Water, water, everywhere, nor any drop to drink

Samuel Taylor Coleridge, The Rime of the Ancient Mariner

OR

Collective action on Analysis software for the nomadic Small Angle Scatterer

CanSAS XI workshop

Freising, Germany

July 8, 2019

Analysis software

Data Analysis eh?
…. So what exactly does that mean?

Only works on Reduced data

(All the instrumental artifacts are removed and only the science is left)

…. Sorta

What about resolution?

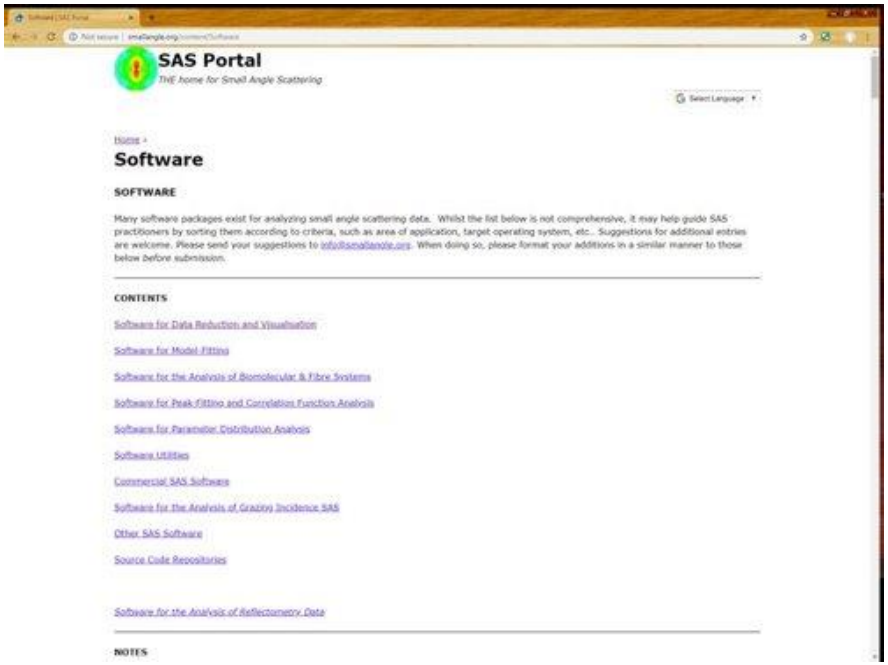
Simulate instrument and sample and always work on raw data?

Suggest leave this to other sessions?

Analysis software

Lots of existing software

http://smallangle.org/content/Software



Problems:

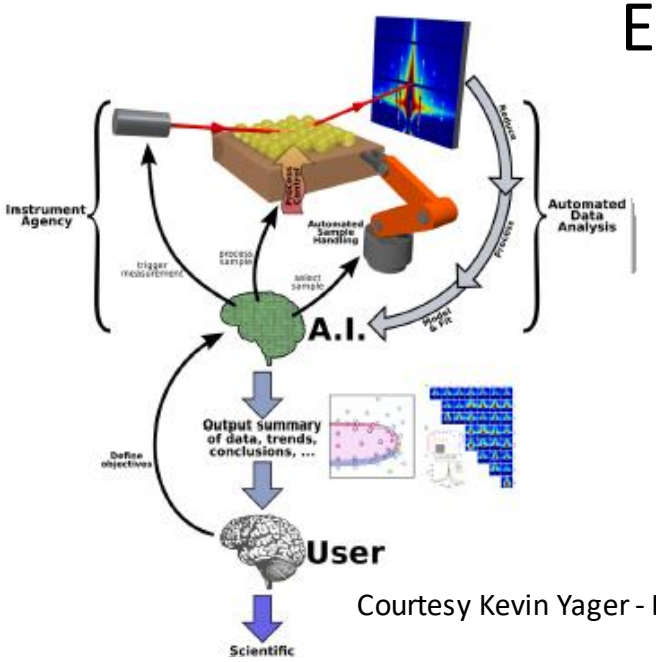
- Limited functionality
- Lack of support
- Difficult to install
- Difficult to use
- Single developer/bus factor
- Etc?

Mostly seen as mundane

….. Booooring!

= no funding to support

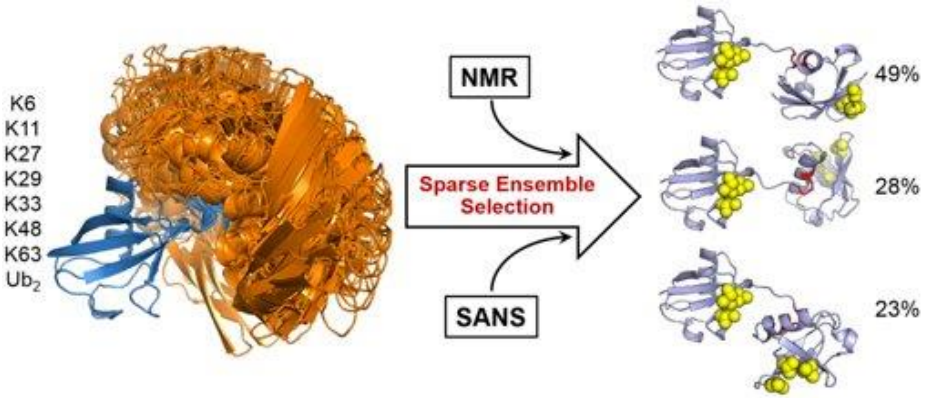
Analysis software
The bright shiny new things

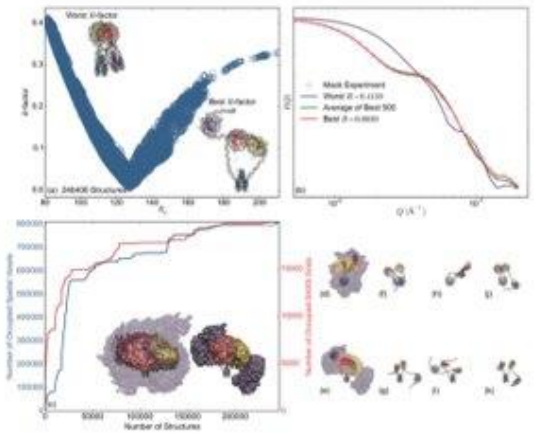


Courtesy Kevin Yager - BNL

Exciting times with lots of Opportunities

- Machine/Deep Learning
  - Soooo 1980s!!
- Molecular simulations
- Correlative Analysis...
- Others?





Courtesy Joseph Curtis - NIST

Analysis software

When the going gets tough the finger pointing starts

Who's Job is it Anyway?

Analysis is where the science is → the USER'S JOB

Scattering is an analysis tool and part of providing the tool should be the necessary software → the FACILITY'S JOB

Data on disk is useless to EVERYBODY

Why can't we all just get along together ….

We need to work together!

Demands on Analysis Tools

New realities of data driven science

- **Reviewers/Editors**
  - Demand for access to source code for inspection (transparency)
  - Traceability from data provenance to final results
  - Reproducibility of results – packaging analysis

- **Funding agencies**
  - Demonstrate long term development and support post grant funding
  - Open source and reusable code (taxpayer funded)

Scientific Software Development and The cyberinfrastructure revolution

- **Never enough resources to achieve the vision we have**
- **No resources for long term maintenance and support.**

Problem:

- To reap the benefit of investment in software developments requires foundational long term support.
- If entity that supports the development also must support the "maintenance" forever, the entity will soon cease to be able to fund new projects.

CONCLUSION: This paradigm is broken!!!

FACTS OF LIFE:

- **Resources are finite**
- **Needs are infinite**

Scientific Software Development

If paradigm is broken how do we as a community fix it?

Collaborations?

- On common areas?
- On fully joint efforts?
- On "division of labor"?

Benefits

- More achieved for total resources
- Lower the cost of long term maintenance

What is/should/could be the role of:

- The scientific community (users)
- Funding agencies
- Facilities

Some Questions for Today

Goal = Start a conversation to

- **Identify**
  - Current Analysis needs and bottlenecks?
  - What are the most promising emerging techniques?

- **Explore**
  - Possible Models for sustainable development and support
    - For current less glitzy needs
    - When the bright new shiny thing becomes less new and shiny.
  - How best to "manage" limited resources
    - Avoiding re-inventing the wheel (and NIH)?
    - Without succumbing to dangers of a monoculture
  - What Collective Action can we take to help maximize the analysis tools support for our Nomadic Small Angle Scattering community given finite resources.

Final thoughts for Today

Some Ideas

- Can we use networking grants internationally to seed better collaborations/interactions?
- Can we define and propose a development and sustainability model that defines the roles that we the users, facilities, industry, and granting agencies can/should play

In the end:

- We can either be passive and accept whatever comes our way or
- Be proactive in trying to influence how the limited resources get used to our benefit.

The latter usually gets you more of what you want...