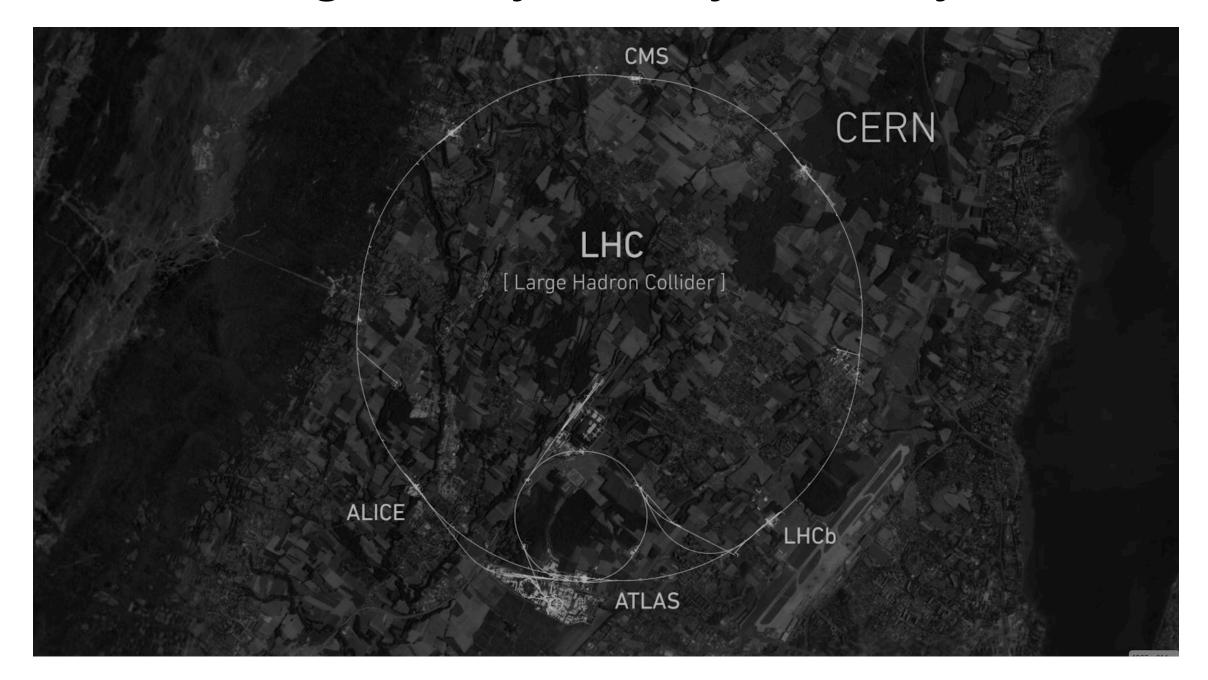
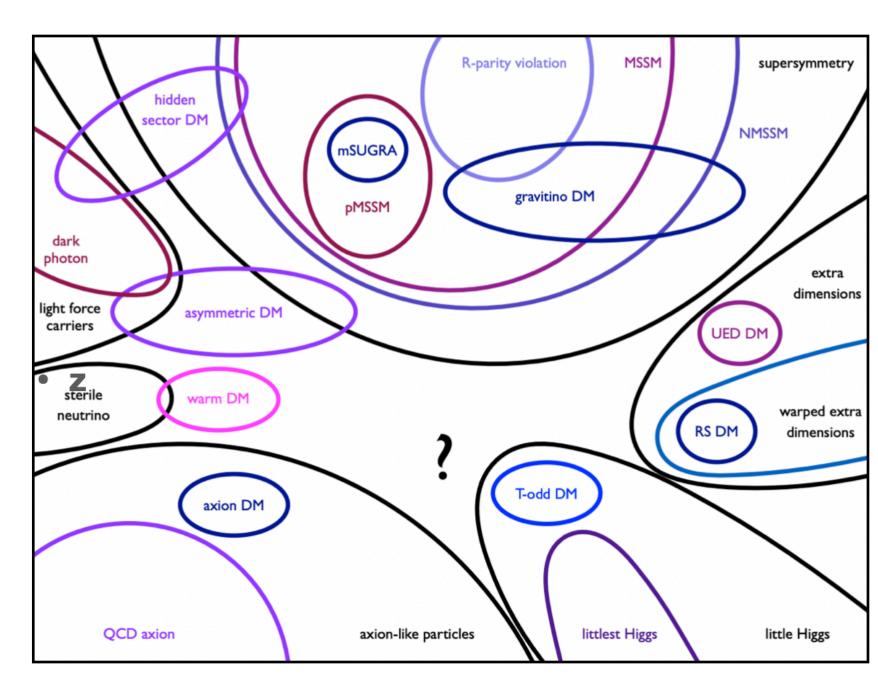
Pheno Interpretation at the LHC

Key Problem: Many variants of Physics beyond the Standard Model are proposed by theorists, but only a few experiments can check them

Answering "is my theory already excluded by the LHC" is difficult for outsiders



Experiment



Theories

Reinterpretation

By preserving and reusing analysis pipelines, we can make answering this

question accessible to the full community HEPData Original Use-case that led to the development of REANA DRPs from archived DataHep data modify & new new new analysis gen, sim & reco result model infer sample model analysis-independent **Analysis Preservation**

archive





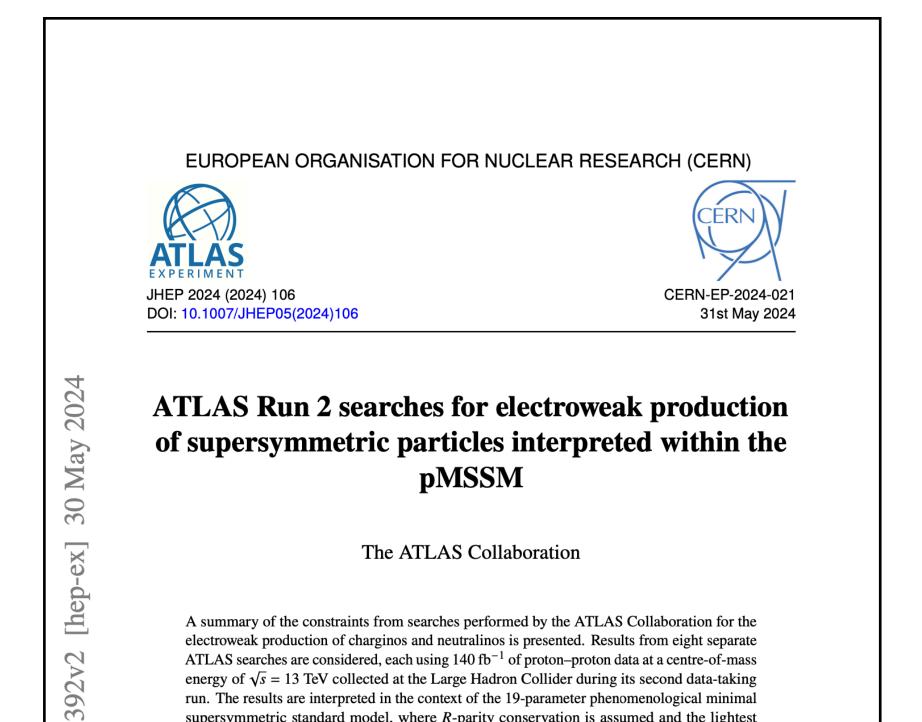
rerun

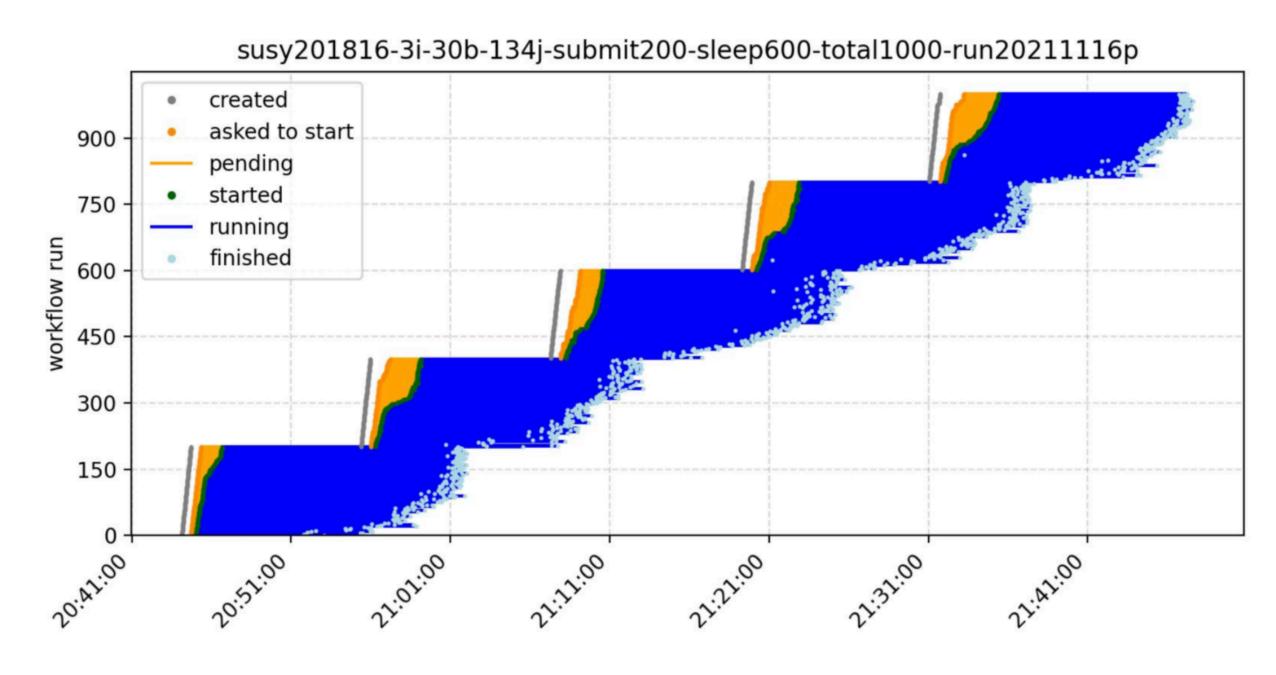
Status Quo and Scaling Tests with REANA

Reinterpretation used successfully internally by LHC Experiments to explore high-dimensional theory parameter spaces. Execution of over O(10k) RANA workflows.

But so far no public infrastructure and interfaces exist to do this.

→ Opportunity for PUNCH to fill an important gap





Questions

Which problems do they solve?

Enables the wider community to answer "is theory X excluded by the LHC" using production grade archived workflows from the experiment

Which gaps do they fill?

So far this has only been done inside of the experiments and not yet on Open Data or using Open Infrastructure

How could they be generalized / what is generic?

Creating a science-driven UI for re-use of data pipelines is a generic requirement. It serves as a example on how to capitalize on rich data products created by large-scale experiments

End-to-end FAIR use case?

This would be the first time we can demonstrate reinterpretation at the LHC using newly available Open Simulation and Open Data

Connection to DRP?

This uses a combination of open workflows (reana), open software (analysis code), open environemnts (docker images), open likelihoods (HepData)

Questions

How is the use case viewed from outside of PUNCH?

There is a broad interest in LHC interpretation (e.g. via the LHC Reinterpretation Working Group) and it's recognized need for the community (esp. given increasing use of AI in HEP). It's been the main driver for development of REANA etc. Also used in context of ESCAPE & EOSC

Define input – procedure – output? Specify workflow.

- 1. User can request new signal to be produced according to a theory specification (e.g. SUSY Model)
- 2. Computing Infrastructure runs simulation up to event reconstruction
- 3. User can select from a catalogue of preserved analyses for reinterpretation
- 4. Workflow is run on REANA using DRPs fetched from HepData and new simulation input
- 5. Result is presented to the user

Operation model? - what is required for the use case, and how to organize that?

Requires large-scale computing resources for simulation execution
Access to Open Data and Open Simulation (from CERN Open Data Portal)
Repository for preserved Workflows (either in PUNCH or e.g. CERN Analysis Preservation Portal)
Large-scale REANA instance for running preserved workflows
Web Services for User-Facing UI

Sustainability

The re-use of preserved pipelines minimized the amount of resources required to produce a new, gold-standard scientific result for a given set of theories, as only a new signal must be simulated and existing data and background estimates are re-used