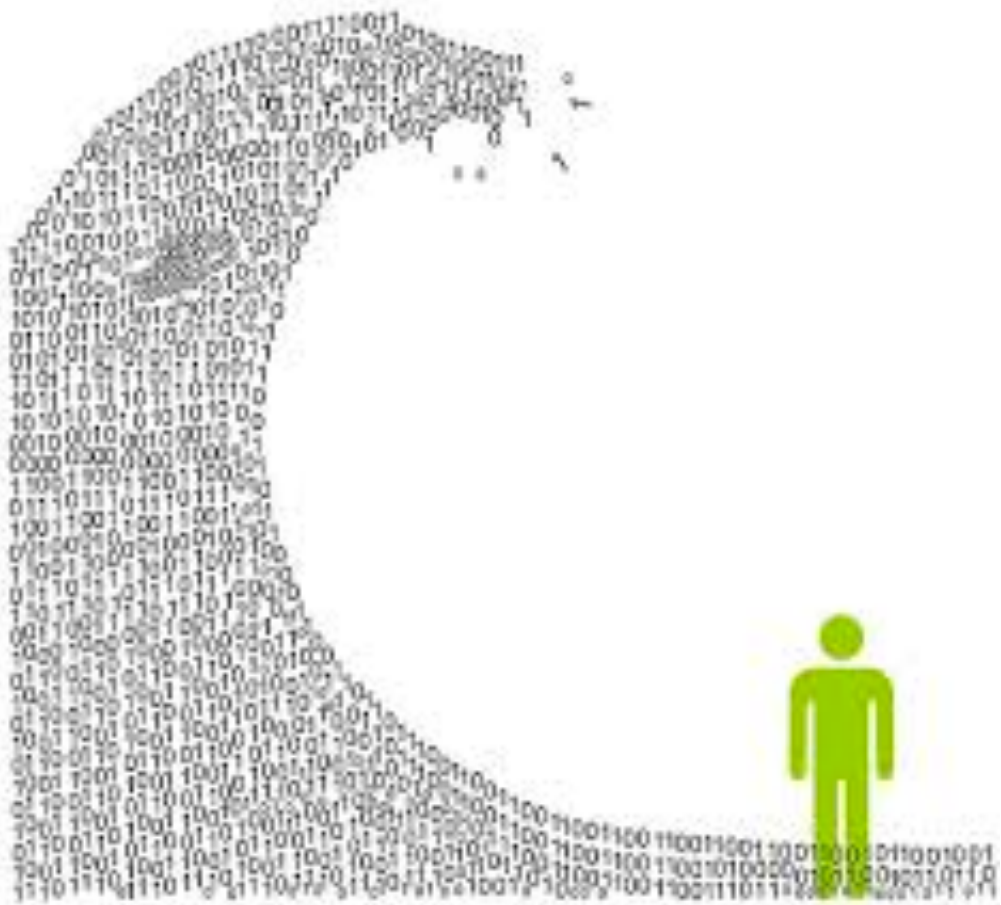
Feedback on software parallel sessions

Common themes I

- Facilities all move to live data processing
- Event mode is used everywhere
- Data volumes are ‘inconvenient’
- Software is being developed using modern methods



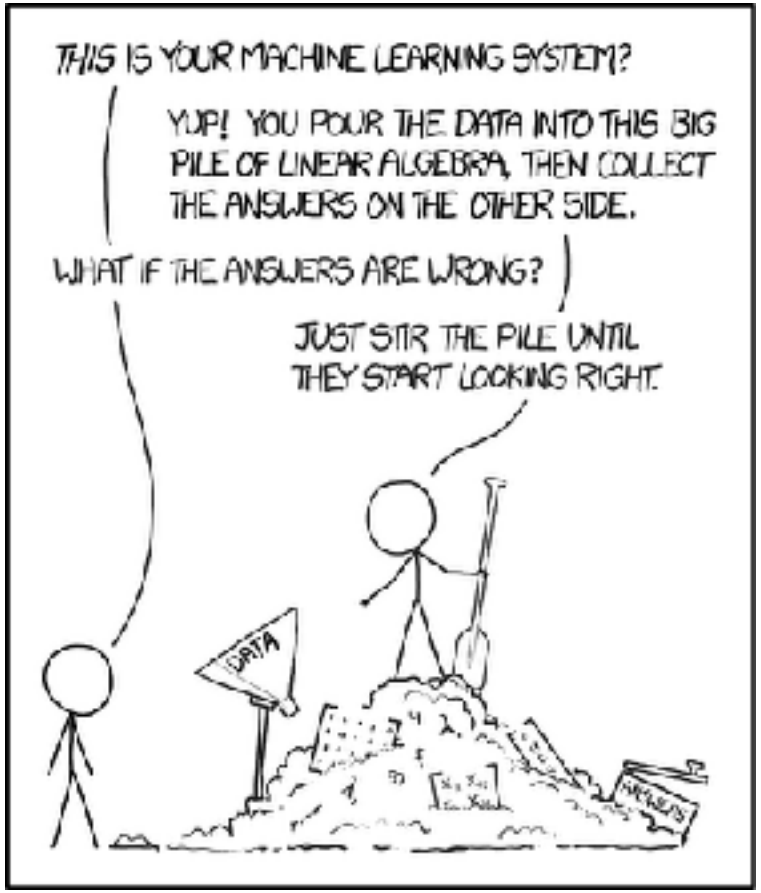
Common themes II

- We are all concerned about key software becoming unmaintained / unsupported.

- There are good examples - Mantid
- There are not so good examples - Powder diffraction

Common Themes III

- New technologies & methods are being adopted across all areas.
  - Additive manufacturing
  - MC simulation & Transport codes used for more that design & safety
  - Machine learning / Deep learning methods
    - Complex and our community are not the experts
    - *Could* provide significant impact



Common themes IV

- Everyone agrees that software increases the scientific impact / efficiency
  - More is More
  - 350 papers from SASView
  - 450 citations for Mantid
- Software is resourced by facilities and in the user community
- Lots of software is developed in unsustainable ways
  - Fixed term funds
  - Evening / Weekend development by ‘the few’


thanks!

Diffraction session notes

Young – DISOLVE software and instrument scientist for NIMROD
- Soper's EPSR is quite old and Soper has retired
- DISOLVE is designed to analyze liquids data with modern computing hardware and larger atomic models

Bilheux – jupyter and imaging
- Notebooks are generated for single tasks in the imaging workflow
- Software is written to address diverse user abilities/expertise
- Users are contacted shortly after proposal is accepted to start process of customizing software and verifying that appropriate software is available before their arrival at the facility

Peterson – event data processing
- The math for event data processing exists
- Can use clustering algorithms to determine what diffraction measurements are similar without models of the atomic structure

Islam – 3d printed collimators
- Collimators can reduce unwanted scattering from sample environment/pressure cells
- Using monte-carlo simulations (mcvine/mcstas) to minimize <num peaks from background>/<num peaks from sample>

Discussion at the end
- Why is jupyter at ORNL only for imaging?
  - Users from other fields like more involvement
  - Imaging has a lot of industrial users
- Single point of failure in diffraction analysis software (GSAS/FullProf)
- ISIS total scattering is still using GUDRUN rather than mantid

LSS

large scale structures topics

reduced data - how much can you run analysis on reduced data, what do you need the full raw data for
1d reduced data is trivial to fit
2d gets harder, GPU Helpful
higher dimensions need clusters

PyQT vs Jupyter for UI

License for commercial use

how to make projects sustainable: Make them developed at at least two facilities
by either spreading developers or in kind development

How to survive topical short term funding

data management

- Though Russell Ewings Focused on IDAaas he had a nice summary of the Bottlenecks that is more far reaching
  - Inconvenient Data
  - Too large for users resources
  - Reduction/analysis is highly interactive
  - Specialized domain specific tools are needed
  - Data should be processed quickly

-Another thing they learned from IDAaas is that there is different expectations of service level from the users (24/7) and the Cloud support staff (more 9-5)
-Username and password should be sensible and easy for users.
-Specific user requests need to be aligned to facility goals in a strategic manner.

-There was a discussion on Fair data usage in regards to EOSC

Other Bottlenecks for production of results
-Instrument hardware is often updated without an eye towards the required computing and data infrastructure upgrades.
-Reuse of software components across different software packages could expedite development and ease support issues.
-UX vs. resource issues. We agree better user experiences for users will expedite the production of results. UX work is often constrained by personel resource constraints. At the very least tasks that compete with industry standard solutions (ie making a search engine to replicate google functionality) should be avoided.

MonteCarlo Methods

Mads Bertelsen on behalf of Peter Willendrup: McStas continues to be strongly used. The next release (version 2.6) will come out approximately the same time as a beta version of 3.0 which is less backward-compatible, but with substantial improvements in the translation into C. It is also needed for future GPU use.

Milan Klausz: BIFROST was simulated using both McStas and Geant4, using the ncrystal component in both, with good agreement for the scattered neutrons, but different results for the transmitted beam.

Kyle Grammer: MCNP can now do TOF with choppers, including rotating components which incorporate scattering or even moderating materials.

Miguel Magan: MCNP6 can now transport slow neutrons in guides with very realistic supermirror models, even including the wavelength-dependence of the gamma emission.

Thomas Miller: Some important tweaks to the ADVANTG tool within MCNP6 is greatly increasing the calculation speed for shielding calculations, particularly when out of line-of-sight.

There is a trend for tools originally created for HEP being adapted for calculation of slow neutrons: Geant4, MCNP
That is strengthened by ncrystal which plugs into Geant4, and the MCPL format for exchanging particle trajectories between packages.
The result is that we have many more tools for simulating instrument performance. That is clearly excellent for allowing cross-package checks and improving the reliability of the results. However, it does raise the question of whether we need a neutron scattering community forum for coordinating all these different packages which are increasingly overlapping in capability, given the finite resources we have.

data processing

Mantid:
- The approach and tools have had to change as the project has grown and more contributors have joined internationally.
- Dependencies can be a major headache - carefully vet what you are importing.
- GUI testing is challenging - using MVC approach in new GUI means that testability improves.

Scipp:
- 0.1 release out : try it!

JParc Live Data:
- Well designed web interface to all servers.
- Noted separate data flow for data to disk vs live data processing - discussion on data integrity that followed.
- Started work on experimental feedback from live processing.

Discussion:
- Focus on performance is key. Users want live processing, but in fact that is in some ways easier than delivering fast processing from disk when long runs are written in event mode.
- Challenges of data integrity/correctness with live processing. Ensuring that the live result is of sufficient quality for the users to take it away and use it, since they will do anyway. Provenance will be key here - processed data must show history.

Spectroscopy Session

- Machine learning applied spectroscopy finding global minima in a large parameter space model
- SpinW - Broadly accepted as the user code for LSWT fits with the existing tools and workflows
- SpinWave genie is actively developed and supported
- ICEMAN containerised storage that makes many small tools accessible for users