Twitter: @AstroTunnell
tunnell@rice.edu

Input from Dark Matter Community*

Prof Christopher Tunnell
Department of Physics, Astronomy, and Computer Science
Rice University

HOW Conference

March 2019

*Organizer's meant my XENON1T/XENONnT/DARWIN collaborations, but there's a bigger untold story here…

Science one slider

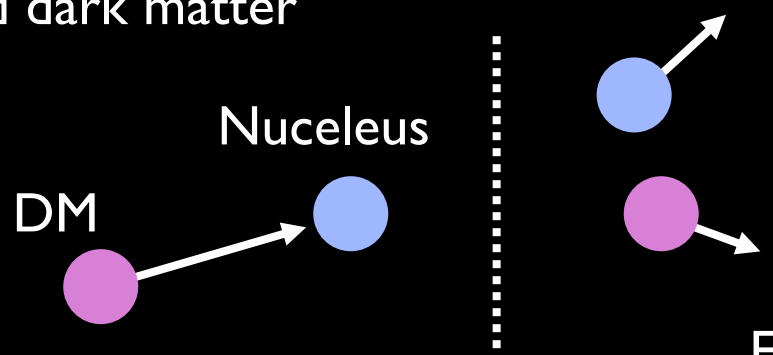
PHYSICAL REVIEW D VOLUME 31, NUMBER 12 15 JUNE 1985

Detectability of certain dark-matter candidates

Mark W. Goodman and Edward Witten
Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544
(Received 7 January 1985)

We consider the possibility that the neutral-current neutrino detector recently proposed by Drukier and Stodolsky could be used to detect some possible candidates for the dark matter in galactic halos. This may be feasible if the galactic halos are made of particles with coherent weak interactions and masses $1-10^6$ GeV; particles with spin-dependent interactions of typical weak strength and masses $1-10^2$ GeV; or strongly interacting particles of masses $1-10^{13}$ GeV.

Theory:
Elastic scattering of galactic-bound dark matter



Experiment:
Detect faint fast keV* signal in tonne-scale background-free cryogenic experiment

*keV not a typo!

Question:

Dark Matter experimentally diverse, so:

- What does each do?
- What are challenges?
- Commonalities within Dark Matter, HEP, industry?
- What 'inputs' do we have?

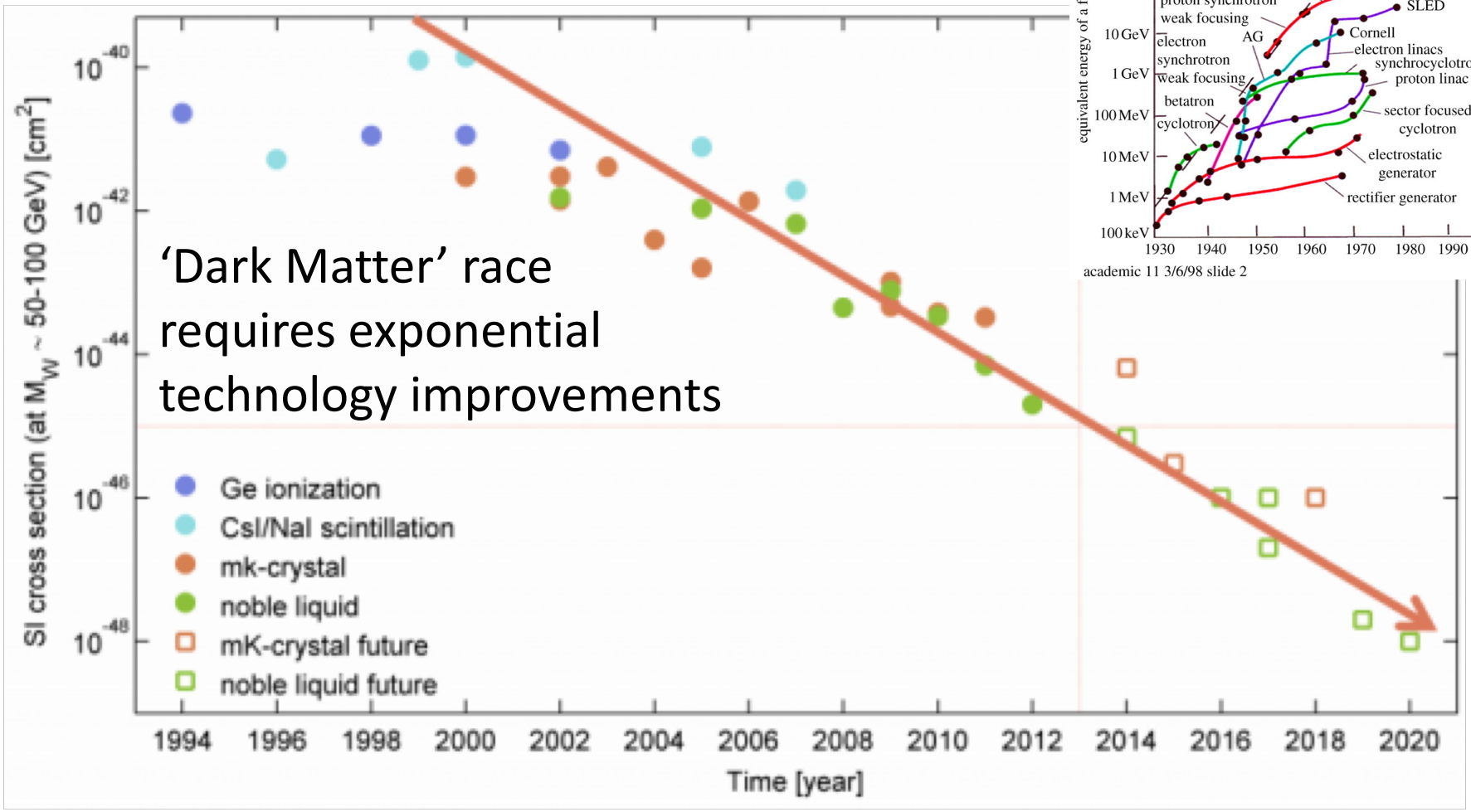
Methodology:

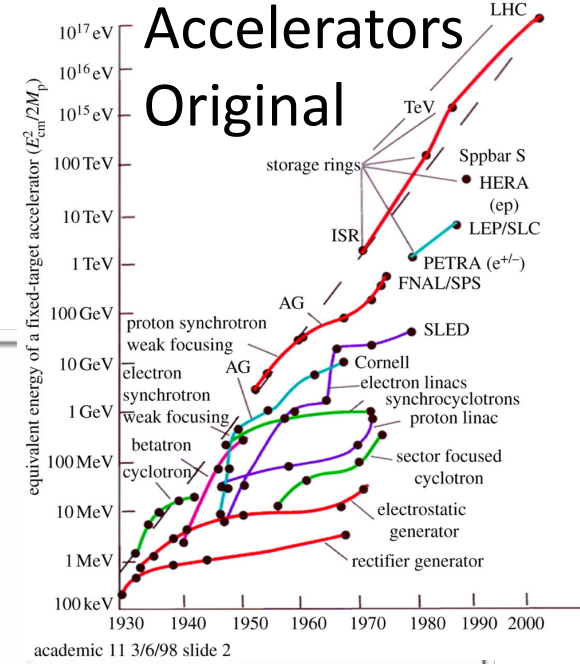
- First (?) review of field's *cyberinfrastructure*
- Approach leading experiments using similar tech*
  DarkSide, PandaX, LZ, nEXO^, XENONnT
- Determine what each experiment does and challenges
- Determine where syngeries exist (if any) to communicate here

^ nEXO is a neutrinoless double-beta decay experiment, but uses very similar technology to PandaX/LZ/XENONnT

*Subjective; SuperCDMS technically much different than noble TPCs despite pushing 1-10 GeV WIMP sensitivity..

Dark Matter's Livingston Plot



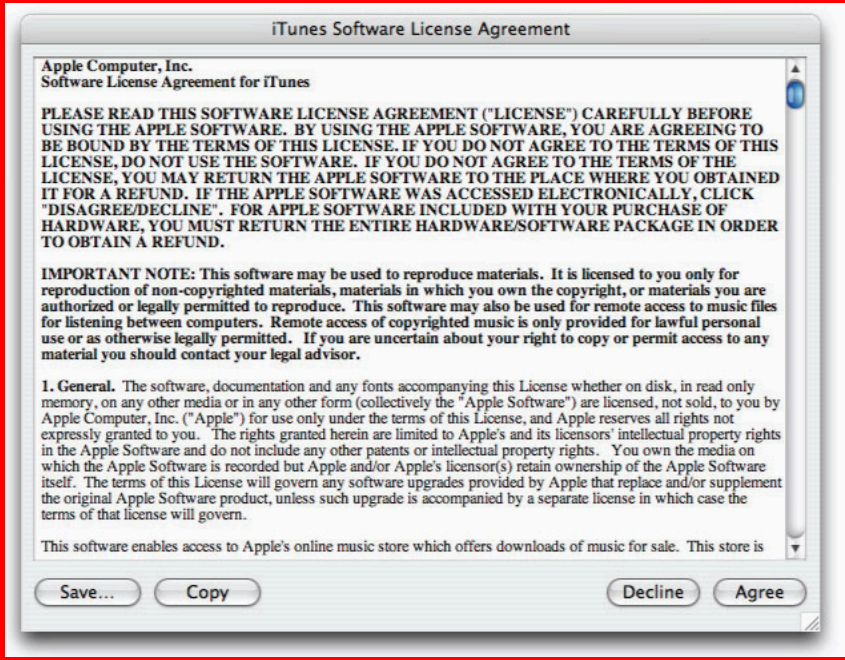


Credit Laura Baudis who updated a figure from her Dark Matter review

Terms and Conditions Disclaimer

Views are my own.

The information contained within is an incomplete survey of the field. The presentation purpose is to communicate the current state of the field, though this is subjective. The bias of the author may unintentionally be present. For official information from the collaboration, please follow the links within this talk or contact the collaborations directly. Using this talk as a reference is unwise as I have selective hearing. By going further, the audience member agrees to not make the various experiments regret talking to the speaker.



iTunes Software License Agreement

Apple Computer, Inc.
Software License Agreement for iTunes

PLEASE READ THIS SOFTWARE LICENSE AGREEMENT ("LICENSE") CAREFULLY BEFORE USING THE APPLE SOFTWARE. BY USING THE APPLE SOFTWARE, YOU ARE AGREEING TO BE BOUND BY THE TERMS OF THIS LICENSE. IF YOU DO NOT AGREE TO THE TERMS OF THIS LICENSE, DO NOT USE THE SOFTWARE. IF YOU DO NOT AGREE TO THE TERMS OF THE LICENSE, YOU MAY RETURN THE APPLE SOFTWARE TO THE PLACE WHERE YOU OBTAINED IT FOR A REFUND. IF THE APPLE SOFTWARE WAS ACCESSED ELECTRONICALLY, CLICK "DISAGREE/DECLINE". FOR APPLE SOFTWARE INCLUDED WITH YOUR PURCHASE OF HARDWARE, YOU MUST RETURN THE ENTIRE HARDWARE/SOFTWARE PACKAGE IN ORDER TO OBTAIN A REFUND.

IMPORTANT NOTE: This software may be used to reproduce materials. It is licensed to you only for reproduction of non-copyrighted materials, materials in which you own the copyright, or materials you are authorized or legally permitted to reproduce. This software may also be used for remote access to music files for listening between computers. Remote access of copyrighted music is only provided for lawful personal use or as otherwise legally permitted. If you are uncertain about your right to copy or permit access to any material you should contact your legal advisor.

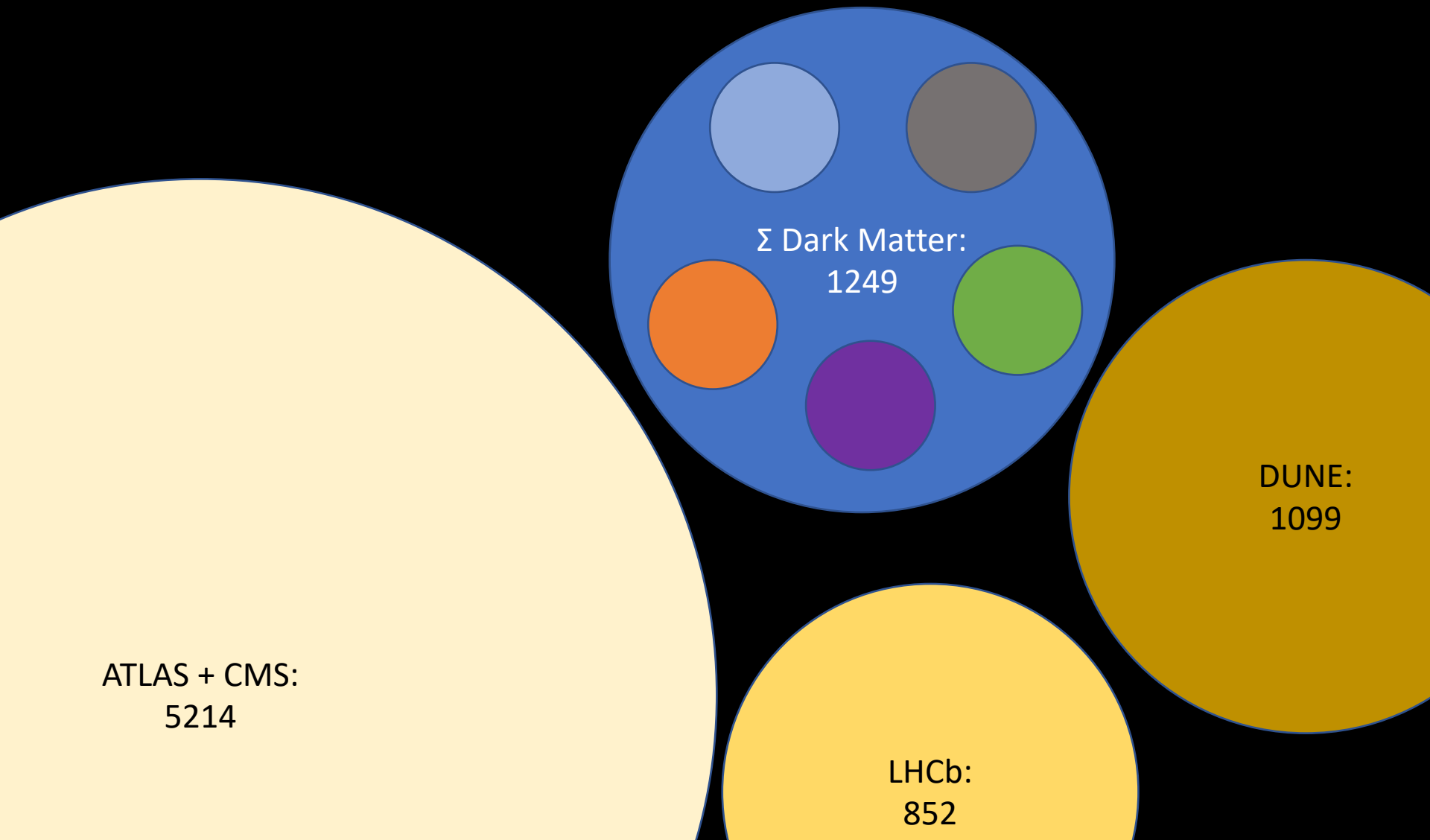
1. General. The software, documentation and any fonts accompanying this License whether on disk, in read only memory, on any other media or in any other form (collectively the "Apple Software") are licensed, not sold, to you by Apple Computer, Inc. ("Apple") for use only under the terms of this License, and Apple reserves all rights not expressly granted to you. The rights granted herein are limited to Apple's and its licensors' intellectual property rights in the Apple Software and do not include any other patents or intellectual property rights. You own the media on which the Apple Software is recorded but Apple and/or Apple's licensor(s) retain ownership of the Apple Software itself. The terms of this License will govern any software upgrades provided by Apple that replace and/or supplement the original Apple Software product, unless such upgrade is accompanied by a separate license in which case the terms of that license will govern.

This software enables access to Apple's online music store which offers downloads of music for sale. This store is

Save... Copy Decline Agree

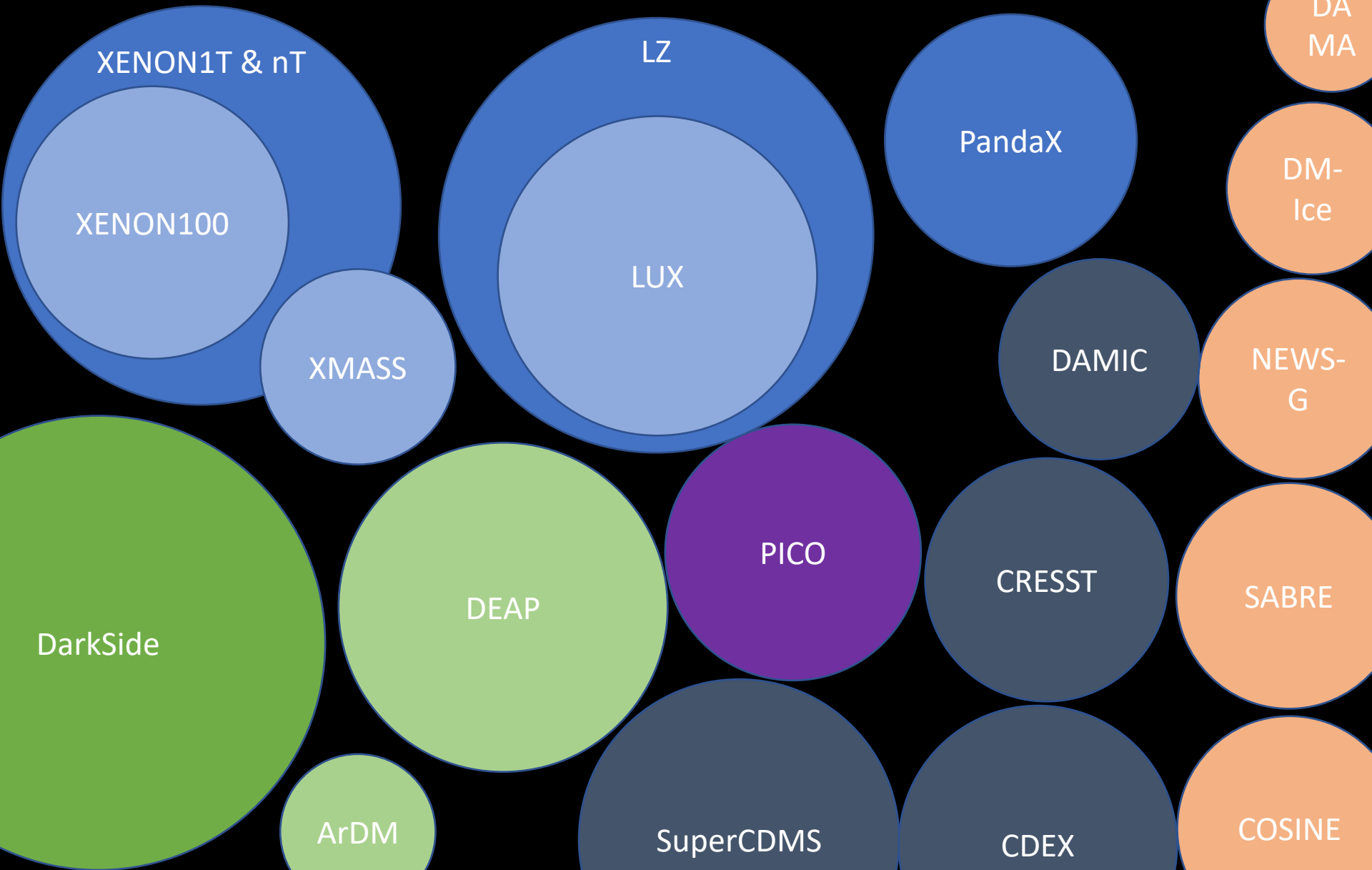
Dark Matter Detection community big

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. I almost certainly missed an experiment . Number of authors also does not correspond to FTEs since not all experiments require collaborators be 100% committed to that experiment. See gist for calculation notes. 16/March/2019



But 1249 Dark Matter users spread out

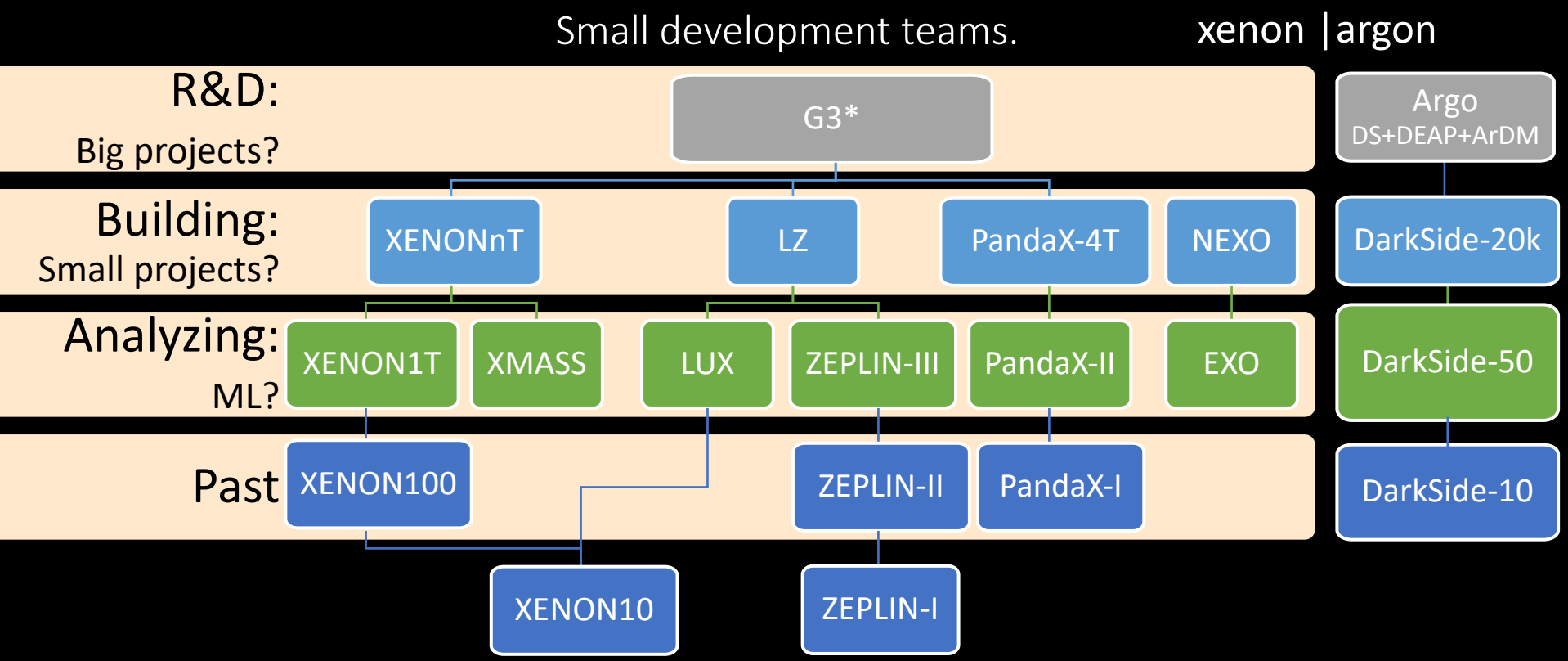
Area corresponds to number of people based on most recent publication from any experiment that has published scientific papers in the last two years. This relied on Inspire-HEP. I almost certainly missed an experiment . Number of authors also does not correspond to FTEs since not all experiments require collaborators be 100% committed to that experiment. See gist for calculation notes. 16/March/2019



DM Experimental statuses… stati?

Dark Matter experiments are HEP-like: O(PB) raw data, analysts often work on reduced data, groups of analysts often work O(year) on publications.

Many phases of experiment and opportunities for collaboration, prototyping, and support by wider particle-physics software community.



*There are concrete proposals for what "G3" would look like. This is on the European roadmap as "Darwin". Young faculty from the US XENON, LZ, nEXO groups are exploring R&D avenues. The Chinese have plans too in their nice new underground lab. The physics and R&D are clear, if the politics of it are still to converge.

Experimental details References used

Preliminary, subjective, not built.
Missing info means didn't discuss yet

PandaX
MC talk link
Data model talk link
Software Github

XENONnT
Data distribution talk link
Software Github 1T nT Ax

Non-analysis open source

DarkSide
arXiv:1707.08145

LZ
TDR arXiv:1703.09144

EXO
Machine learning following

Experimental details
Data handling

Preliminary, subjective, not built.
Missing info means didn't discuss yet

PandaX-II

- 100s TB
- JinPing->SJTU
  - Limited sites simplifies things
  - Ships AWS snowball style it sounds

XENONnT

- PBs
- Highly distributed, ~10 sites
  - Distributed difficult… computing
  - Rucio and OSG
  - Gardner UChicago group provides *substantial* support incl. B. Riedel

DarkSide

- **10s of PB**
- (See next slide)

LZ

- PBs
- SURF->NERSC->UKDC
  - Limited sites simplifies things
  - *Subjective: DOE labs must make this easier? Same story with nEXO.*

EXO

- 10s TB
- Easy and great SLAC support
  - Bigger challenge getting lots of MC from NERSC

Experimental details Software

Preliminary, subjective, not built.
Missing info means didn't discuss yet

XENONnT

- In-house Python (much faster than C++ :P)
  - agile paradigm, evangelically open source, MongoBD
  - 500 MB/s trigger, processor/analysis
  - numba e.g. strax
- Advertisement: I view our software as R&D / research in own right, is their grant $$$?

LZ

- Gaudi
  - Only software engineer fulltime this
  - @you: more framework support for small experiments?
- NERSC: Difficult small experiments adapt to multithread or new CPU/GPU architectures

PandaX

- In-house C++, low bus factor
- Live processing with JS monitoring
- MongoDB
- Subjective: seem to do neat things, though limited details.
- Seem interested in openness and collaborating (maybe through HSF?)

DarkSide

- ART and DIRAC
- Main focus is on software trigger development (argon detectors huge!)
- User management difficult for the NSF experiments (incl. XENON)
  - *Where is David Groep (Nikhef) at this meeting?*

EXO

- Sniper and great SLAC support

Experimental details
Simulations

Preliminary, subjective, not built.
Missing info means didn't discuss yet

PandaX

XENONnT

DarkSide

Everybody uses geant4… but even if this (or any other) code we make was open source, different frameworks would make it of limited copy-paste value.

LZ

EXO

Experimental details Machine learning

Preliminary, subjective, not built.
Missing info means didn't discuss yet

PandaX

- CNNs article

XENONnT

- Used but only published in masters theses

DarkSide

- Dedicated efforts ramping up

LZ

- Used extensively in LUX, starting in LZ

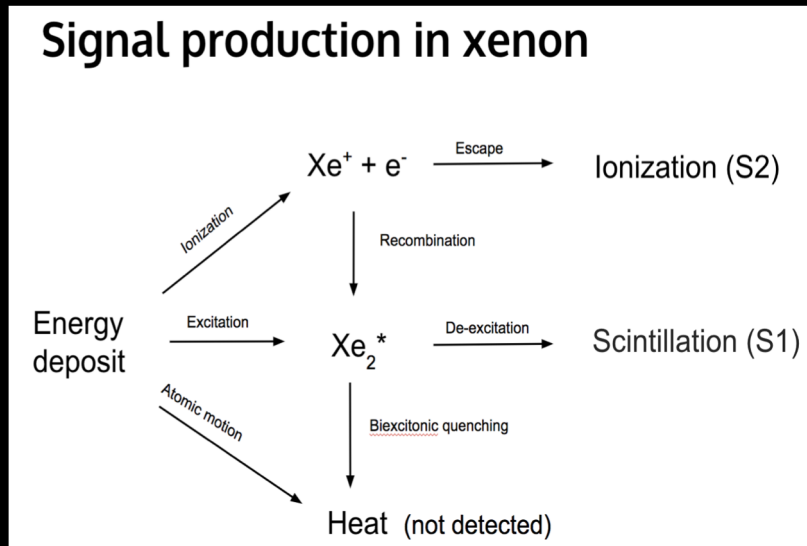
EXO

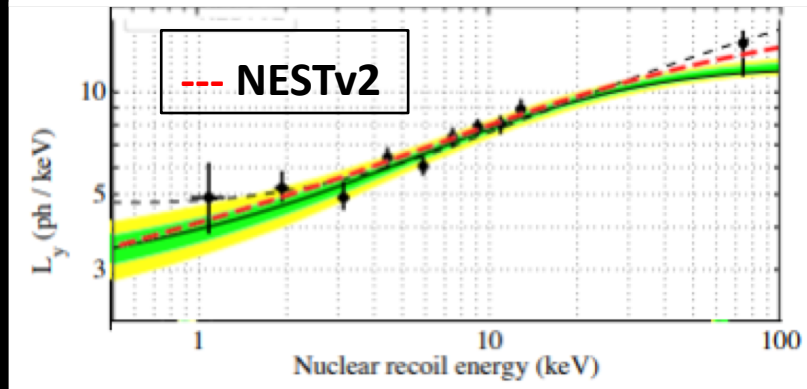
- High energy reconstruction article
- Major challenge:
  - Using Python ML codes starting from ROOT (said see uproot)

NEST example of successful intercollaboration software

- NEST simulates interactions of particles with xenon and argon atoms
  - Members from LUX, LZ, XENON1T/nT, **DUNE**, nEXO
  - Actively developed with v2.0 recently released
  - Henry Schreiner (IRIS-HEP) helped us develop nestpy
  - Vital to field in interpreting our data and designing new experiments
  - **Challenge**: for years NEST unfunded* which complicates support/usability experiments, meaning experimental FTEs often wasted reinventing wheel then making subtle microphysics mistake
- Are there other such niches? That we can get support for? So we can focus on detector-specific software?
  - Statistics? (astrophysics, form factors, flux, publish likelihoods)
  - Signal processing and templating?
  - Software triggers?
  - HEP-Python challenge with ML?
  - Simulations?
  - ?

*SciDAC fail, OAC pending





First Computational Dark Matter Direct Detection Workshop?

- Other subareas of our field have dedicated workshops
  - TPC design
  - PMTs
  - Calibration
  - Statistics
- Most students and postdocs code, but never discussed
- Limited communication on software, computing, machine learning
  - 'HEP' name discouraging?
- Broad interest in initiating a workshop
  - Communicate to improve all experiments without changing balance of power between experiments?
  - Support PhDs/PDs for cybertraining?
  - Unconference?
- Will have first organizational call in coming month with representative of each experiment.
  - Details all TBD by consensus.
  - Interested? Email.

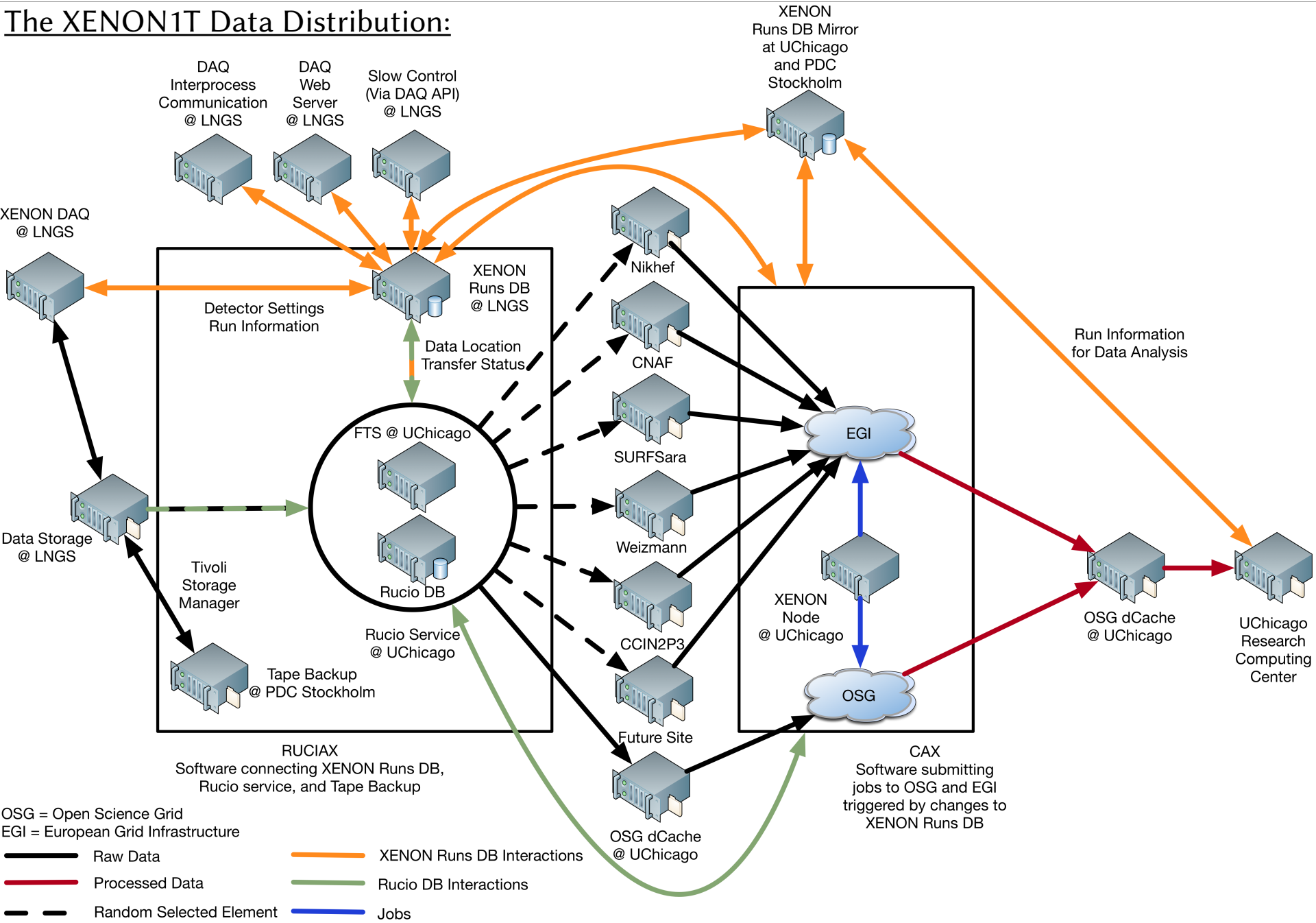
Summary:
"The Dark Matter community is substantial and could be engaged more by this community for mutual benefit."

Looking for Dark Matter job? Can connect you.

This talk was partially based on long discussions with a half dozen experiments, which only worked due to the openness, support, and significant time of many busy people ranging from spokespeople to research scientists to students. Outside XENON, thanks to those willing to speak Xun Chen, Jianglai Liu, Pengwei Xie, Maria Elena Monzani, Igor Ostrovskiy, Cristiano Galbiati, Emilija Pantic, Andrew Renshaw, Nektarios Benekos, Valerio Ippolito, Marco Rescigno. Within XENON, long chats unrelated to this talk and public talks over years with Rob Gardner, Benedikt Reidel, Boris Bauermeister, Evan Shockley, Jelle Aalbers, and many more. Any misrepresentation is my fault.

Backup slides on different experiments

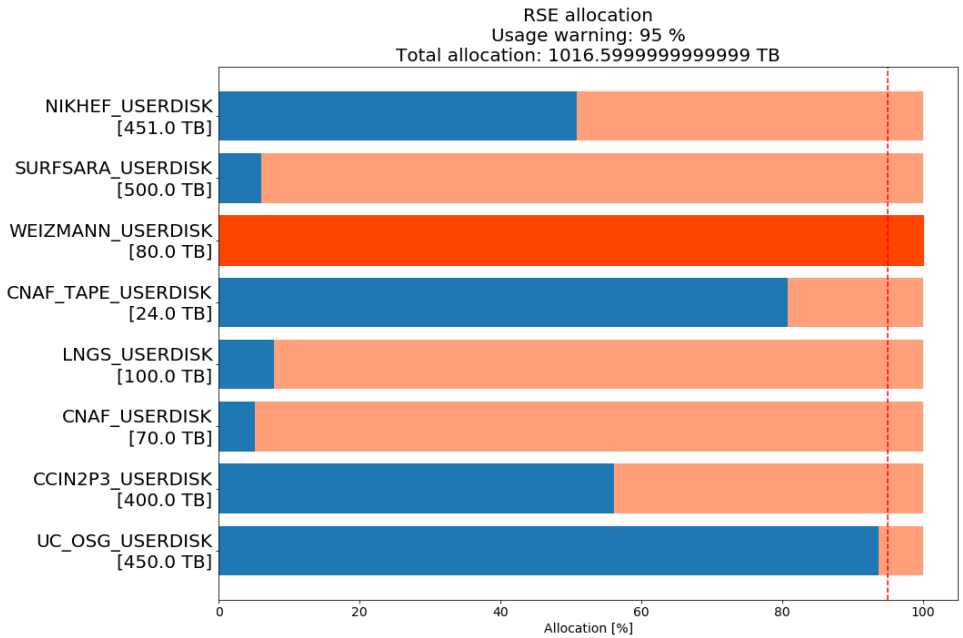
The XENON1T Data Distribution:

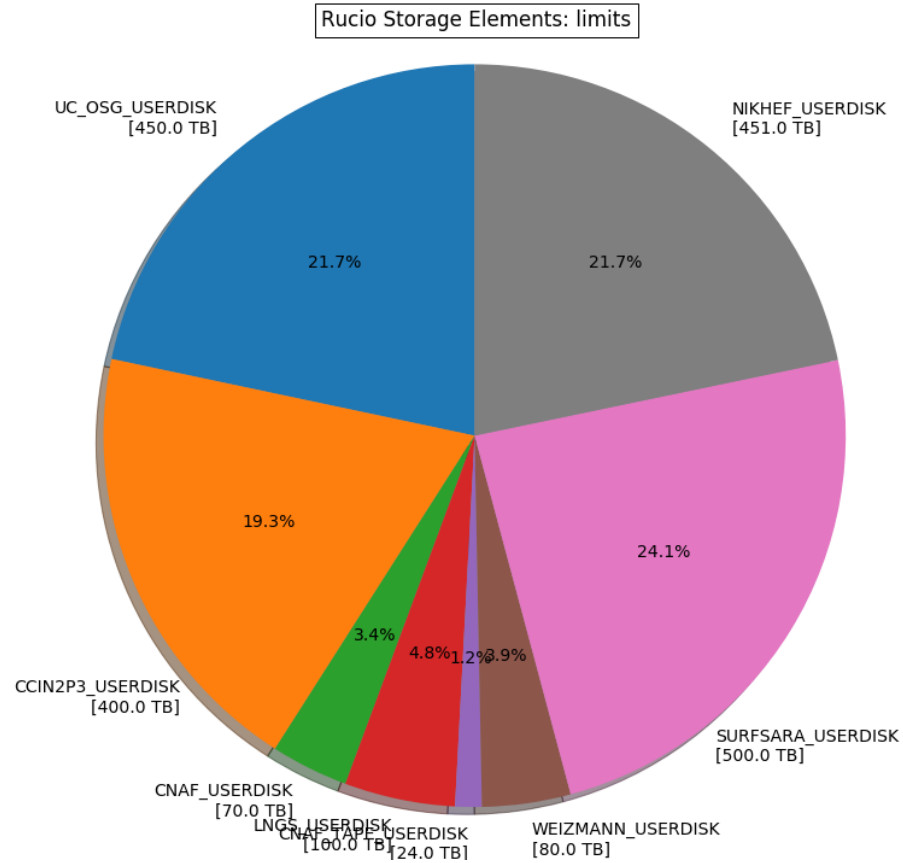


Slide from talk of Boris Bauermeister (U. Stockholm) at 1st Rucio Community Meeting

The XENON1T Disk Allocation and Requirement

- Data have two copies:
  - US: OSG dCache at UChicago (hold only relevant data)
  - Europa: One of several computing centers
- Tape copy in Stockholm Independent from Rucio

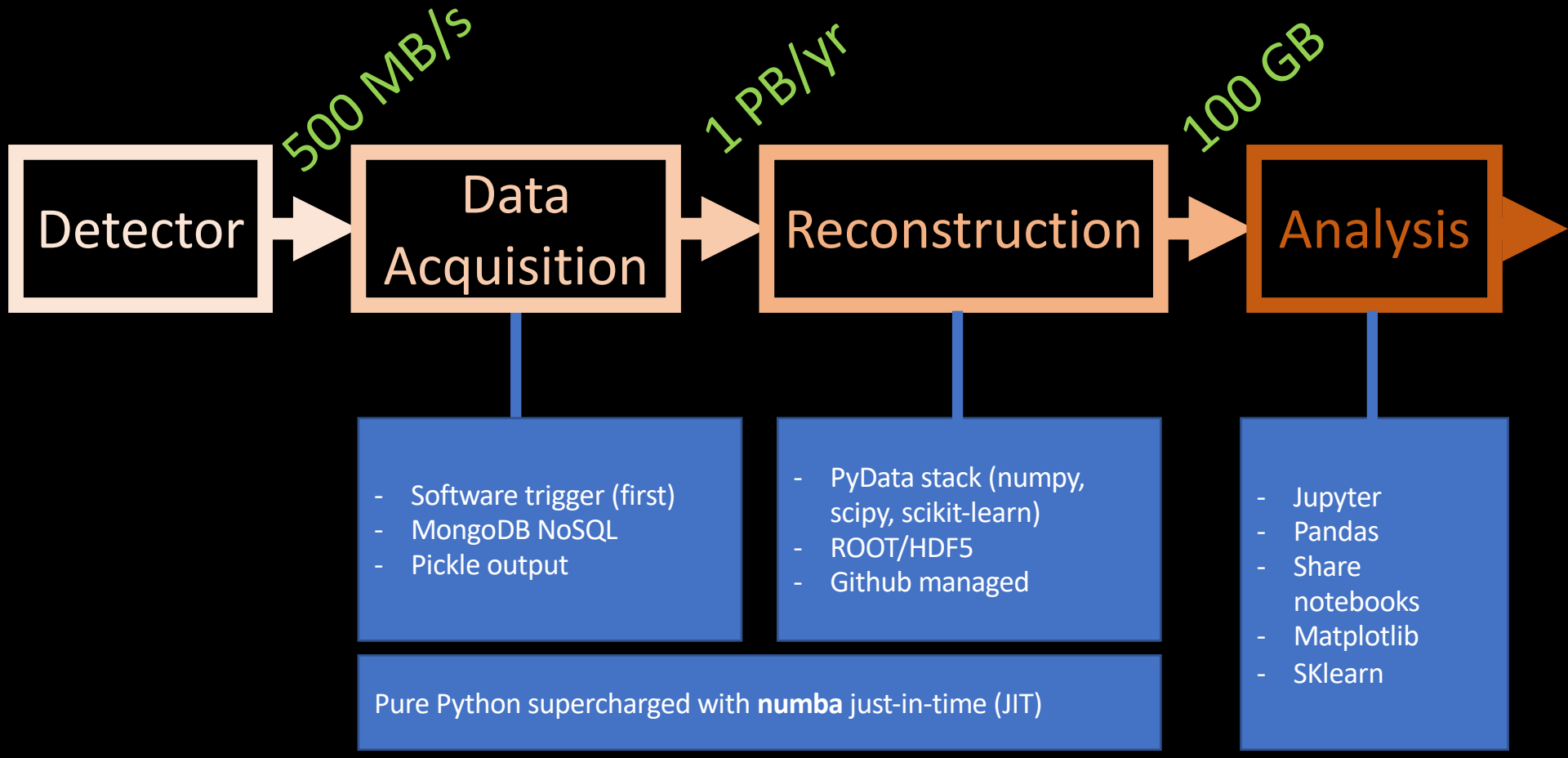




- In total: ~2 PB available
- Distributed worldwide
- Connected to computing centers

Slide from talk of Boris Bauermeister (U. Stockholm) at 1st Rucio Community Meeting

Data Flow Essentials



Numba: Python just-in-time compiler

- Few 'array-oriented' compilers though common use case and hardware optimizations exist.
- Wasn't possible few years ago, **Python faster than your C++ code.**

```python
@vectorize
def sinc(x):
    if x==0.0:
        return 1.0
    else:
        return sin(x*pi)/(pi*x)
```

```llvm
; ModuleID = 'sinc_mod_7b29370'

define double @sinc(double %x) {
Entry:
  %0 = fcmp oeq double %x, 0.000000e+00
  br i1 %0, label %BLOCK_12, label %BLOCK_16

BLOCK_12:                                         ; preds = %Entry
  ret double 1.000000e+00

BLOCK_16:                                         ; preds = %Entry
  %1 = fmul double %x, 0x400921FB54442D18
  %2 = call double @llvm.sin.f64(double %1)
  %3 = fmul double %x, 0x400921FB54442D18
  %4 = fdiv double %2, %3
  ret double %4

BLOCK_47:                                         ; No predecessors!
  ret double 0.000000e+00
}

declare double @llvm.sin.f64(double) nounwind readonly
```

Use in XENONnT: give it a rethink

Throughput in uncompressed raw data / core

Pax: 0.3 MB

XENON1T eventbuilder: 3 MB

Strax: 100 MB

Aalbers