to researchers
 - Demonstrator of selected end-to-end pipelines for direct, indirect detection and collider searches for Dark Matter integrated in ESCAPE/EOSC infrastructure
 - Pipelines produce results that will help characterise discovery / constrain Dark Matter
 - Integrated in the JENAS Expression of Intent [iDMEu](#)
- Some open questions for discussion:
 - How to implement infrastructures that are ready for ML-aware pipelines (eg accelerators)?
 - How will this TSP interface itself with other entities that support/develop DM research / open science in astronomy and astrophysics?
 - E.g. ESA, http://www.esa.int/About_Us/Digital_Agenda/Open_Science
 - It is our duty as scientists to make our research FAIR (and our code sustainable)
 - But do we (PIs) / the system (funding agencies) offer sufficient reward?
 - A concern of many: maintaining code is necessary but is often done on a voluntary basis
 - Need a **healthy system of incentives** including permanent positions for these tasks



Foundations needed to exploit synergies



Common theory ground

instrumentation
(accelerators, beams,
detectors, vacuum &
cryogenics,
control & automation...)

data acquisition,
computing,
data sharing
& open science



& more...

EuCAPT

Talk at EPS-HEP / ECFA session 2019, [CERN EP Newsletter](#)



- Many DM discussions, from **Granada** to the **ApPEC-ECFA-NuPECC JENAS meeting** held in Orsay in October 2019
 - Talk on ESCAPE (G. Lamanna) in plenary programme
 - [HEP Software Foundation meeting](#) on possible software synergies



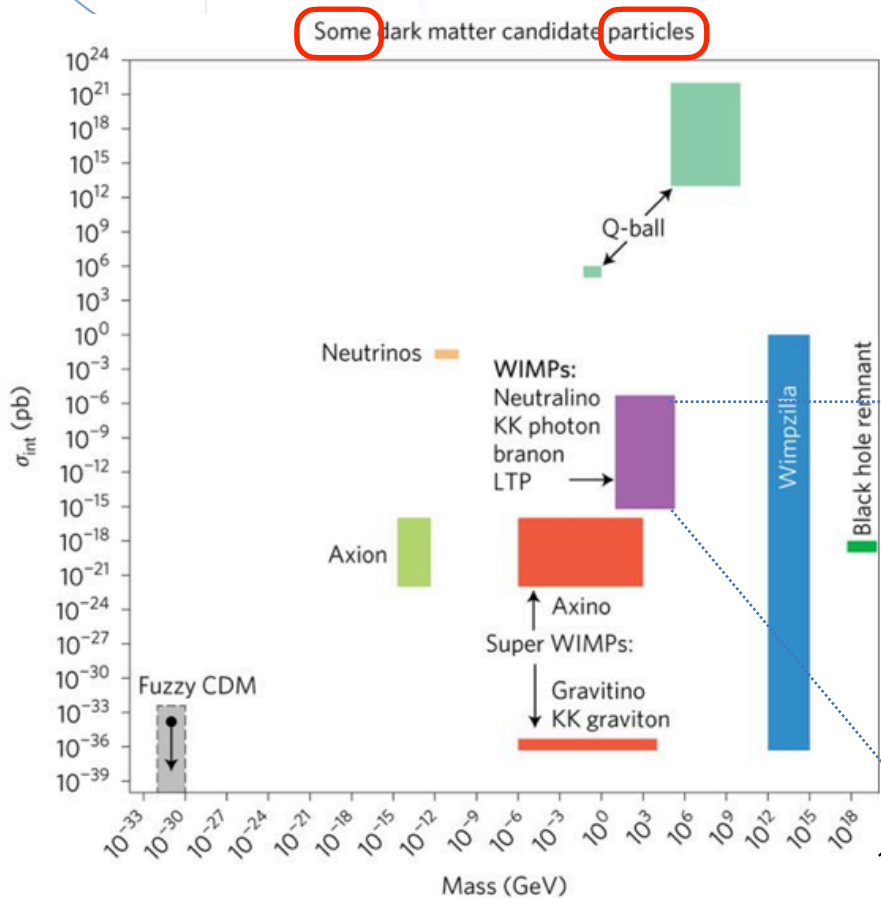
- JENAS prompted a new initiative centered around **dark matter**: <https://indico.cern.ch/e/iDMEu>, also featured in ESCAPE [newsletter](#)
 - *iDMEu* aiming to build a discussion platform to facilitate collaboration of existing groups/efforts
 - *Dark Matter Test Science Project* targeting data, software and tools sharing where necessary/useful
 - Points of contact between *iDMEu* and *TSP*:
 - participation of DM community to software catalogue
 - help with common repositories of data and final results (e.g. versioning)
 - e.g. [DMTools](#), [DM Limit Plotter](#)



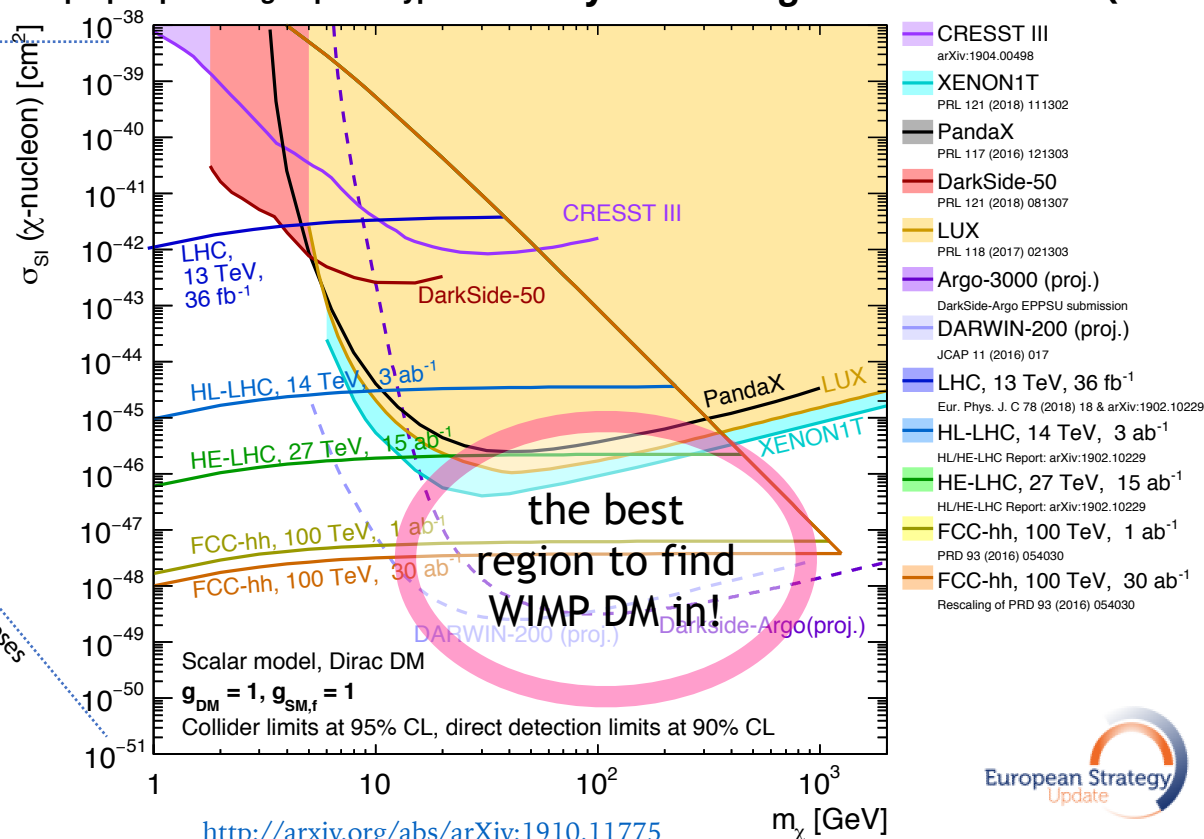
Different kinds of DM, and synergies

Many hypotheses for dark matter

- many ways to detect it
- many different experiments
- many different data / workflow needs
- many different data / result sharing policies



example plot predicting a specific type of **Weakly Interacting Massive Particles (WIMP)**



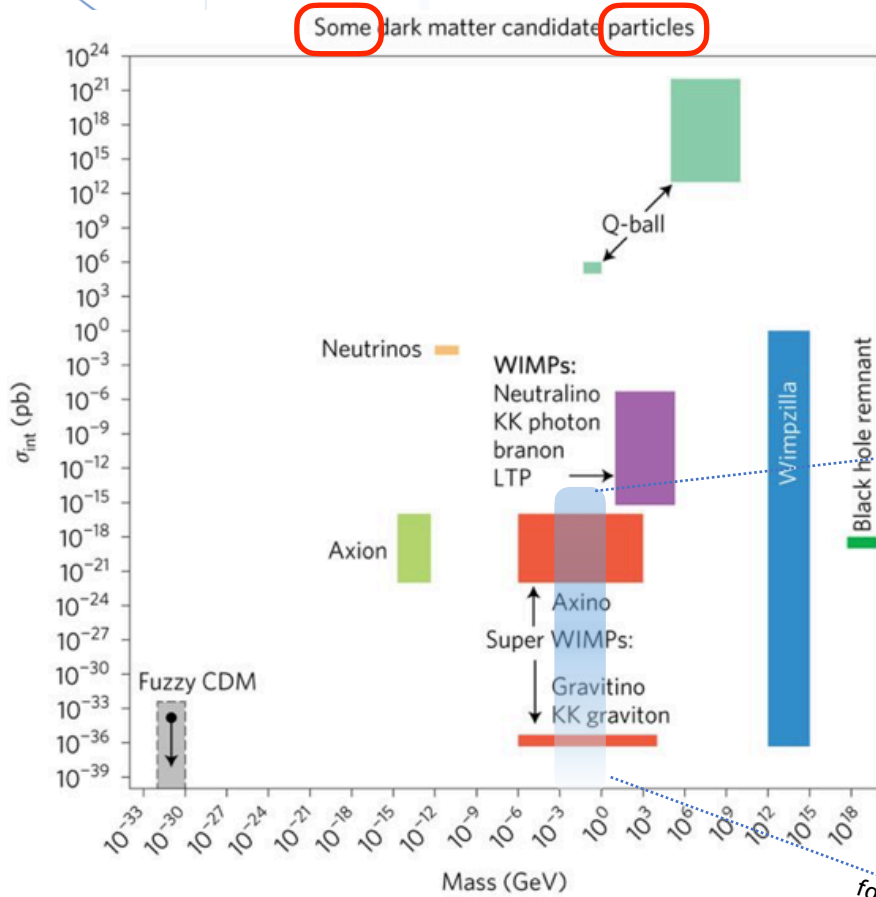
for illustrative purposes
no 1:1 correspondence

<https://www.nature.com/articles/nphys4049>
adapted from [The Dark Matter Scientific Assessment Group](#)

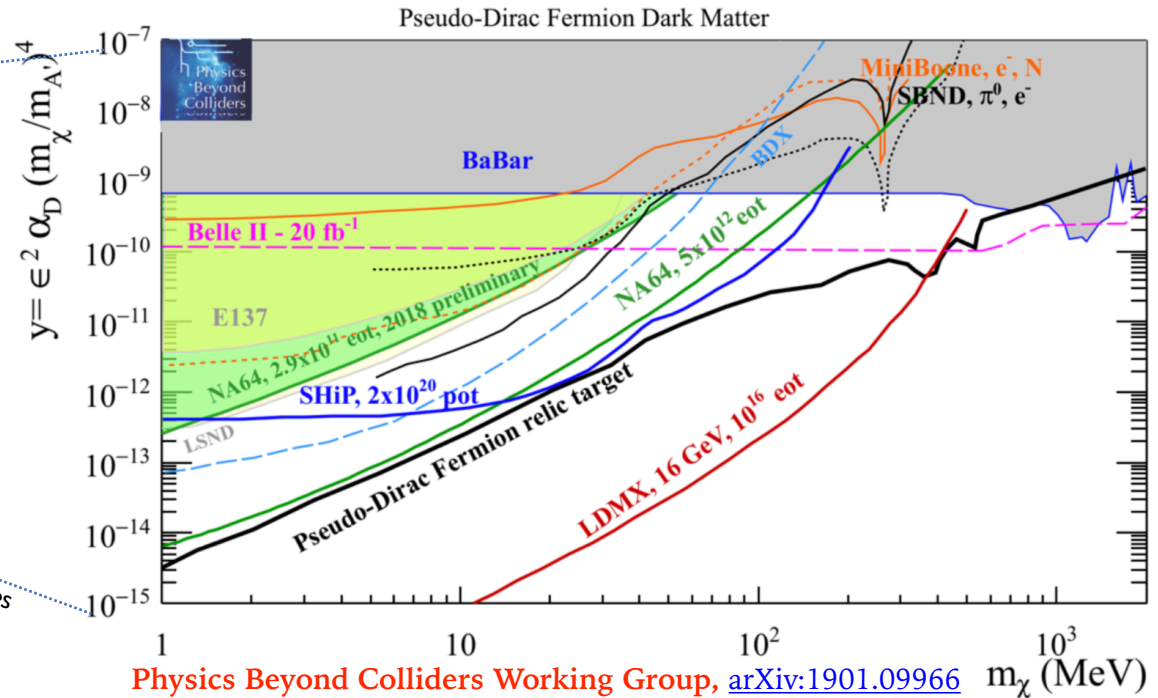
Different kinds of DM, and synergies

Many hypotheses for dark matter

- many ways to detect it
- many different experiments
- many different data / result sharing policies
- many different data / workflow needs



Dark photon portal model



<https://www.nature.com/articles/nphys4049>

adapted from [The Dark Matter Scientific Assessment Group](#)

for illustrative purposes
no 1:1 correspondence

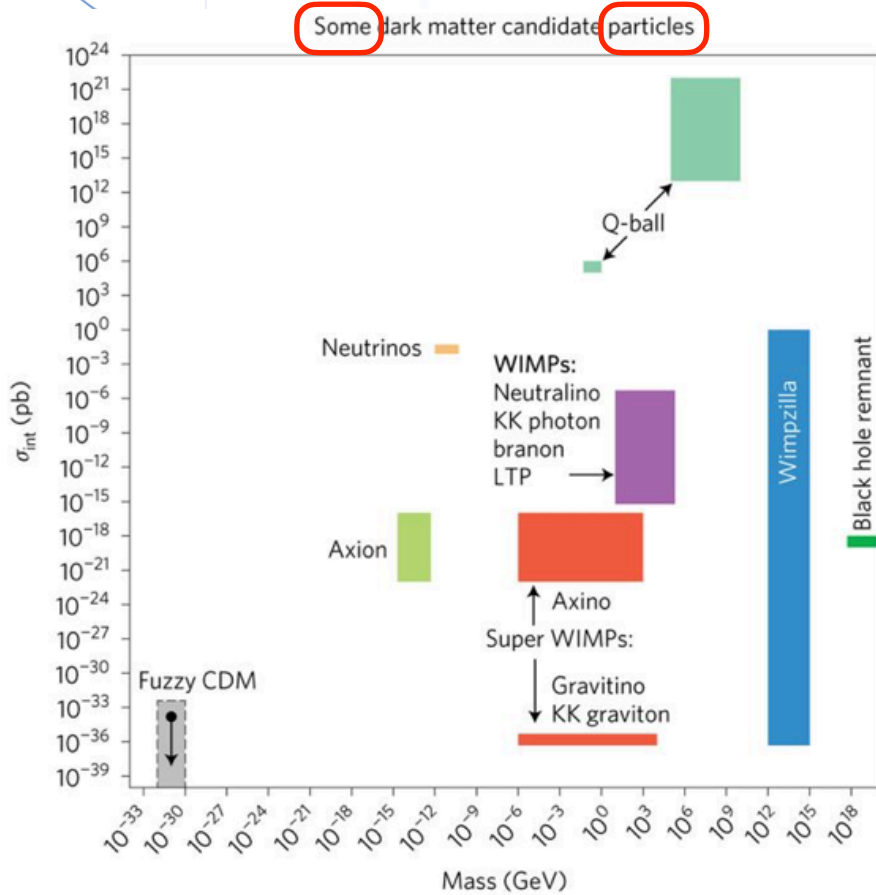
Physics Beyond Colliders Working Group, [arXiv:1901.09966](https://arxiv.org/abs/1901.09966)



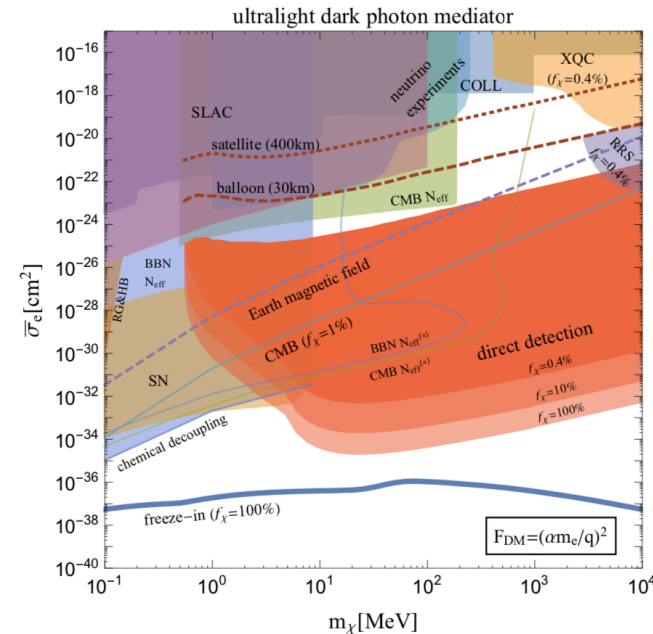
Different kinds of DM, and synergies

Many hypotheses for dark matter

- many ways to detect it
- many different experiments
- many different data / result sharing policies
- many different data / workflow needs



Constraints from astrophysics



<https://www.nature.com/articles/nphys4049>

adapted from [The Dark Matter Scientific Assessment Group](#)

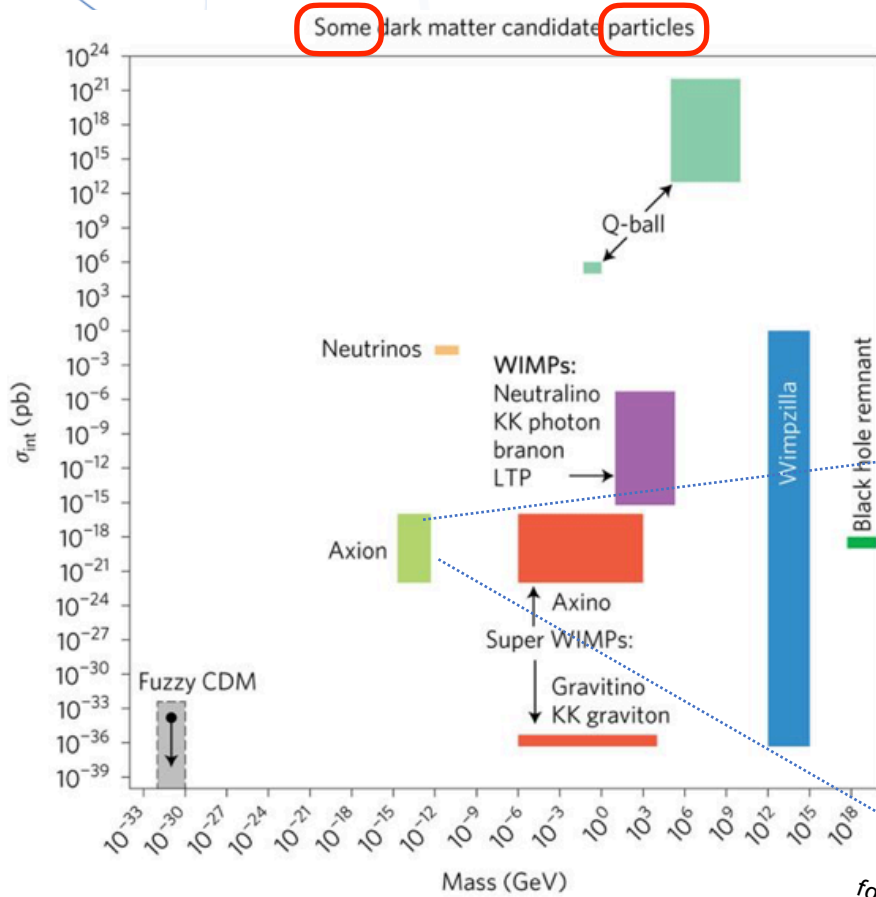
Physics Beyond Colliders Working Group, [arXiv:1901.09966](https://arxiv.org/abs/1901.09966)



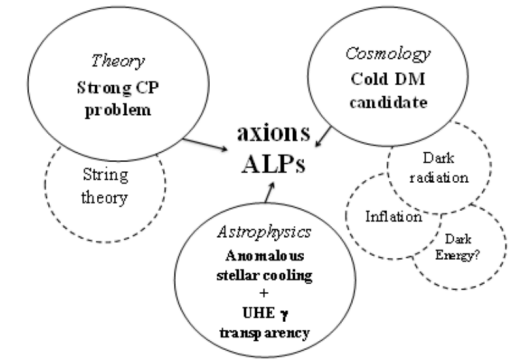
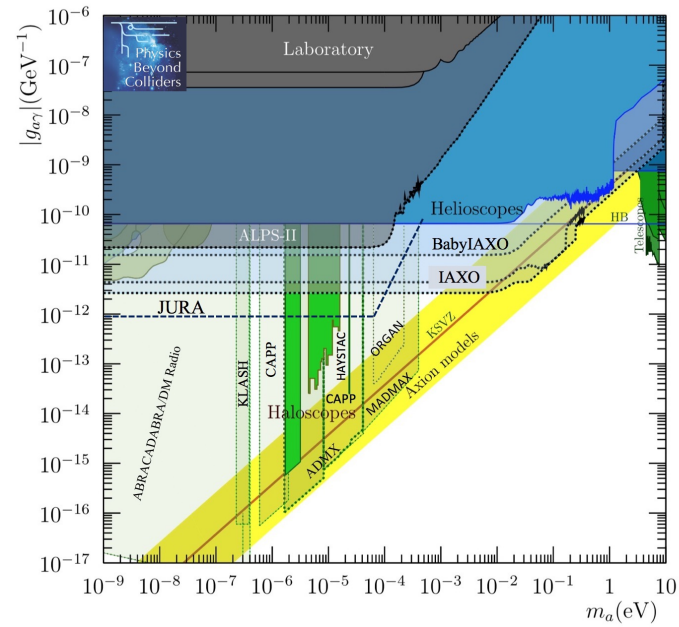
Different kinds of DM, and synergies

Many hypotheses for dark matter

- many ways to detect it
- many different experiments
- many different data / result sharing policies
- many different data / workflow needs



Axion models



for illustrative purposes
no 1:1 correspondence

<https://www.nature.com/articles/nphys4049>
adapted from [The Dark Matter Scientific Assessment Group](#)

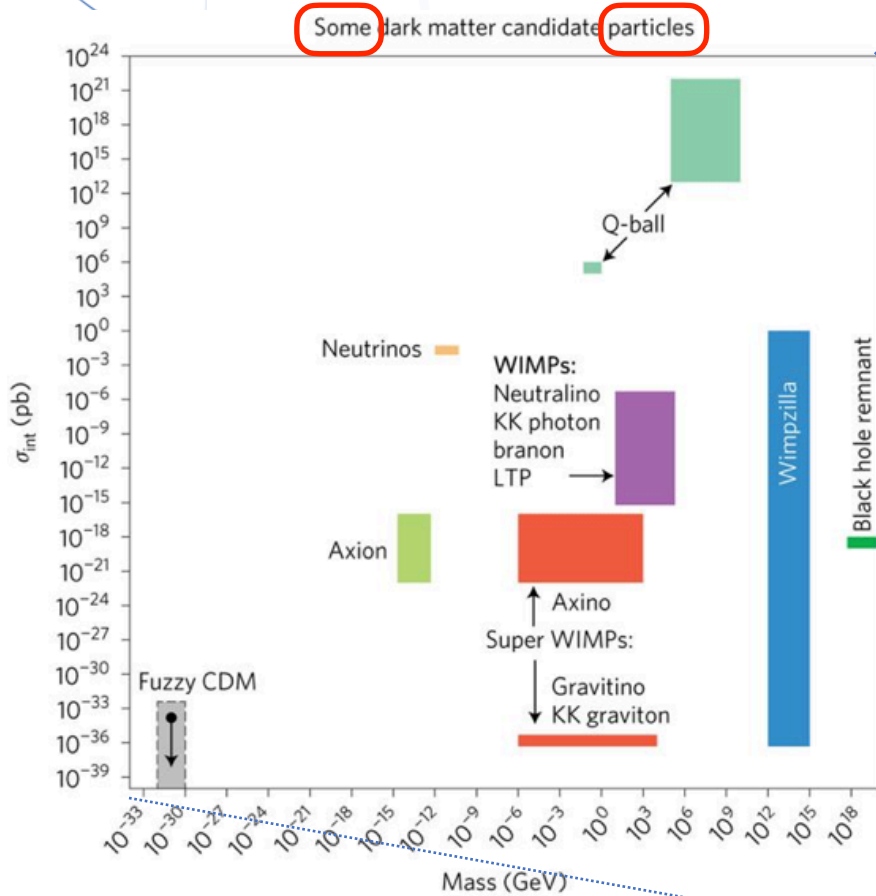
Physics Beyond Colliders Working Group, [arXiv:1901.09966](https://arxiv.org/abs/1901.09966)



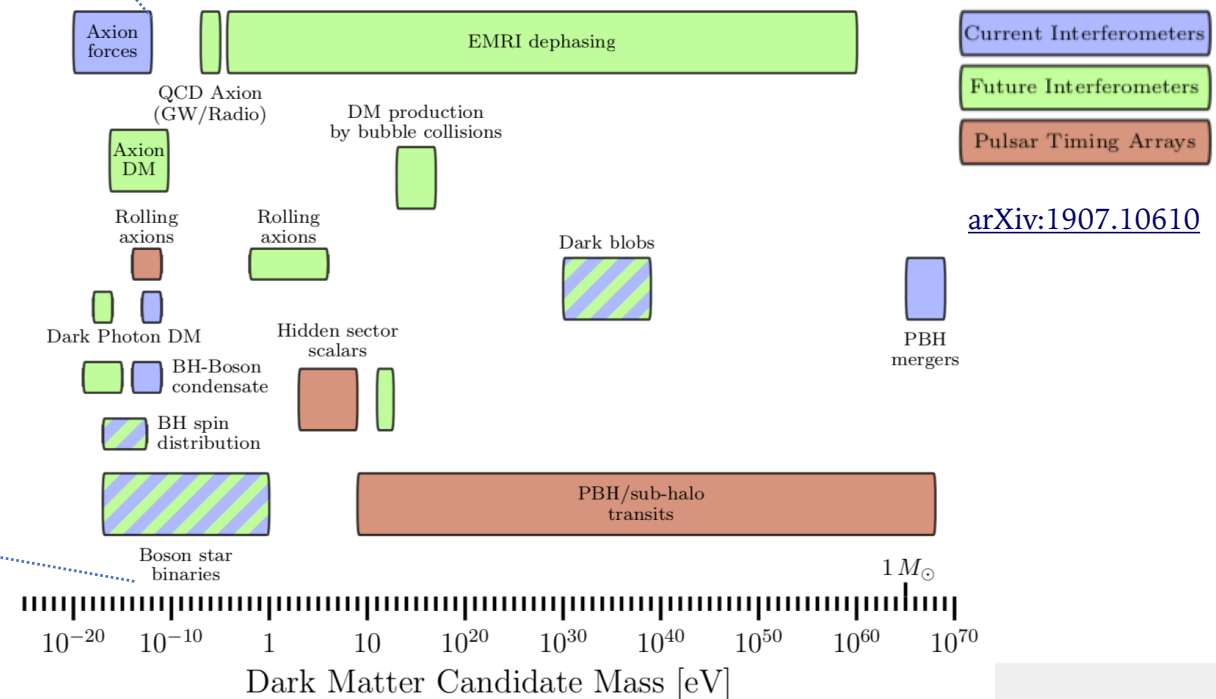
Different kinds of DM, and synergies

Many hypotheses for dark matter

- many ways to detect it
- many different experiments
- many different data / result sharing policies
- many different data / workflow needs



Models relevant for GW experiments



[arXiv:1907.10610](https://arxiv.org/abs/1907.10610)

<https://www.nature.com/articles/nphys4049>

adapted from [The Dark Matter Scientific Assessment Group](#)

Different kinds of (WIMP) communities

C. Tunnell, HSF/WLCG/OSG workshop 2019

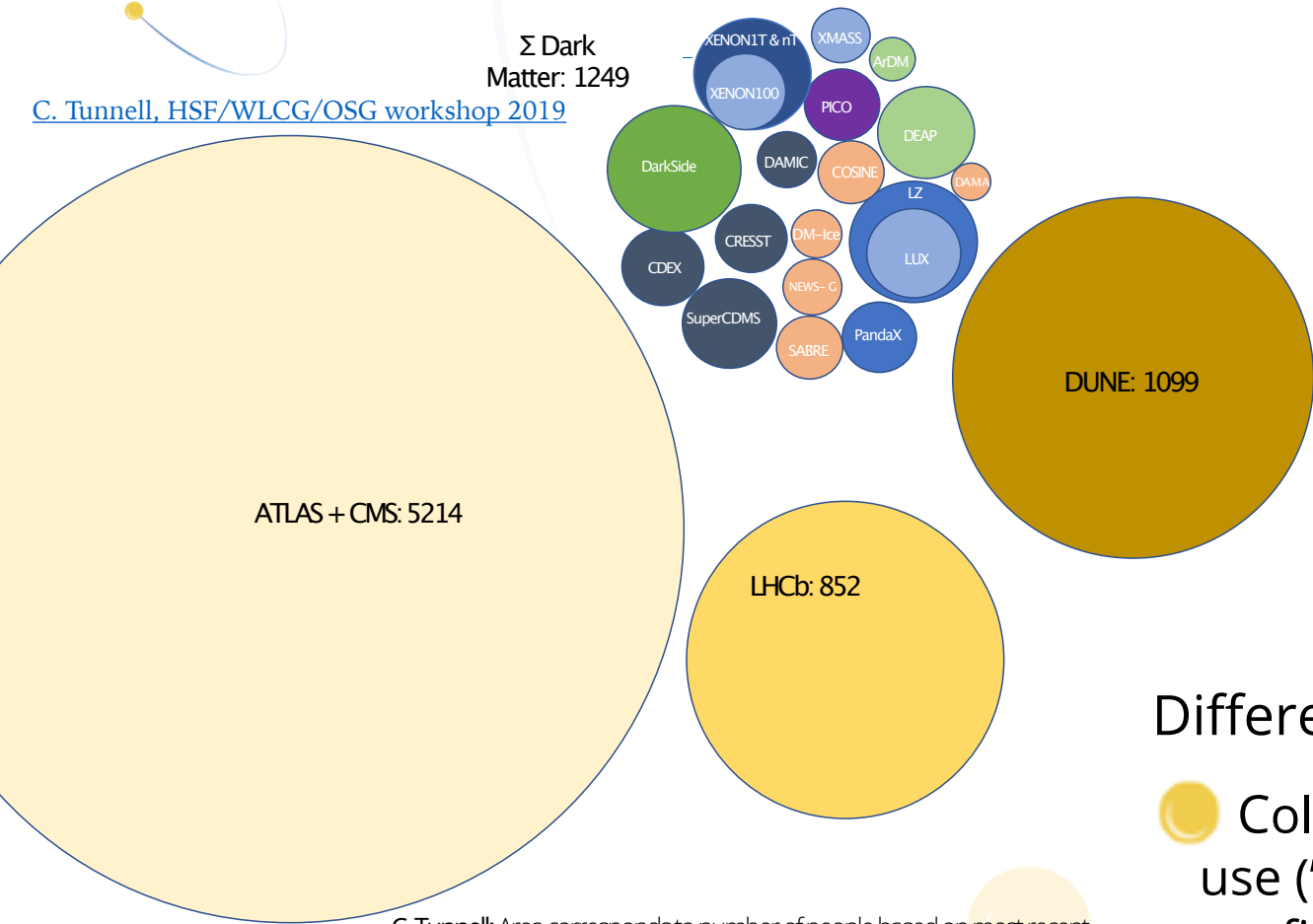


Diagram only representing **collider and direct detection**

- Differences in collaboration variety and size
- Differences in data volumes:
 - Colliders: "Big Data" volumes (>> PB)
 - DD: smaller data volumes (~TB/PB)
- Synergies in statistical analysis and interpretation of results

Different modus operandi for **indirect detection**

- Collaborations e.g. Fermi release data for general use ("observatory mode"), but also perform high-profile analyses themselves

C. Tunnell: Area corresponds to number of people based on most recent publication from any experiment that has published scientific papers in the last two years. This relied on Inspire-HEP. See gist for calculation notes. 16/ March/2019

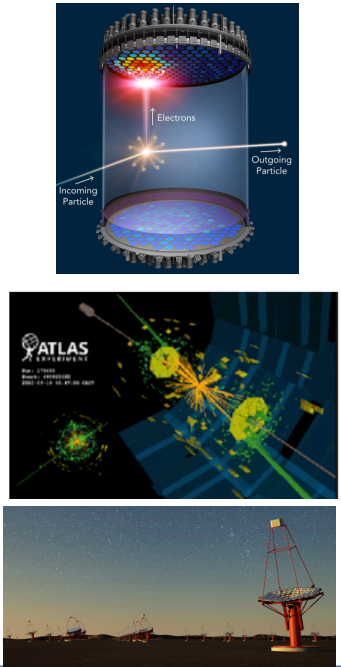


(Different) end-to-end WIMP analysis workflows

- Simplified abstraction of workflows to fit in this slide
- Different collaborations may implement software at different stages (e.g. not all boxes will come at the same time)

Generation & simulation of events

Experimental data



Data processing
(including reconstruction & calibration)

Analysis of events/
distributions
(including background subtraction, background estimation, statistical analysis)

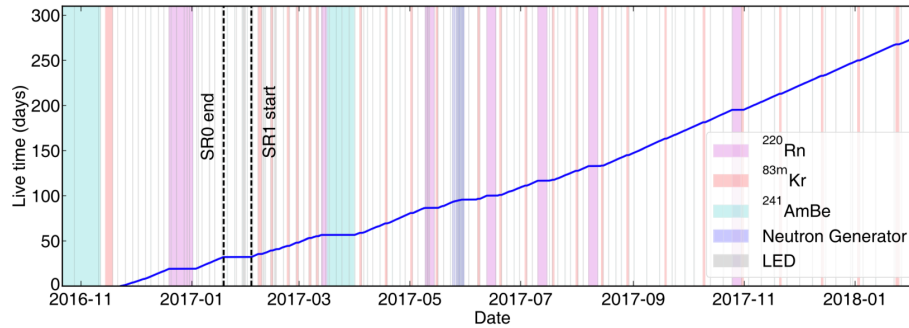
Interpretation of results

Combination of results
with other searches/
experiments

Comparison of results
with other searches /
experiments



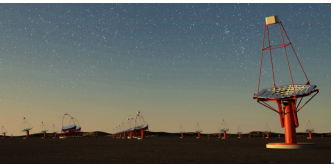
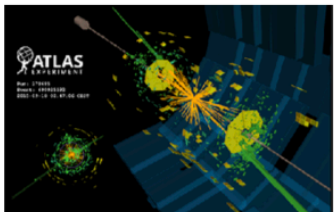
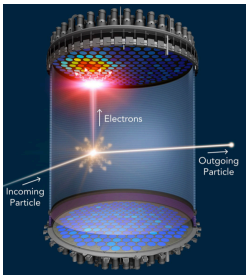
(Different) end-to-end WIMP analysis workflows



[XENON 1T, PRD 100, 052014 \(2019\)](#)

Generation & simulation of events

Experimental data



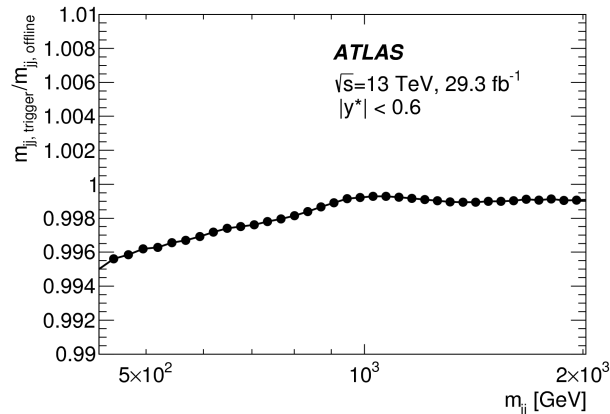
Data processing
(including reconstruction & calibration)

Analysis of events/ distributions
(including background subtraction, background estimation, statistical analysis)

Interpretation of results

Combination of results
with other searches/
experiments

Comparison of results
with other searches /
experiments

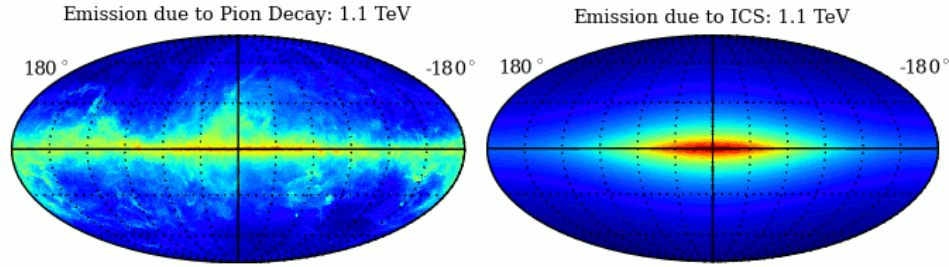


[ATLAS, Phys. Rev. Lett. 121, 081801 \(2018\)](#)



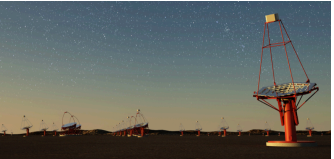
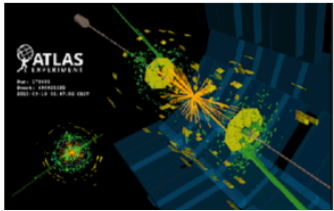
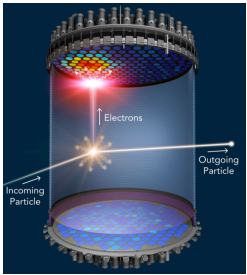
(Different) end-to-end WIMP analysis workflows

Generation & simulation of events



Credit: Galprop, HAWC website

Experimental data



Data processing
(including reconstruction & calibration)

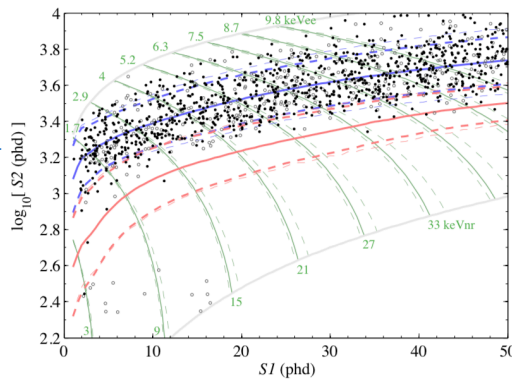
Analysis of events/
distributions
(including background subtraction, background estimation, statistical analysis)

Interpretation of results

Combination of results with other searches/
experiments

Comparison of results with other searches /
experiments

[PhyStatDM, 2019](#)
[LUX, PRL. 118, 021303 \(2017\)](#)

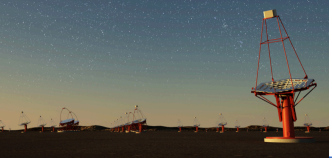
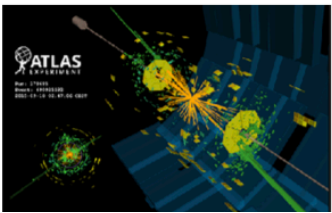
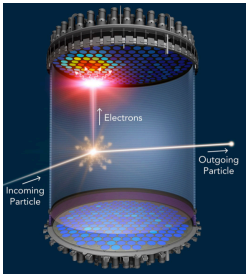


(Different) end-to-end WIMP analysis workflows

<https://arxiv.org/abs/1704.03910>

Generation & simulation of events

Experimental data



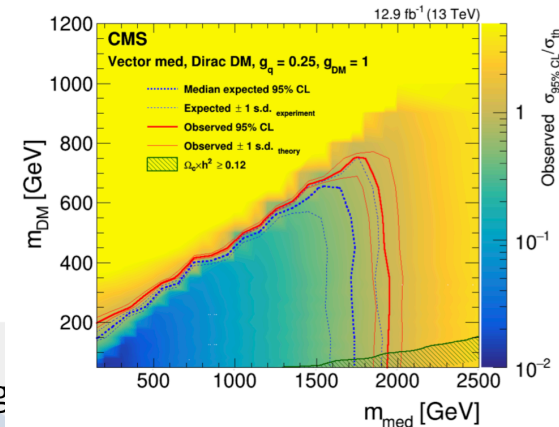
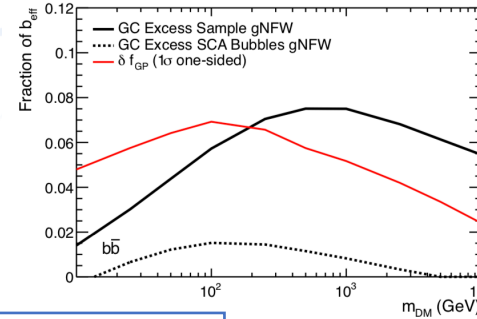
Data processing
(including reconstruction & calibration)

Analysis of events/ distributions
(including background subtraction, background estimation, statistical analysis)

Interpretation of results

Combination of results
with other searches/
experiments

Comparison of results
with other searches /
experiments



Caterina Dog

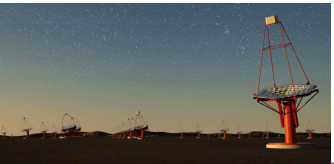
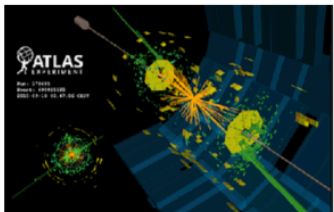
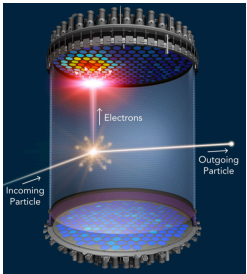
7/07/2020



(Different) end-to-end WIMP analysis workflows

Generation & simulation of events

Experimental data



Data processing
(including reconstruction & calibration)

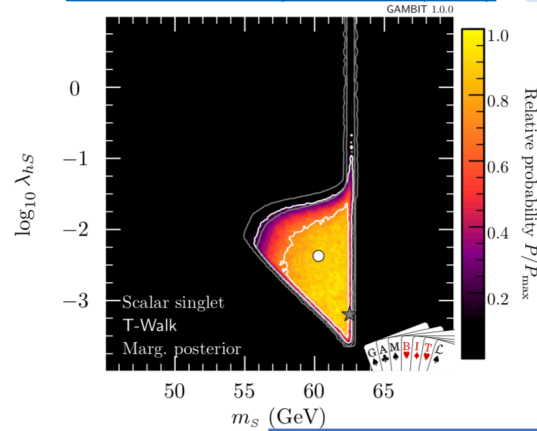
Analysis of events/
distributions
(including background subtraction, background estimation, statistical analysis)

Interpretation of results

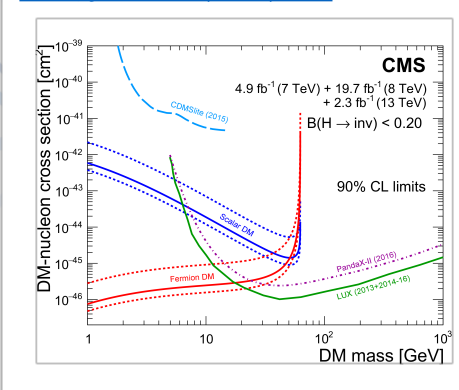
Combination of results with other searches/experiments

Comparison of results with other searches / experiments

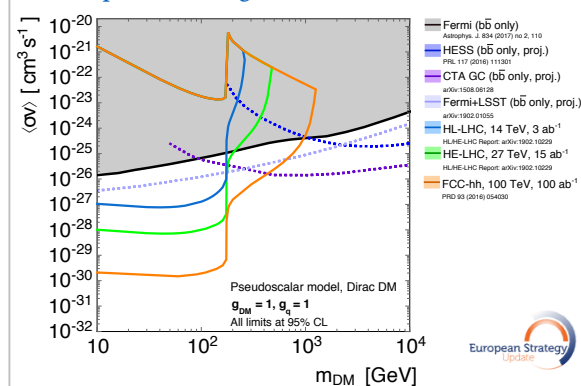
GAMBIT Coll., EPJC 77, 568 (2017)



CMS, JHEP 02 (2017) 135



<http://arxiv.org/abs/arXiv:1910.11775>



Caterina

7/07/2020

Challenges for Test Science Project

- Not possible to find a one-size-fits-all solution in either case...so work in parallel
- Idea (not original, see [DANCE workshop](#)): review what is done by various collaborations, finding points of contact

Generation & simulation of events

Experimental data

Data sharing and data processing challenges

Data processing (including reconstruction & calibration)

RUCIO data management software [shared](#) across collaborations (LHC/ DUNE/KM3NeT/...)

Data analysis, preservation and interpretation challenges

Analysis of events/ distributions (including background subtraction, background estimation, statistical analysis)

Interpretation of results

Combination of results with other searches/ experiments

V. Poireau et al.

Ongoing work between Fermi-LAT, HAWC, HESS, MAGIC & VERITAS [GitHub](#), [ICRC Proceedings](#)

Comparison of results with other searches / experiments

ESCAPE WP2, WP4, WP5

ESCAPE WP3, WP5, WP6



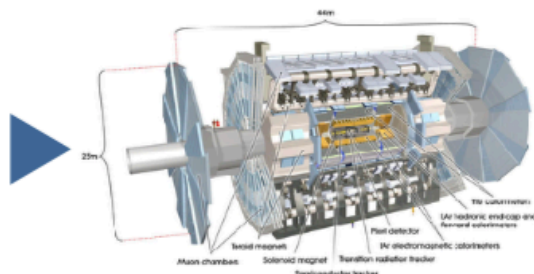
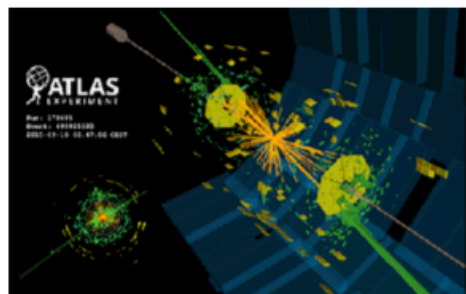
Where the ATLAS experiment pipelines take place

Collisions at ~ 40 MHz
(~ 1 MB of info each)

Hardware trigger
outputs ~ 100 kHz

Software trigger
outputs ~ 1 kHz

**Enormous amount of data
to manage and analyse**



Online **Offline**

Event selection
(trigger)

Object
reconstruction
and calibration

Data analysis

On-detector
hardware processors
+
Trigger farm
@ CERN

Computing Farm
@ CERN

Worldwide LHC
Computing Grid
+
Local computing farms
+
Laptops!

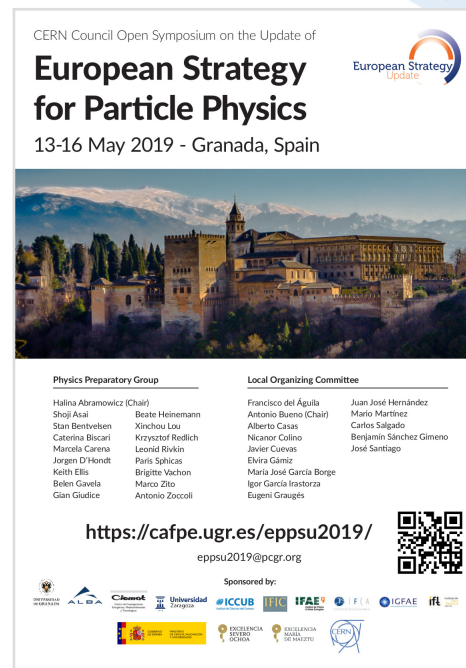
Different pipelines for each of the steps

Support for synergistic initiatives

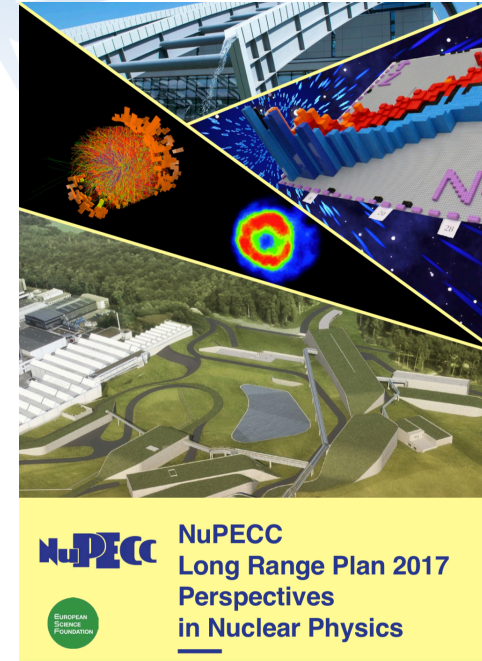
Astroparticle (APPEC)



Particle (ECFA)



Nuclear physics (NuPECC)

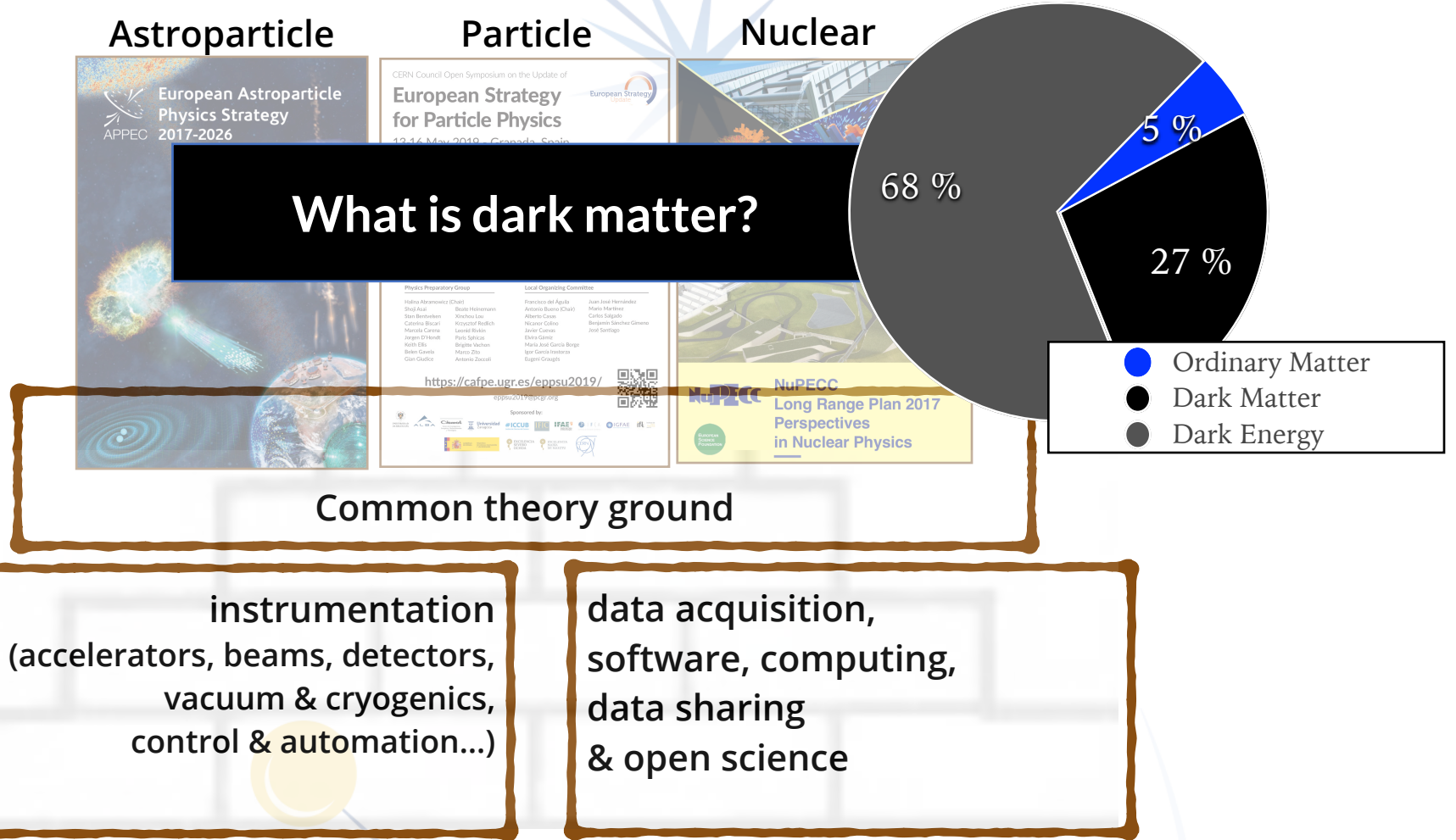


Astroparticle, particle and nuclear physics in Europe have **strategies and plans** that **recognize the importance of synergies** between the different fields

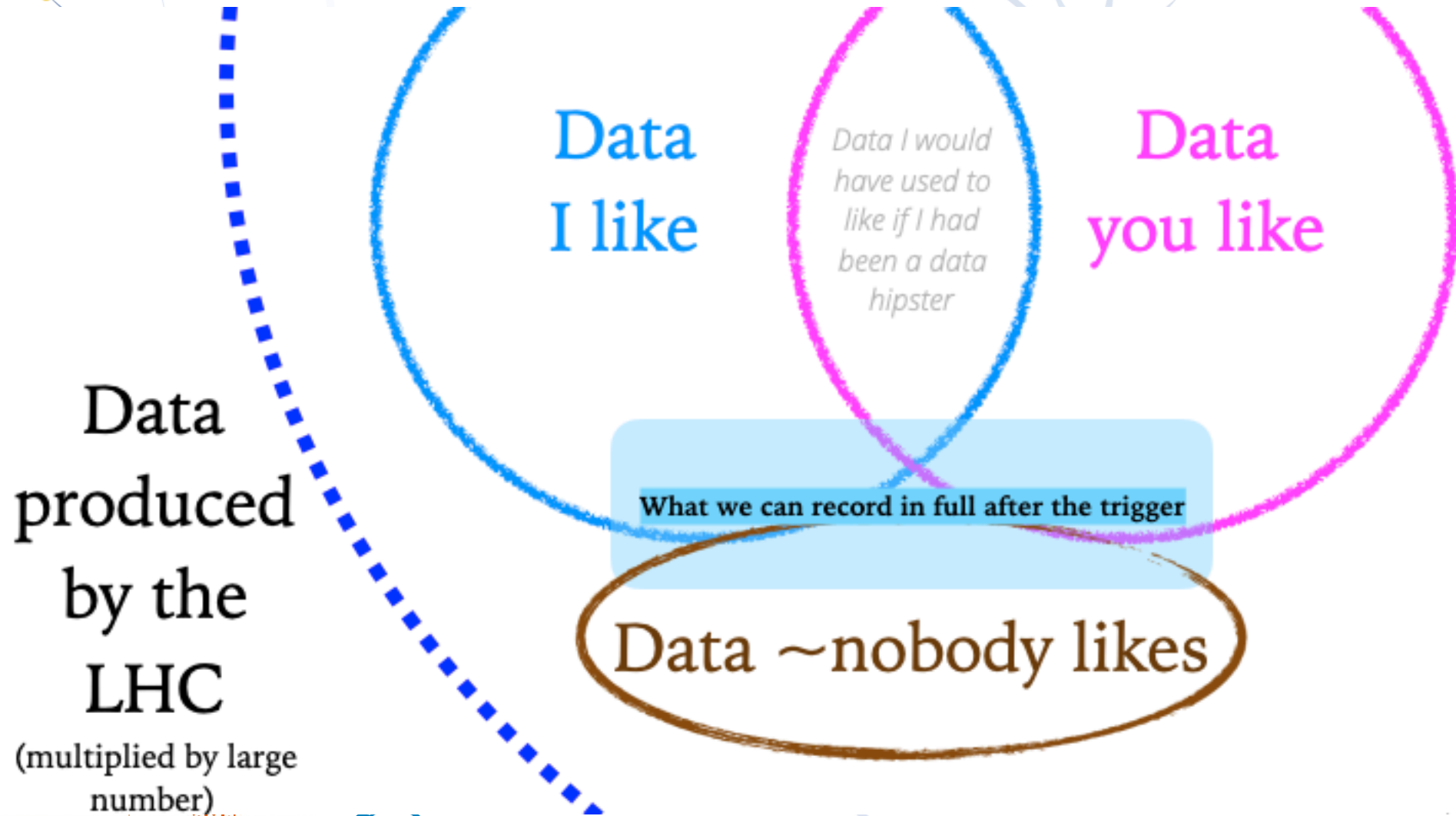
US: *Snowmass* effort starting soon, (for *DM at Colliders* WG: )



Foundations for scientific questions



Part of the ATLAS data problem...



“How to make the most of the data with real-time analysis”: not the topic of the TSP, but the topic of my research

