

# Trusting dark matter results

---

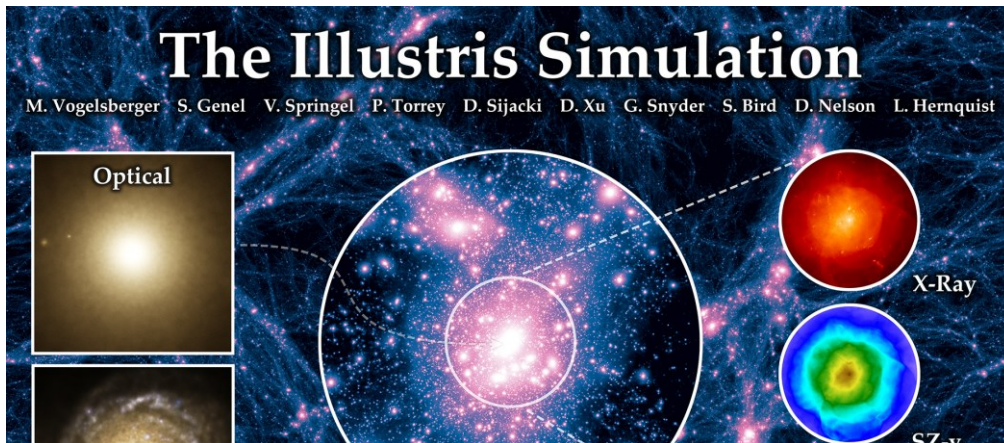
THREE SHORT STORIES ABOUT INCREASING TRUST IN THE DIFFICULT  
REALM OF DARK MATTER ANALYSIS

CONTACT: [AMY.ROBERTS@UCDENVER.EDU](mailto:AMY.ROBERTS@UCDENVER.EDU)

# Dark Matter

A simulated universe with a weakly-interacting, ~~40 GeV~~ particle  $\sim 0.5 \text{ MeV}/c^2$  mass density looks a lot like our universe.

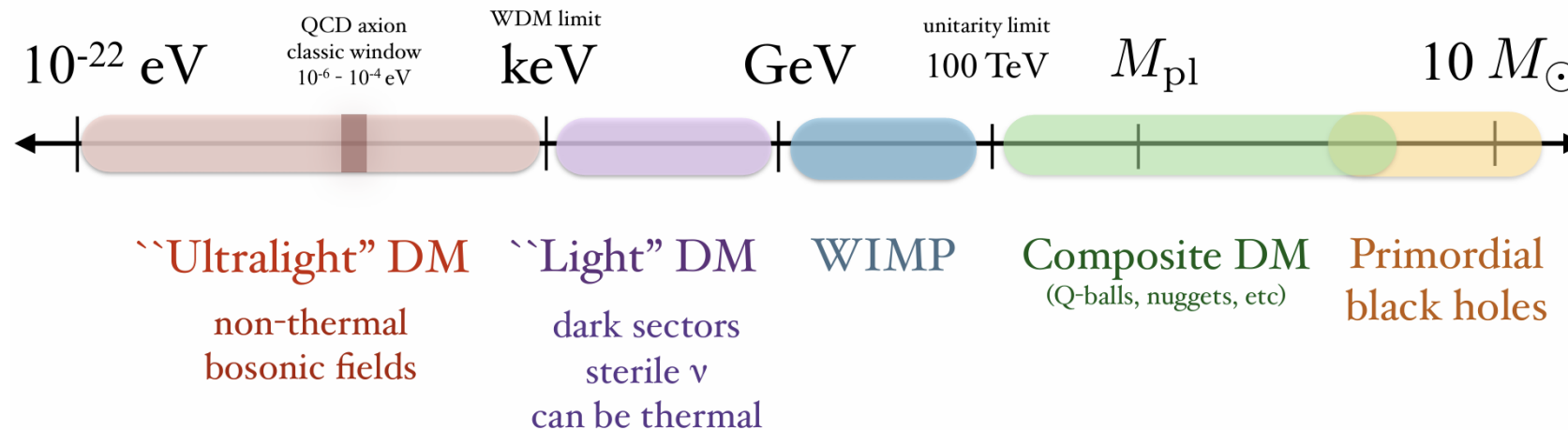
[http://www.illustris-project.org/movies/illustris\\_movie\\_rot\\_sub\\_frame.mp4](http://www.illustris-project.org/movies/illustris_movie_rot_sub_frame.mp4)



# Dark Matter

## Mass scale of dark matter

(not to scale)



There are many dark matter candidates: fun for detector designing, difficult to blind analysis

Image from [1904.07915 \(arxiv.org\)](https://arxiv.org/abs/1904.07915)

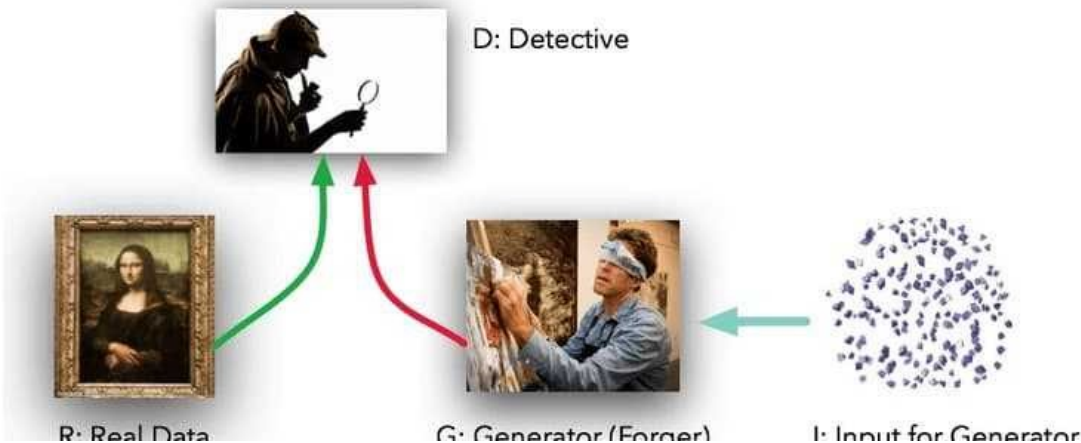
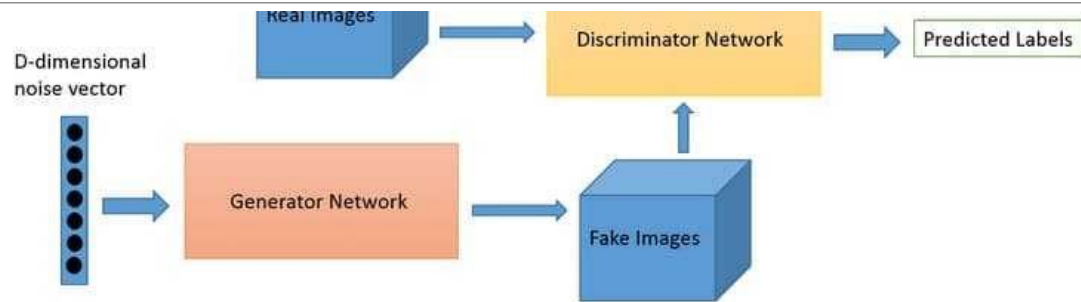


# Story 1. How can we mitigate bias in our analyses?

---

- We don't know what to expect for a dark matter signal. We do know the models rarely involve peaks :(
- Several dark matter collaborations (SuperCDMS, LUX) have used “salting” to blind their data
- See Maria-Elena Monzani's excellent talk at [PHYSTAT-2019-Monzani.pdf \(cern.ch\)](#) for additional details
- Create “fake” signal data and insert it into the data stream. You should recover the injected signal when you analyze. This is a way to build trust in the analysis.
- This is great! We can create salt for all our signal models!
- **The problem** is that creating salt takes a long time

# Story 1. How can we mitigate bias in our analyses?



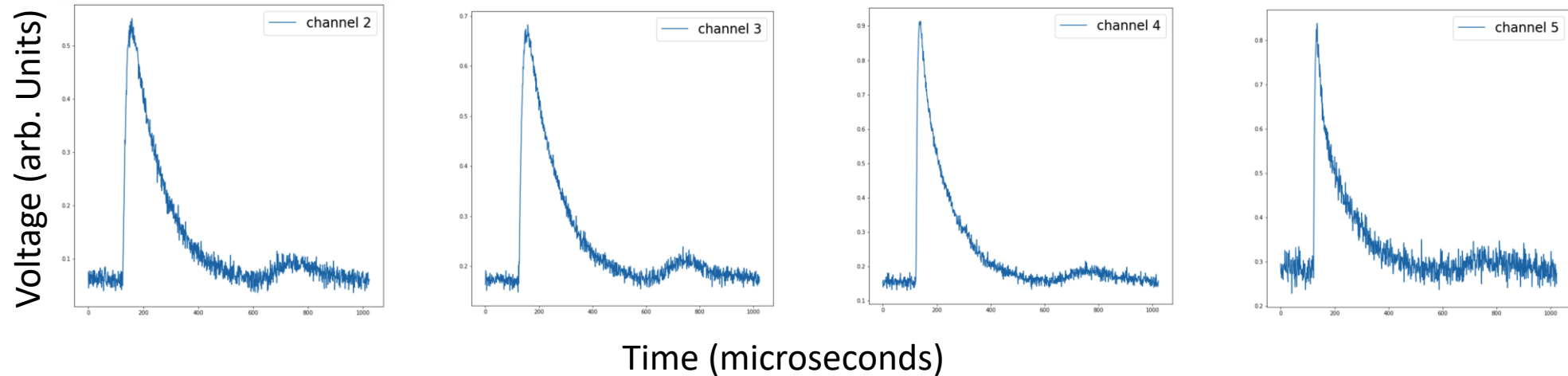
Generative Adversarial Networks (GANs) are specifically built to create fake data.

We're working on using GANs to create raw data that's useful as salt.

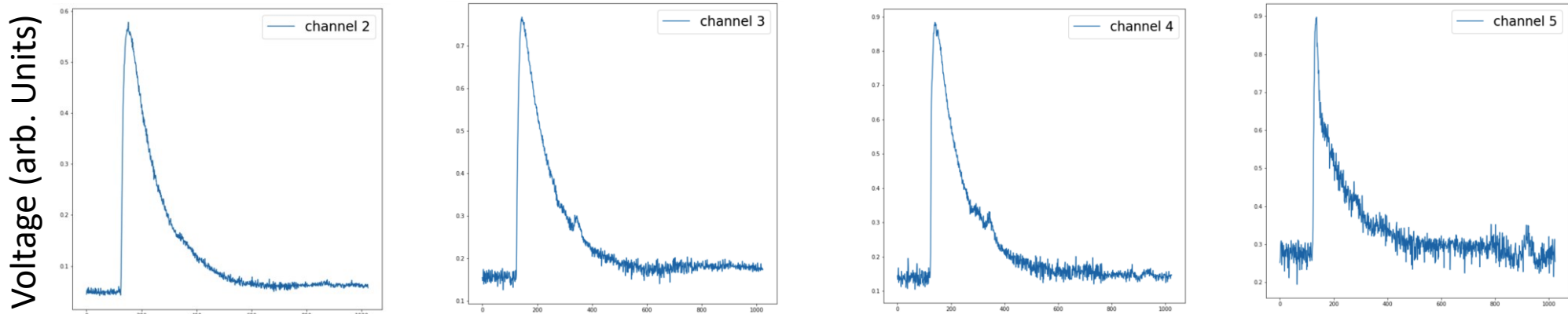
From [Decrypt Generative Adversarial Networks \(GAN\) | AI Summer \(theaisummer.com\)](https://theaisummer.com/gan/), with elements from [Generative Adversarial Networks for beginners – O'Reilly \(oreilly.com\)](https://oreilly.com/catalog/errata/csp/generative/)

# Story 1. How can we mitigate bias in our analyses?

Real Data

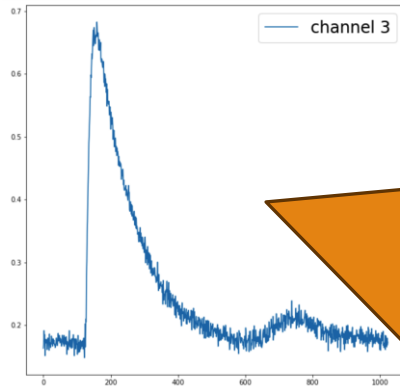
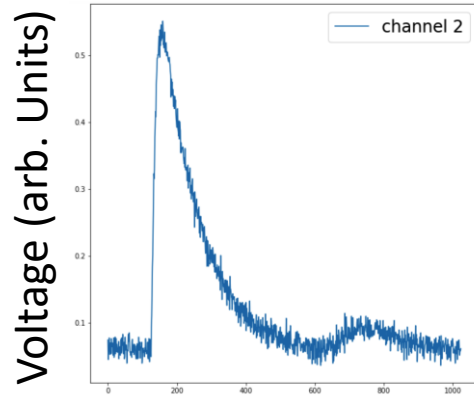


Fake Data

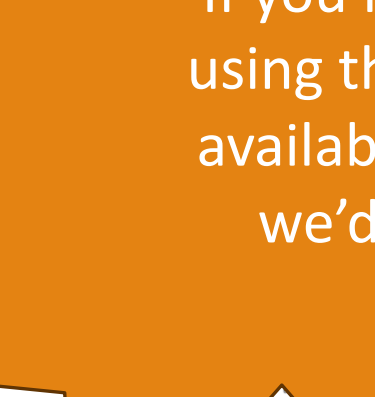
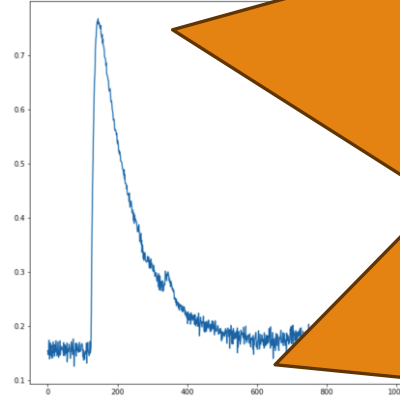
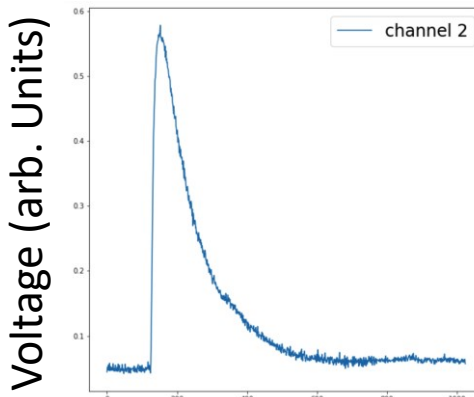


# Story 1. How can we mitigate bias in our analyses?

Real Data



Fake Data



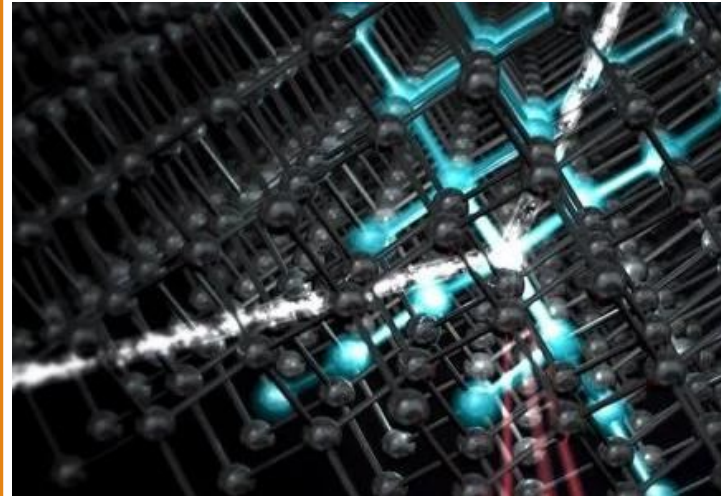
This work uses [TimeGAN](#) extensively and we're currently testing the fake data as salt for SuperCDMS.

If you're interested in using this library (soon available through pip) we'd love to talk!

# Story 2. What are our detectors telling us?

SUBTEXT: DATA AND ANALYSIS PRESERVATION CAN HAVE SIGNIFICANT SCIENTIFIC VALUE.

WE ARE USING DATA MORE THAN 10 YEARS OLD TO UNDERSTAND GERMANIUM RESPONSE TO DEPOSITED ENERGY!



A particle deposits energy by interacting with a nucleus, and that energy turns into vibration (phonons) and charge (ionization). To reconstruct the energy, we need the expected ratio (yield). The fraction that goes into phonons itself has a variance, labeled “F” in the figure.

**Imagine we see an event at 3.5 e/h pairs. How do we interpret this?**

With increasingly high resolutions, understanding energy response is increasingly important.

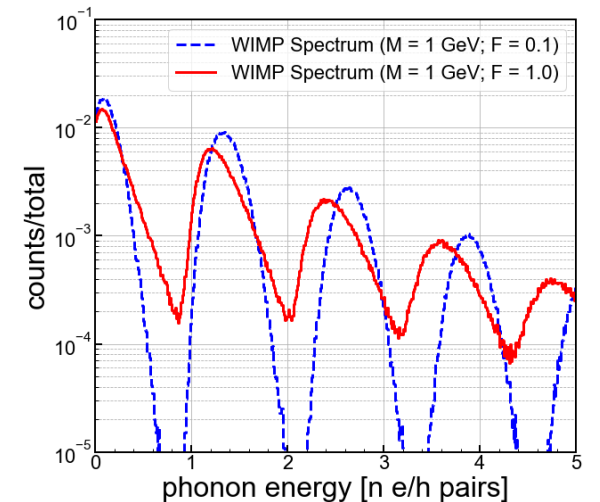
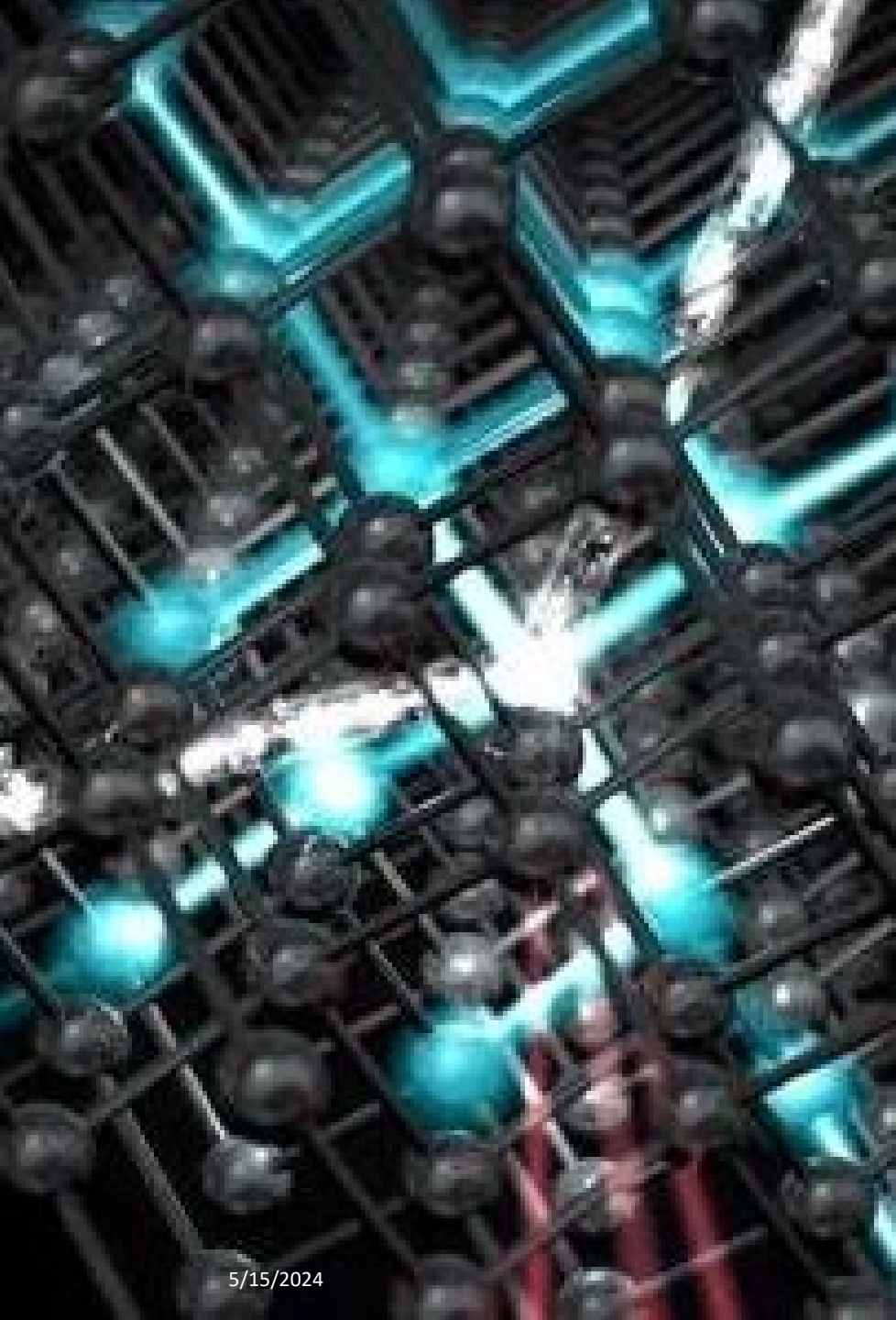


Figure from Anthony Villano





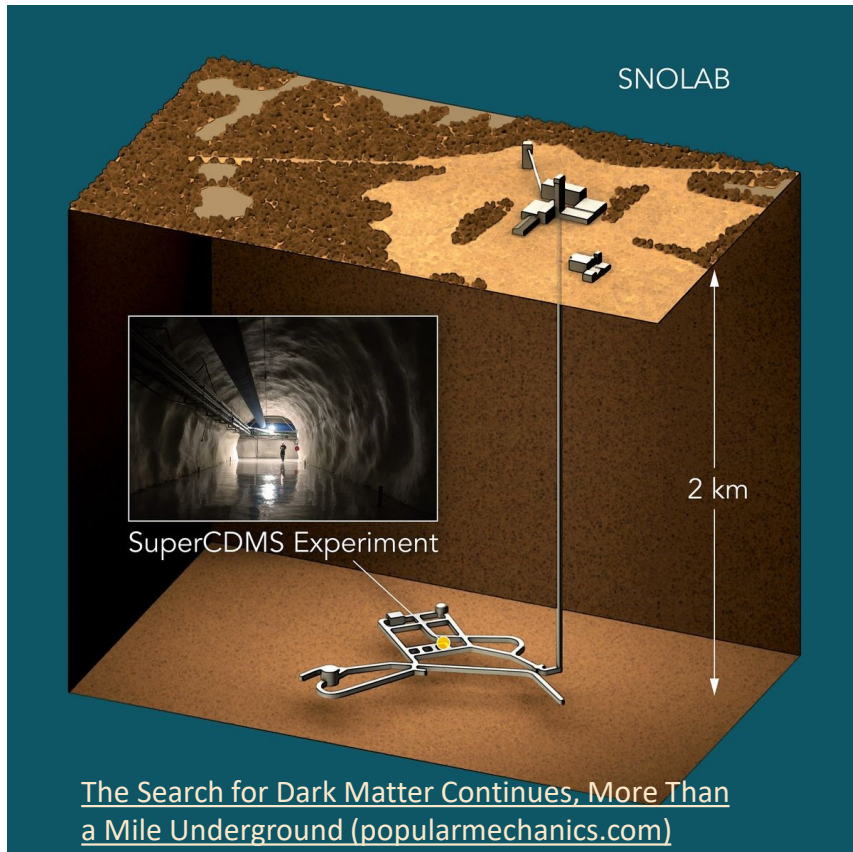
## Story 3. How can we constrain our backgrounds?

---

Some of our backgrounds are specific to our experiment.

But other backgrounds are shared by everyone at an underground facility. If I see a germanium nucleus recoil, is it a neutron or something more interesting?

# Luckily, there are many detectors near our dark matter detector: [SNOLAB experiment posters](#) | SNOLAB



Dark matter

## DEAP-3600

Detection with light

DEAP-3600 uses a vessel of liquid argon to look for dark matter by sensing the ultraviolet light produced when the argon atoms are excited by particle interactions.

The light is detected by sensors that surround the vessel and analyzed to determine what type of particle caused it.

The spherical acrylic vessel in the centre of DEAP had to be shipped to the lab in three separate pieces because it was too big to fit in the mine cage in

Dark matter

## SENSEI

Detection with CCDs

SENSEI is sensitive to low-mass dark matter using CCD (charged couple device) technology to search for rare particle events.

When a particle interacts with the CCDs, there is a small energy change which is captured by its millions of pixels. SENSEI can count every electron within a pixel which leads to accurate measurements and no background noise.

CCDs are commonly found in digital cameras, but the ones in

Dark matter

## NEWS-G

Detection with gas

NEWS-G uses a spherical copper vessel filled with a noble gas to search for dark matter. When a particle enters the sphere it ionizes some of the gas, creating electrons.

The sensor in the middle of the sphere is kept at a high voltage to attract the electrons, which creates a charge that can be measured.

The pieces of the detector made in Europe had to be shipped to SNOLAB by sea because a freight

Dark matter

## PICO

Detection with bubbles

PICO uses a bubble chamber filled with fluid which is superheated to look for dark matter. When a particle interacts, the fluid boils and creates a bubble.

Bubbles are captured with cameras and microphones, and studying them can tell scientists about which particle caused it.

The fluid is a refrigerant so even

Neutrino

## SNO+

Detection with light

SNO+ uses a liquid scintillator to detect neutrinos. When a neutrino hits the detector, it creates charged particles that cause the scintillator to give off light which is detected by thousands of sensors surrounding the vessel.

This detector will be sensitive to neutrinoless double beta decay, anti-neutrinos, solar neutrinos, and supernova neutrinos.

Neutrino

## HALO

Detection with lead

HALO is a dedicated supernova detector that uses lead blocks and helium detectors which record interactions created when neutrinos hit the lead and produce neutrons.

Part of SNEWS (the supernova early-warning system), HALO and other detectors around the world alert astronomers to supernovae so they can view them with telescopes.

Combining information from these detectors can help “veto” dark matter – if they see something when we see something, it’s not dark matter. If we’re up for a tough project, combining data could provide additional constraints on shared environmental backgrounds.

17 Institutions

# NSDF-SLAC

Stanford Linear Accelerator Center (**SLAC**)  
Super Cryogenic Dark Matter Search (**SuperCDMS**)  
Sudbury Neutrino Observatory Lab (**SNOLAB**)  
National Science Data Fabric (**NSDF**)

04 MAR 2024



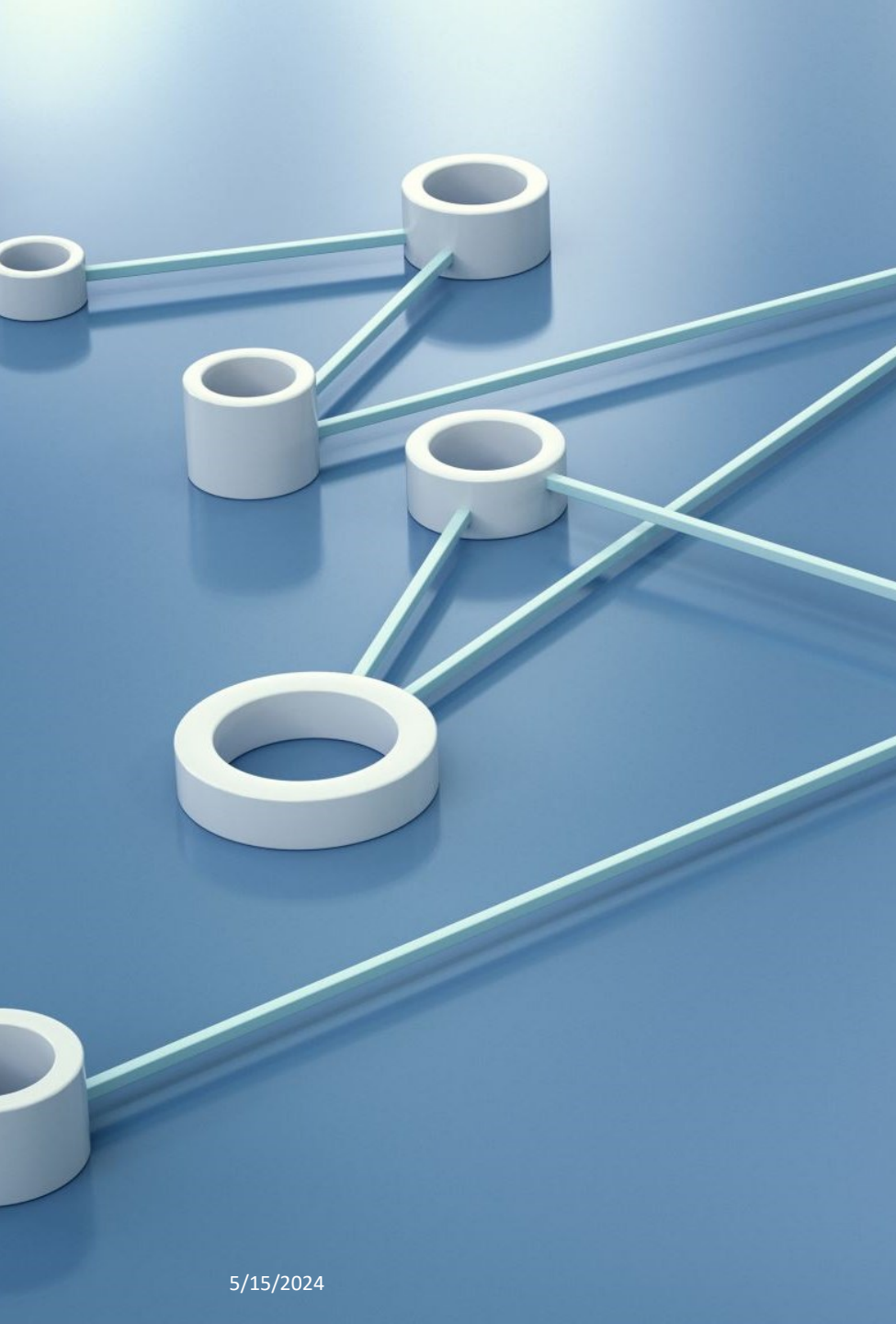
: 1.2 billion years

# Describing data formats to get a common interface

The screenshot displays a software interface for data analysis. On the left is a file explorer showing a directory structure with folders like 'archive', 'common', and 'database', and files like 'sample1.iso' and 'sample2.wim'. The main area is divided into several panels:

- Code Editor:** Shows a schema definition for a 'midas\_header' structure. The schema includes fields like 'seq', 'id', 'type', and 'value' for various data types (u2, u4, str).
- Object Tree:** Displays a hierarchical tree of the loaded data, showing nested objects like 'triggerBlk', 'eventHdr', 'numPrimitiveHdr', 'primitiveBlk', 'nDetectors', 'detectorBlk', 'detectorHdr', 'readoutHdr', 'numChannelsHdr', and 'channelBlk'.
- Hex Editor:** Shows a hex dump of the data with corresponding ASCII characters on the right. The data appears to be XML content.
- Converter Panel:** A panel with an 'info panel' and a 'converter' section. The 'converter' section has a table with columns for 'Type', 'Value (unsigned)', and '(signed)'. The table lists various data types like 'i8', 'i16le', 'i32le', 'i64le', 'float', 'double', 'unixts', 'ascii', 'utf8', 'utf16le', and 'utf16be'.

The PONDD project is a collaborative project to improve access to custom binary data sets. Support for small, non-standard, binary datasets already exists with open-source projects like Kaitai and DFDL. The PONDD project aims to add similar support for GB-scale files.



# In conclusion

---

- Preserving data can lead to better science, particularly as we develop new detectors and analysis methods. Also we need to train our students!
- Sharing data across collaborations is becoming technically feasible and offers new ways to constrain backgrounds.
- Please contact me if you're interested in collaborating!