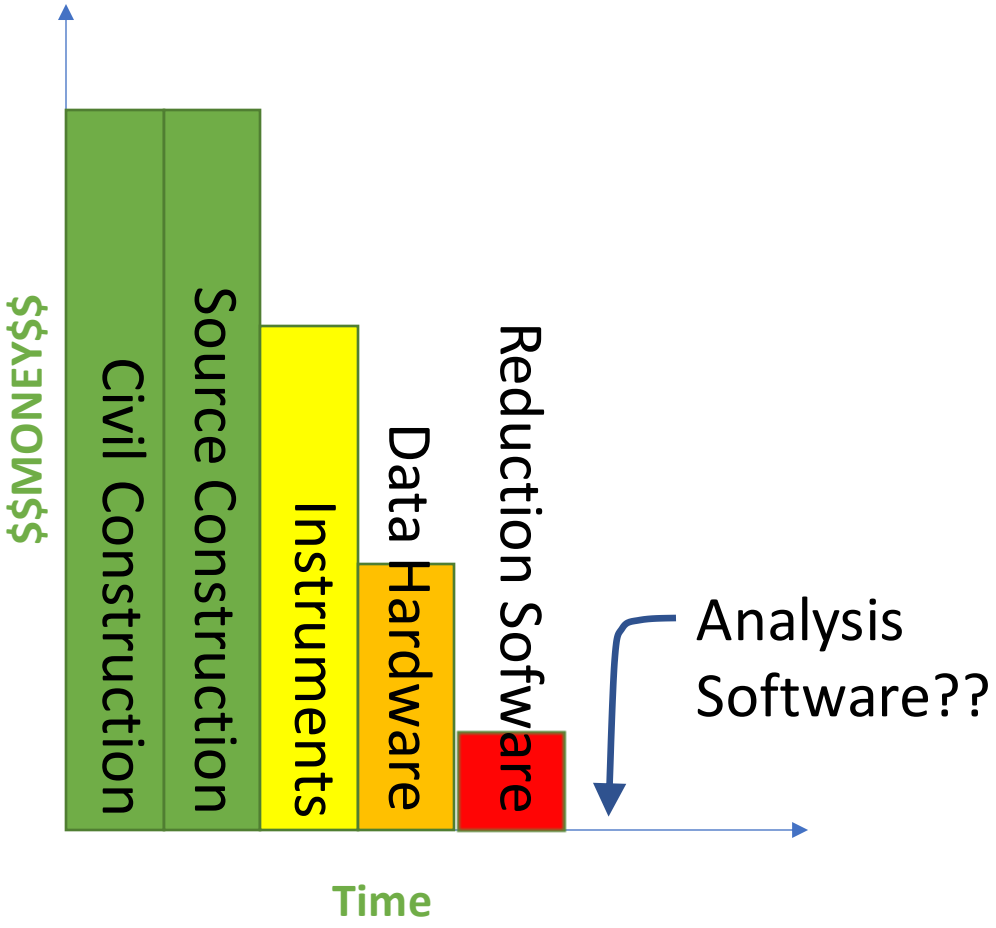
Addressing the Resource Problem Through Community

SasView

UT ISIS NIST diamond OAK RIDGE National Laboratory ansto NEUTRONS FOR SCIENCE ess TU Delft

The resource Problem Part I

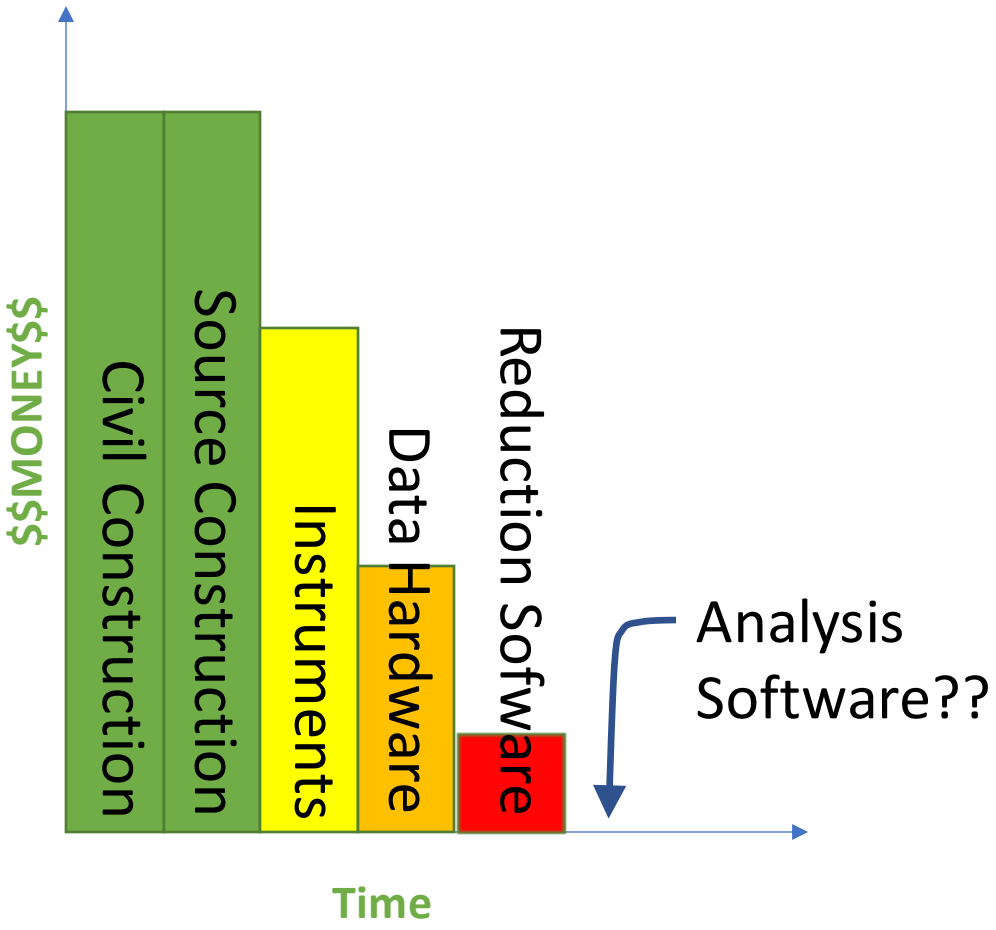


Analysis? Let's define that

"Works on data from which all instrumental artifacts have been removed"
…. sortof

→ Instrument Independent Data

The resource Problem Part I



Analysis Software - Who's Job is it Anyway?

Scattering is an analysis tool and part of providing the tool should be the necessary software

→ the FACILITY'S JOB

Analysis is where the science is

→ the USER'S JOB

Or maybe …..

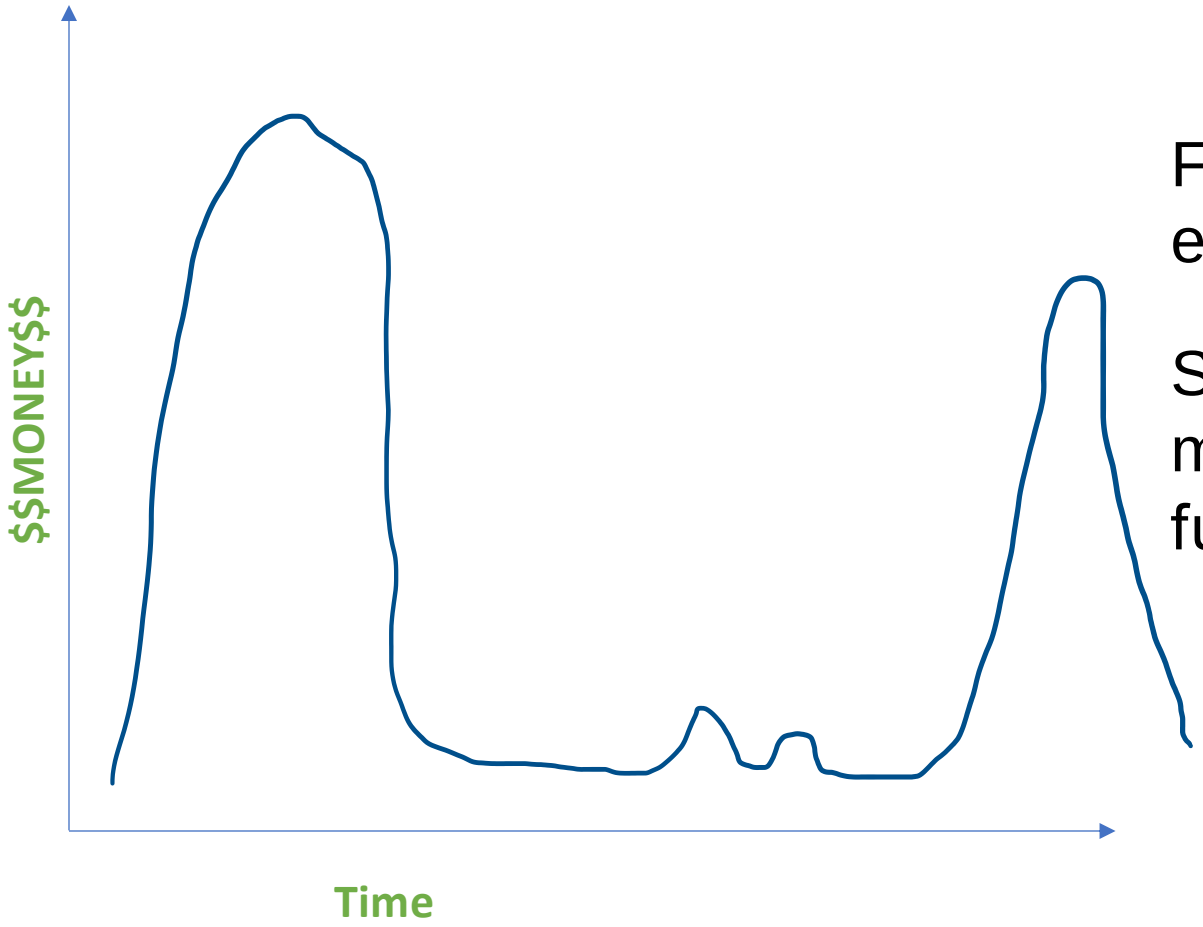
→ The Domain Science funding agencies?

The resource Problem Part II
The feast and famine roller coaster

The Valley(s) of death

Analysis Software - Who's Job is it Anyway?

- The domain science funding agencies

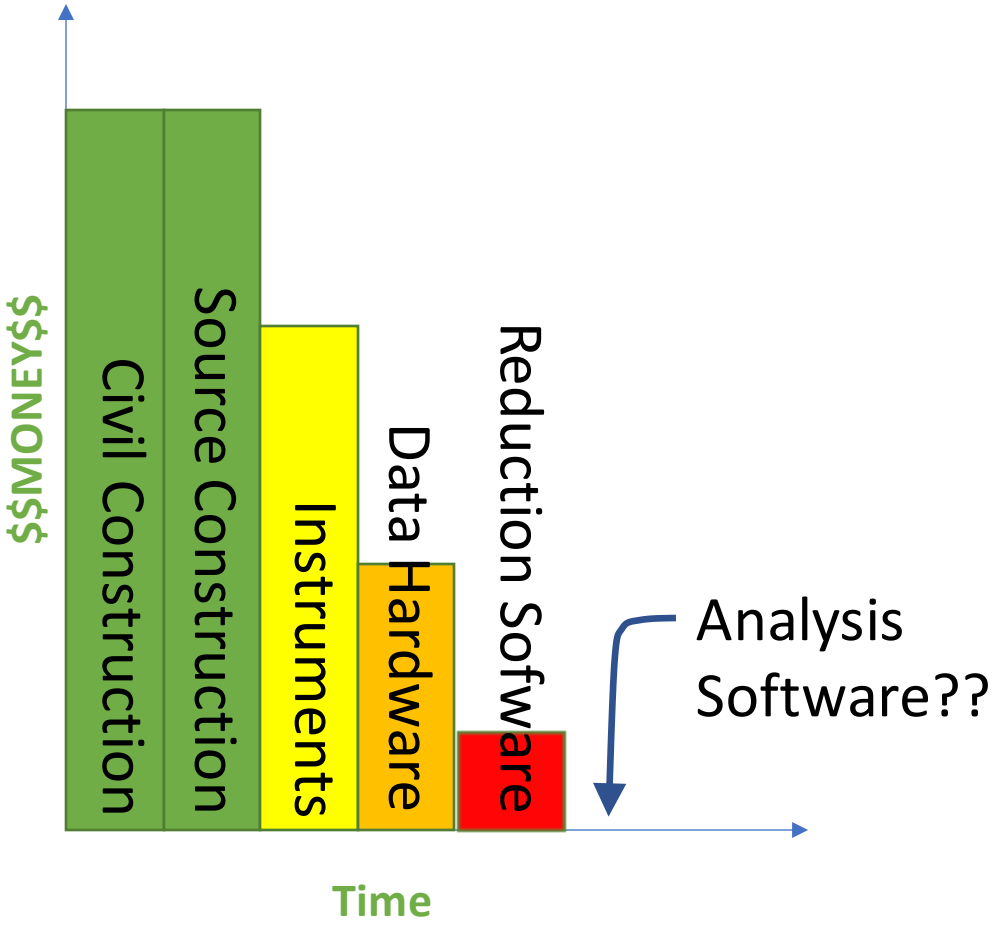


Facility directors discretion, NIH, NSF, DOE, etc.

Special funding (grants) do no fund long term maintenance and ongoing development. They fund "big new (transformative) ideas"

The resource Problem Part III

The unbounded problem



Fundamentally all these elements are relatively well defined problems…

EXCEPT ...

Analysis Software is really unbounded. The needs are nearly infinite and constantly evolving.

HOWEVER --- Analysis as defined here is also uniquely universal and ripe for collaborative pooling of finite resources

…. But beware the monoculture?

The Problem Statement

- A way to focus limited resources on top priorities (most useful to the science being done) in a world of infinite possibilities
- A way to harness funding for bold new ideas without losing the investment in the valleys of death
- A way to provide sustained maintenance and development in an uncertain funding environment (thriving through the famines)

FACTS OF LIFE:

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

So, How Can we Solve the Problem?

Step One: Collaboration.
Let's all work together for the common good

The Reality: Collaboration is REALLY HARD!!

Lesson 1: Work with me NOT for me

Serendipity: Some history

SansView was DANSE project output
~ 8.5% of funds were for SANS
+ BUMPS … see later

2006

2011

2012 NIST Supported initial transition from NSF funding

2013 Transition to Community project.

1st Code Camp at NIST April 2013

2014

2nd Code Camp at ISIS April 2014

2015 Move to GitHub

Rename to SasView

v3.0 released

v3.1 released

3rd Code Camp at ESS Feb 2015

2016

4th Code Camp at TU Delft March 2016

v4.0 released

5th Code Camp at ORNL Oct 2016

2017

v4.1 released

6th Code Camp at ILL/ESRF April 2017

7th Code Camp at DMSC October 2017

8th Code Camp at ESS Sept 2018

2018

v5.0b1 released

v4.2 released

1st SasView User Meeting at SAS2018

v4.2.1 released

2019

v5.0b2 released

9th Code Camp at ILL/ESRF March 2019

Some More Lessons Learned

- Nobody owns → fewer egos
- No proprietary platform (open source python – eventually some C)
- Must align with the needs and the priorities of each person/facility
- Community building is integral/essential to the effort
- "Small" money can be blessing
  - More freedom for experimentation
  - Provides better motivation for collaboration

Getting to collaboration was a huge amount of work but …. there was also a LOT of luck

The SasView Aproach

An
"open, collaborative, community development"
platform for
Small Angle Scattering Data Analysis

Open, Collaborative, Community Development

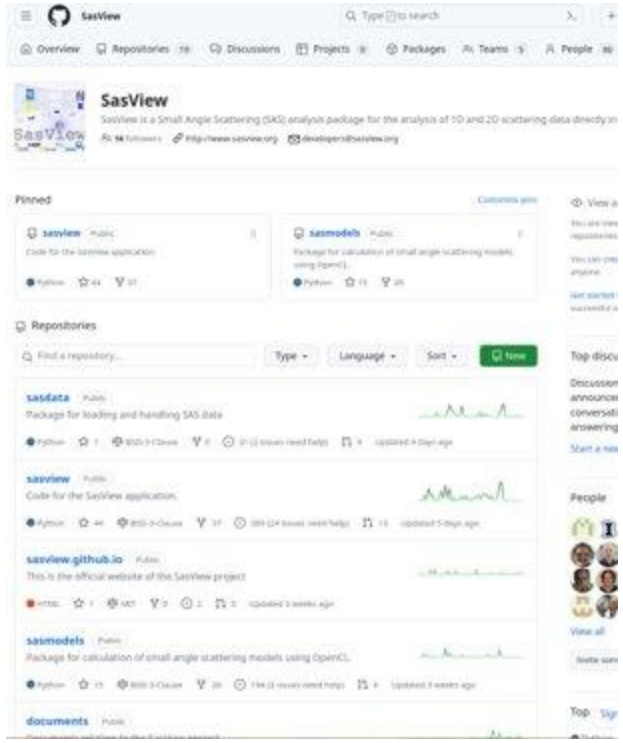
Code is open source and publicly hosted at Github

Released under BSD 3-clause license

(Zenodo) DOI for **each release**

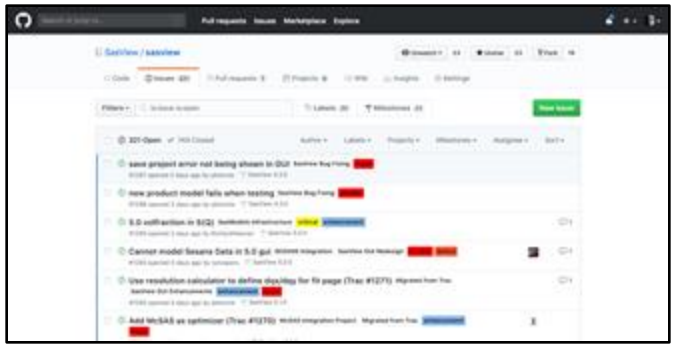
Rolling **5 Year Roadmap**

Code Hosting, Issue Tracking, Developer Wiki & CI on Public Github repos



https://github.com/SasView

https://www.sasview.org



python™

matplotlib

Open, Collaborative, Community Development

- Twice monthly zoom calls
- Regular ‘camps’ & ‘hackathons’
- Developer's mailing list
- SasView slack
- Expertise sharing and helping
- Small leadership team to facilitate

Collaboration also builds Community

http://www.sasview.org

http://github.com/SasView

Open, Collaborative, Community Development

Ask not what the community is going to do for you,
ask what you can do for the community

- P. Butler, March 2019

No MOU ... all are invited and welcome

Two Basic "Rules"

He who pays the piper gets to choose the tune…

Those who bring the resources (time and effort, or funds to buy time and effort) choose what to work on.

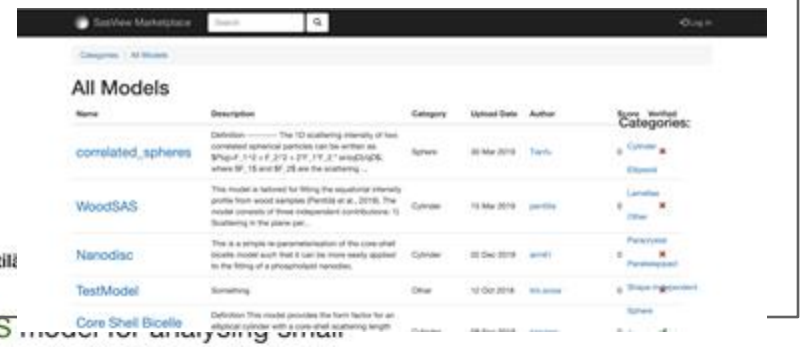
And ...

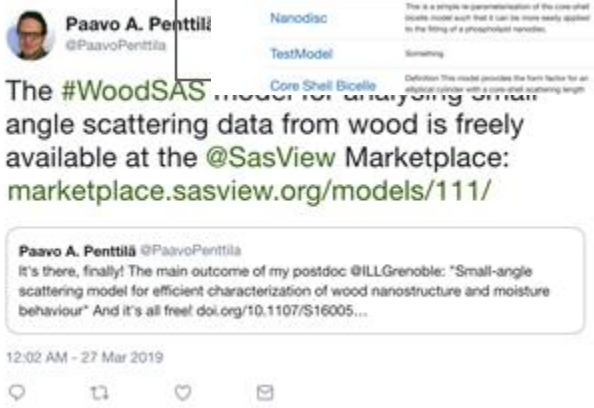
You cannot break existing experiences ...

- New dependencies vs long term maintenance (sustainability)
  - Code quality vs long term sustainability
  - Changing/degrading the current user experience for the existing user base

http://marketplace.sasview.org

Model Marketplace for Users to share their models

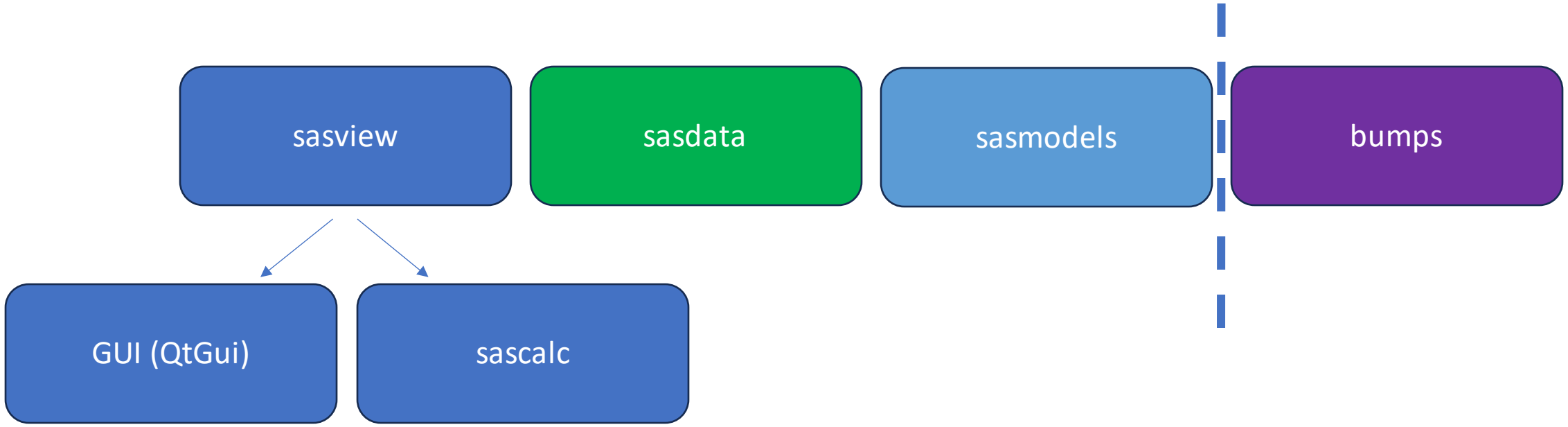




http://www.sasview.org

SasView Today


Basic Structure of code base



- *Pip install sasmodels sasdata*
- Need to finish separating GUI from sascalc … discussion on how
- Documenting entry points to GUI functionalities

DATA MANAGEMENT

Recently moved to sasdata

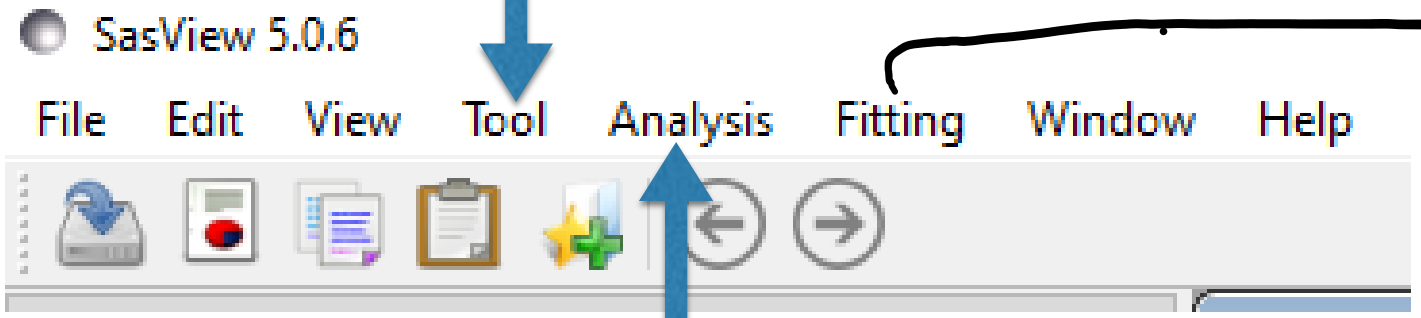
- Most all 1D data types supported on input
- Standard data supported (cansas1D and NXcanSAS)
- Some image and odd data supported through converters
- All data output in NXcanSAS format
- Calculated curves associated with data (PQ, SQ, desmeared model, etc)
- Project saving capabilities (currently a bit fragile)

SasView Toolkit for SAS data Analysis

Tools

- Data Operation
- SLD calculator
- Density/Volume calculator
- Slit Size Calculator
- Kiessig Thickness Calculator
- Q Resolution Estimator
- Generic Scattering calculator
- Orientation Viewer
- Python Shell/Editor
- Image Viewer
- File Converter

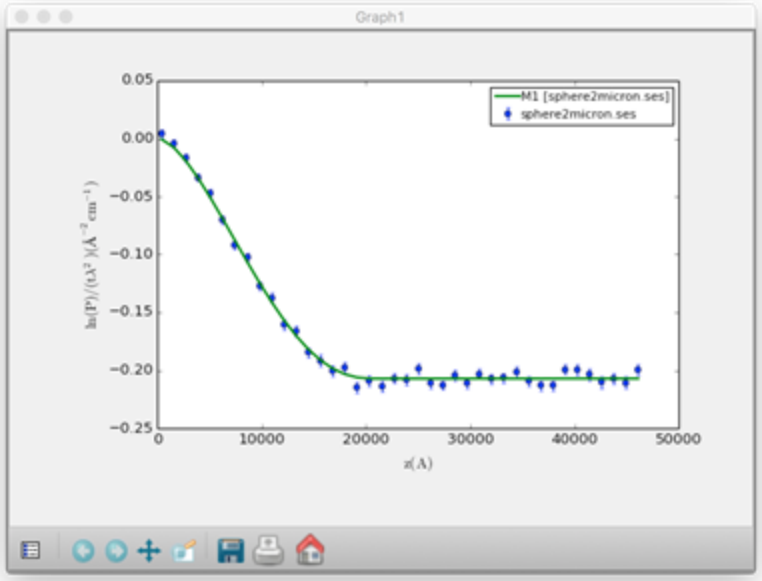
Perspectives on the data



Analaysis

- Fitting
- Invariant
- Pr Inversion
- Correlation Function

Fitting in SasView: 1D



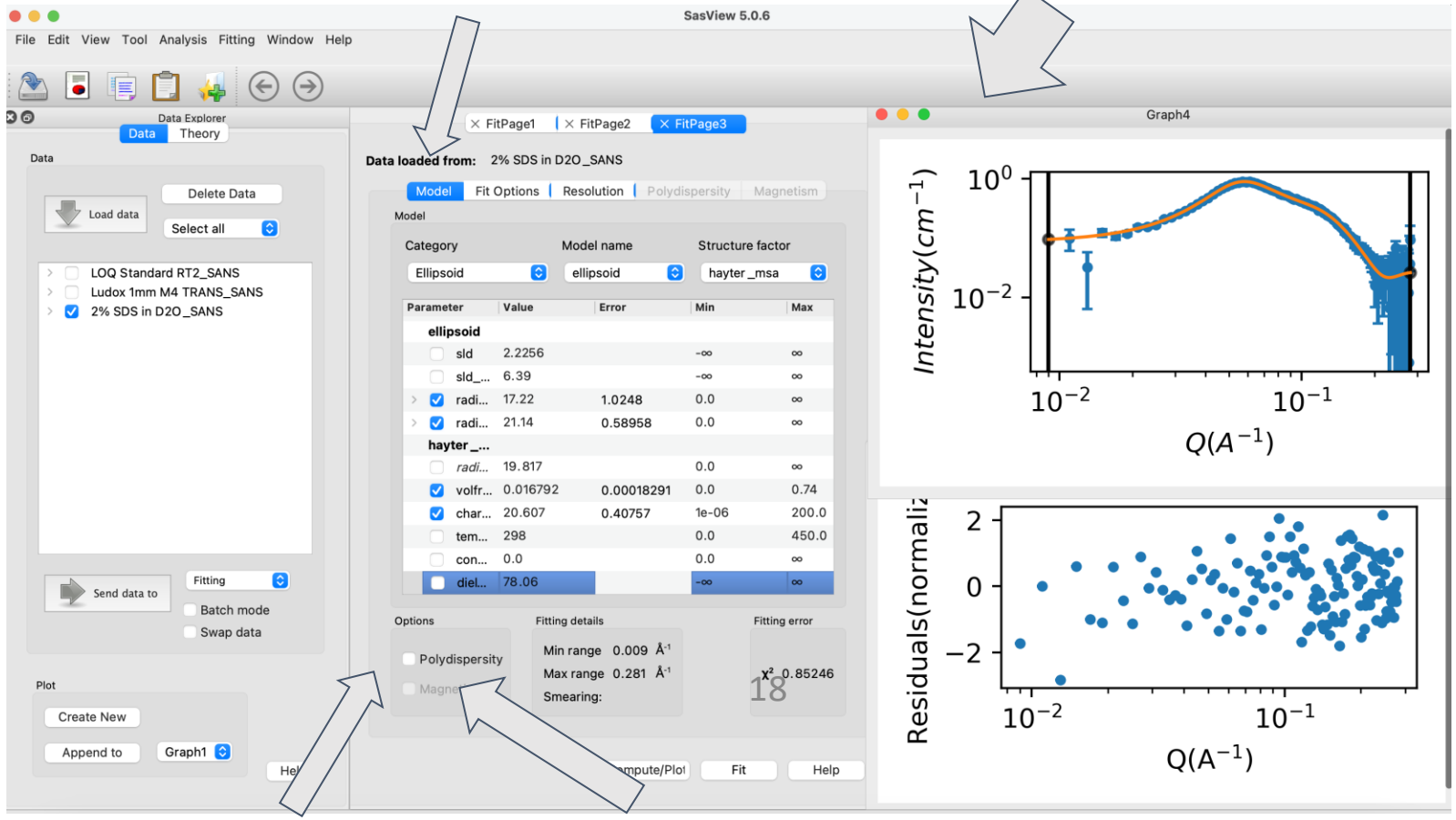
Also SESANS data

Wide choice of built-in models (> 70)
P(Q), S(Q) & P(Q)*S(Q)

Single, batch and simultaneous
1D and 2D fitting

Data management
Common data formats
supported, including
NXCansas & *cansas1D*

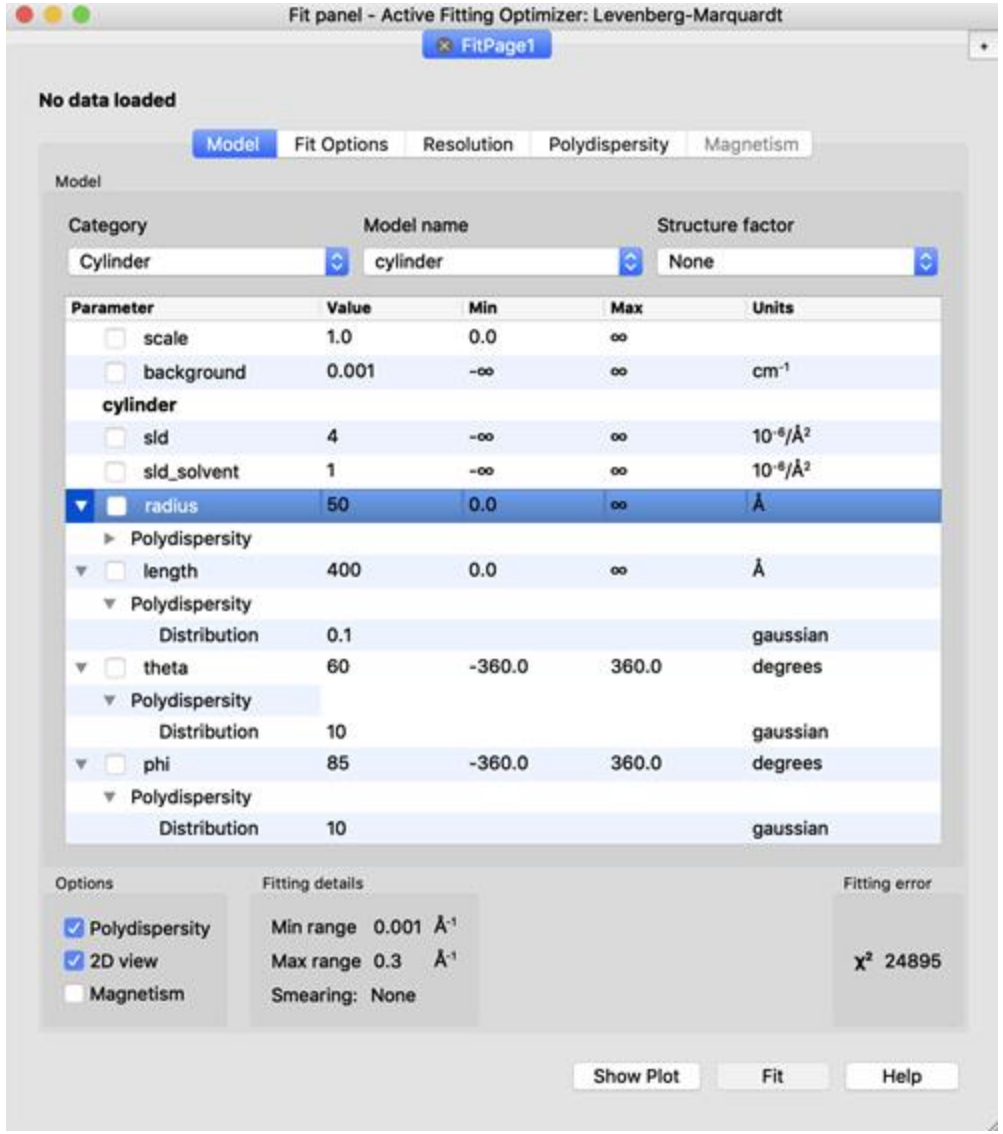
Analysis Tool Choice
&
Plotting

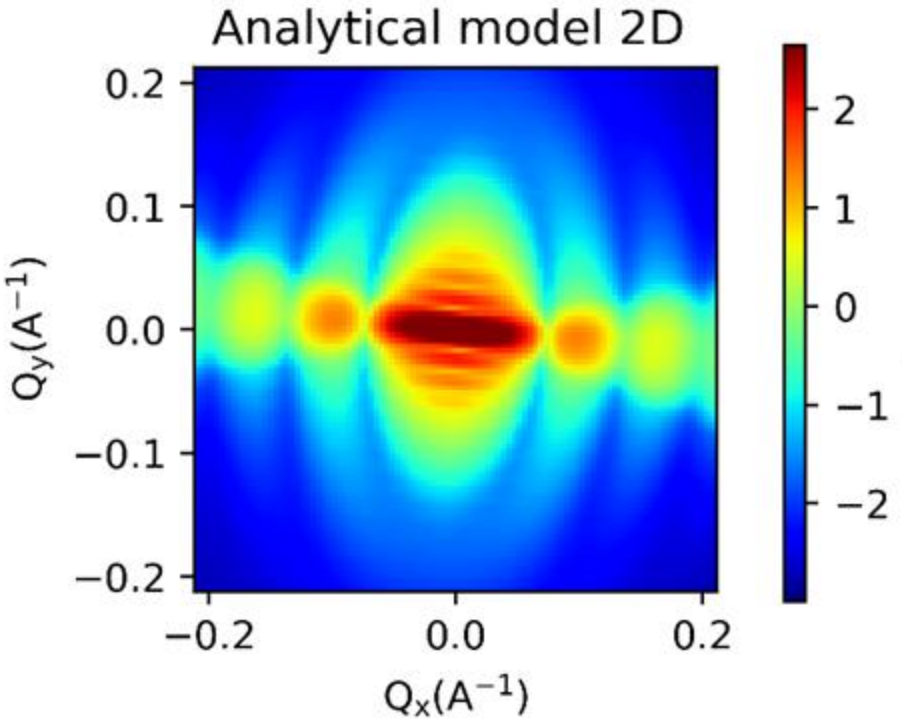


Polydispersity (choice of distribution and distribution parameters)

Resolution smearing (pinhole and slit)
Automatically from data or provide parameters

Fitting in SasView: 2D



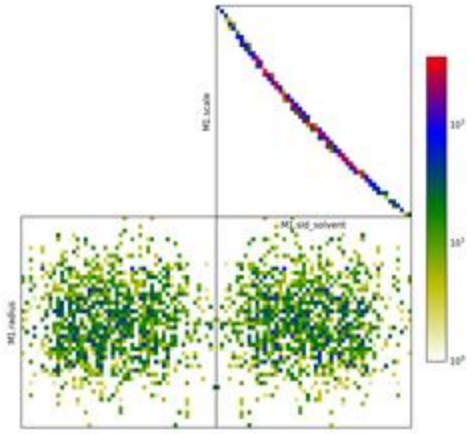


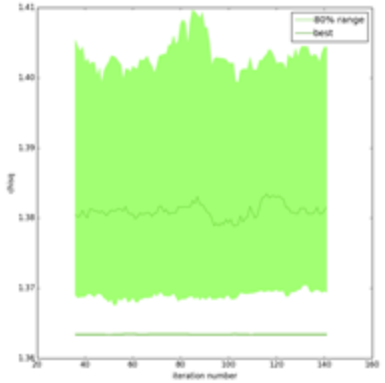
Orientational polydispersity = “jitter”

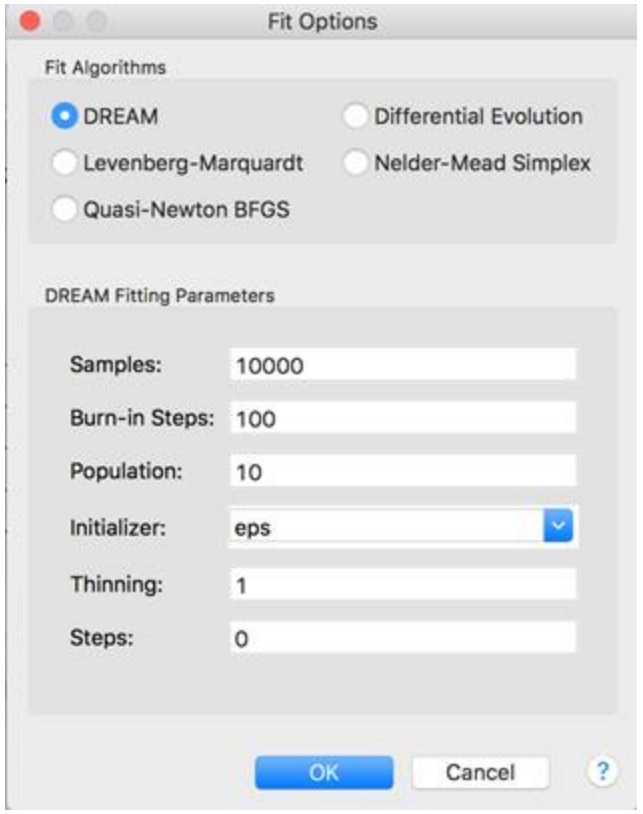
Decouples the frame for the object’s orientation with respect to the beam and the “jitter” around the axis of the object.

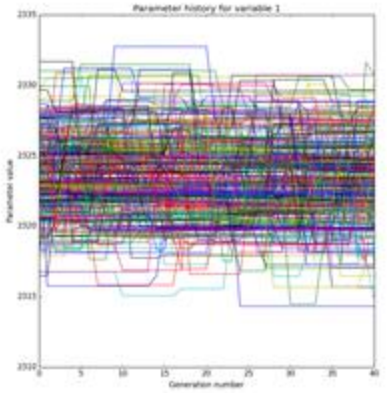
Turning on GPU Option highly recommended for fitting

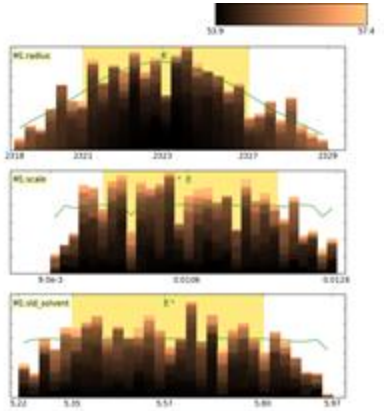
Choice of Fitting Algorithms





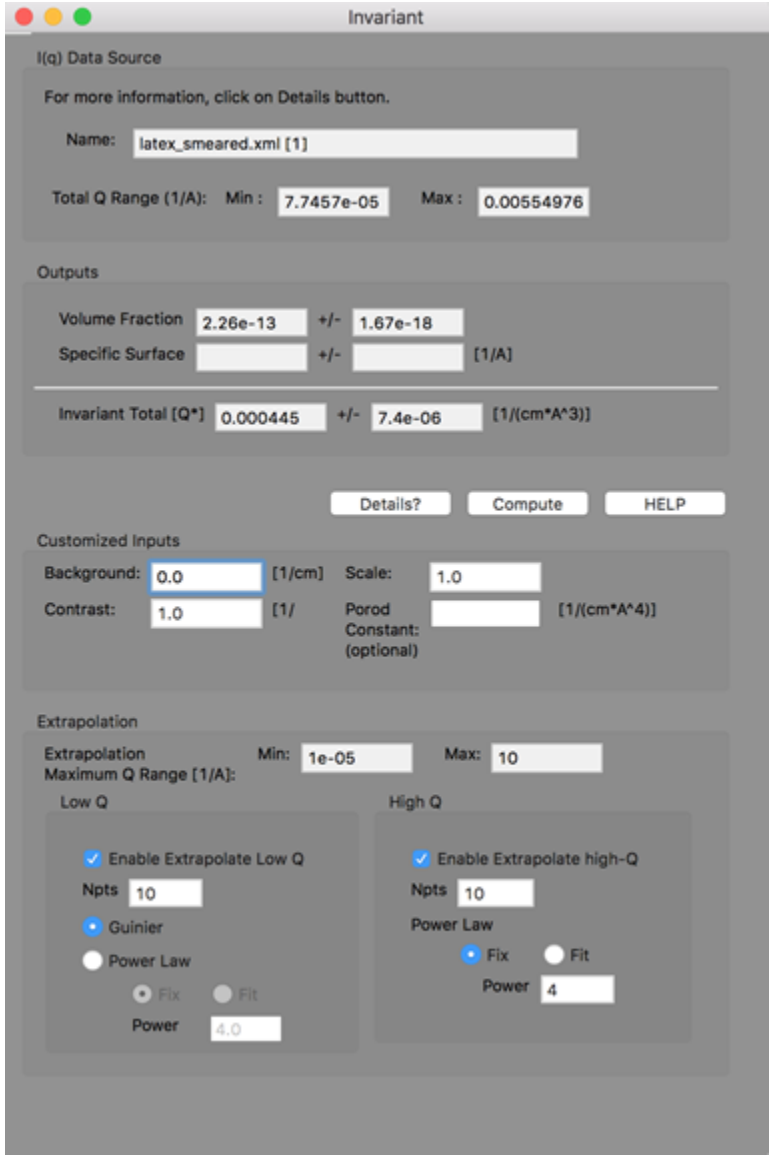


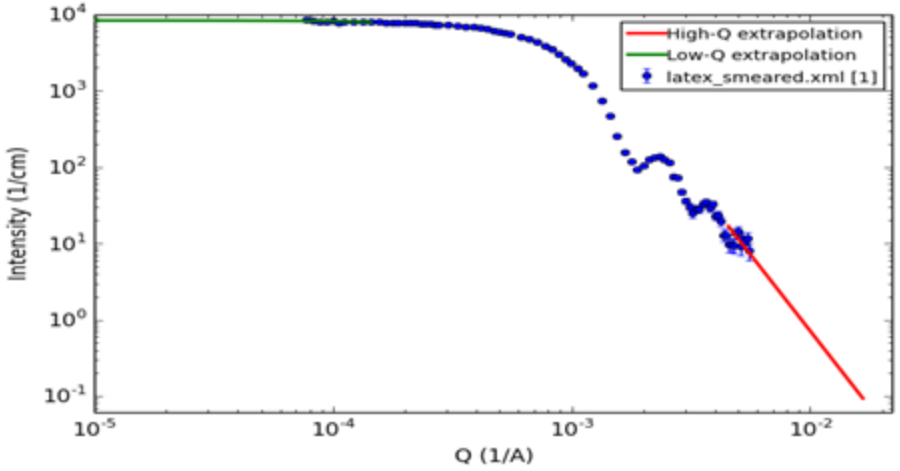




Uses bumps package from P. Kienzle
(also has DANSE origins)

Invariant Analasis





P(r) Inversion

P(r) control panel

I(q) data source

Name: 34854_db_rear.xml

Estimate background level

Slit parameters

Height | Width | [A^(-1)]

Q range

Q min 0.015 | Q max 0.35 | [A^(-1)]

Parameters

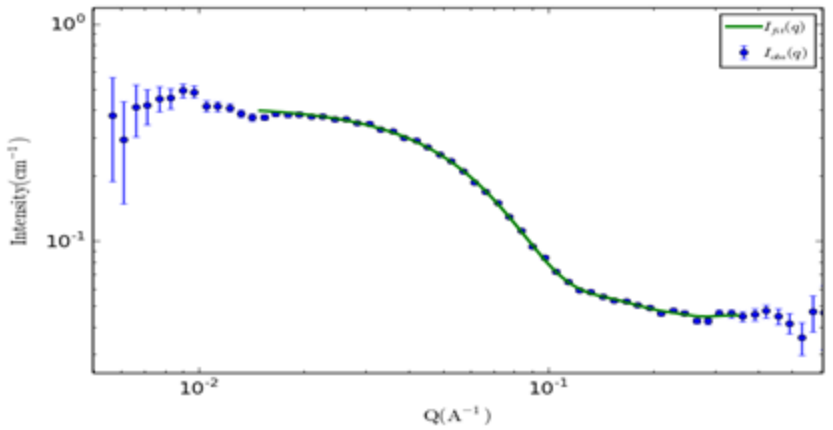
P(r) is found by fitting a set of base functions to I(Q). The minimization involves a regularization term to ensure a smooth P(r). The regularization constant gives the size of that term. The suggested value is the value above which the output P(r) will have only one peak.

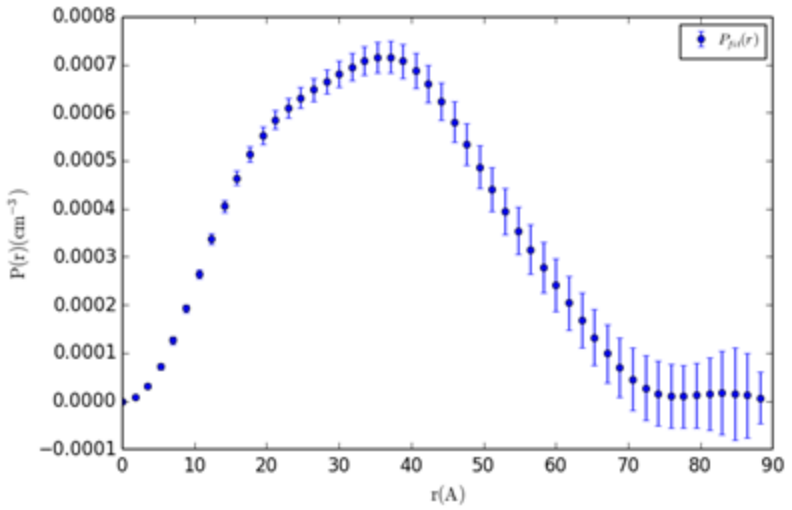
		Suggested value
Number of terms	15	15
Regularization constant	4e+09	4e+09
Max distance [A]	90	Explore

Outputs

Rg	27	[A]
I(Q=0)	0.37	[A^(-1)]
Background	0.046	[A^(-1)]
Computation time	0.0011	secs
Chi2/dof	46	
Oscillations	1.3	
Positive fraction	1	
1-sigma positive fraction	0.99	

Compute HELP



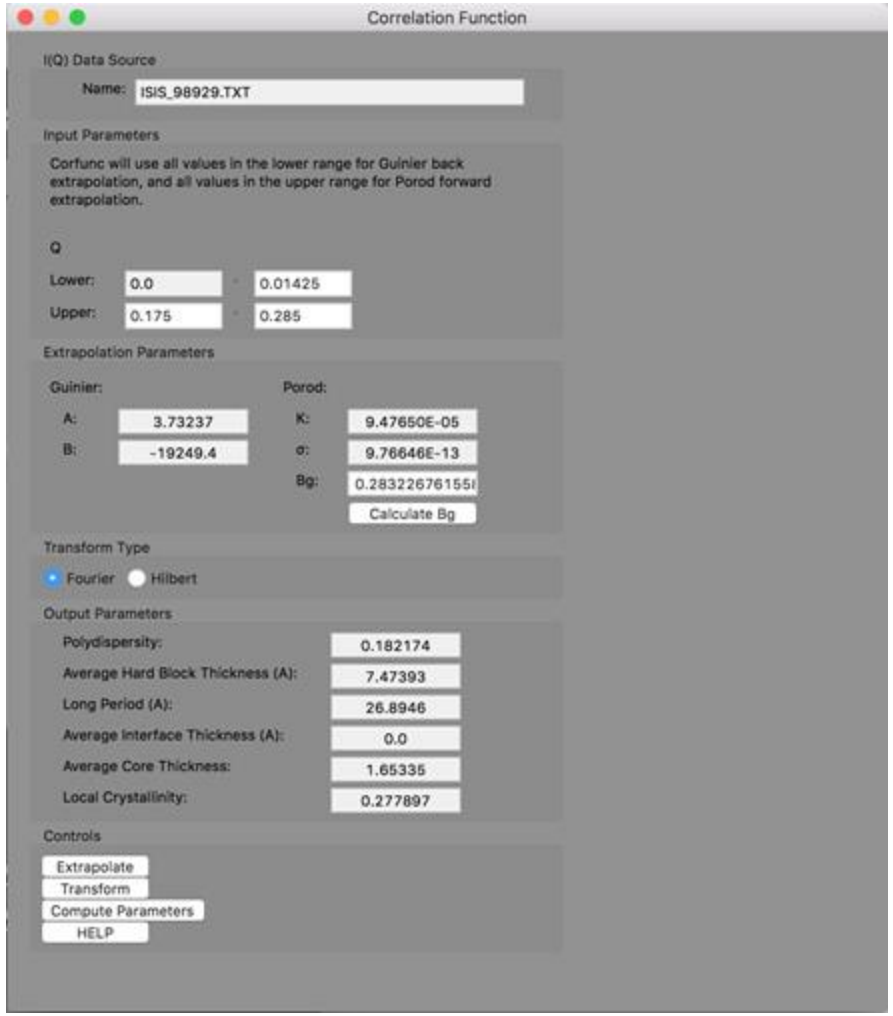


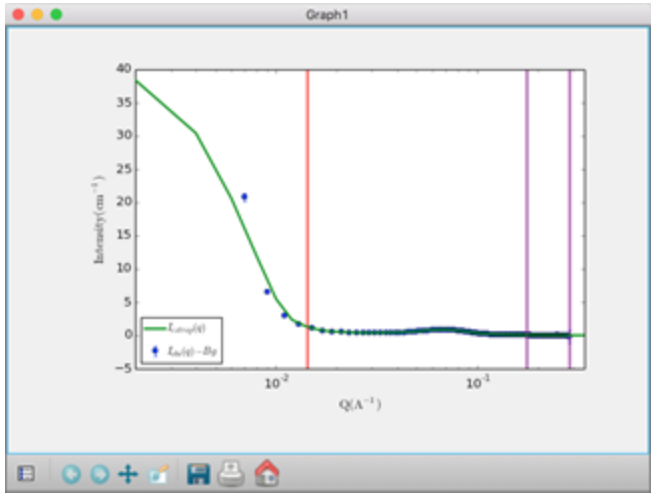
Correlation Function Analysis

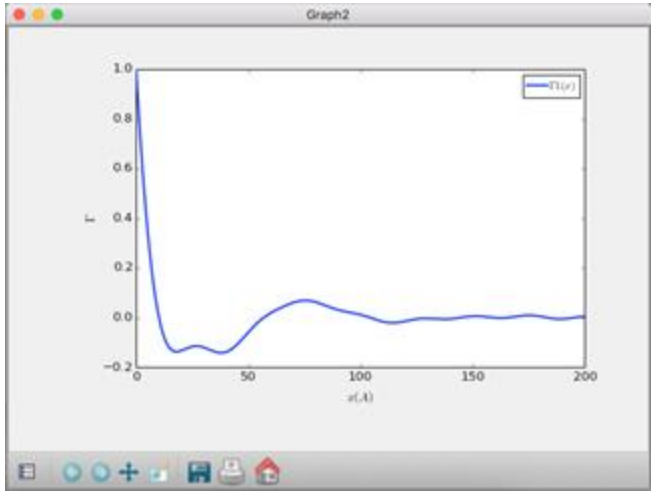
Recently refactored

CCP13 (Fiber Diffraction) legacy code (Fortran)

(ISIS summer student)



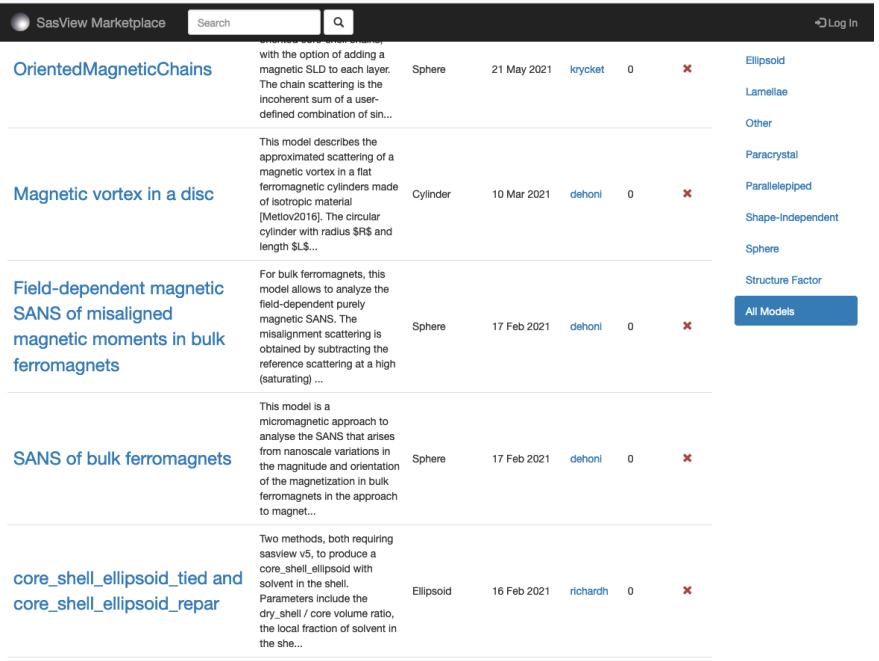


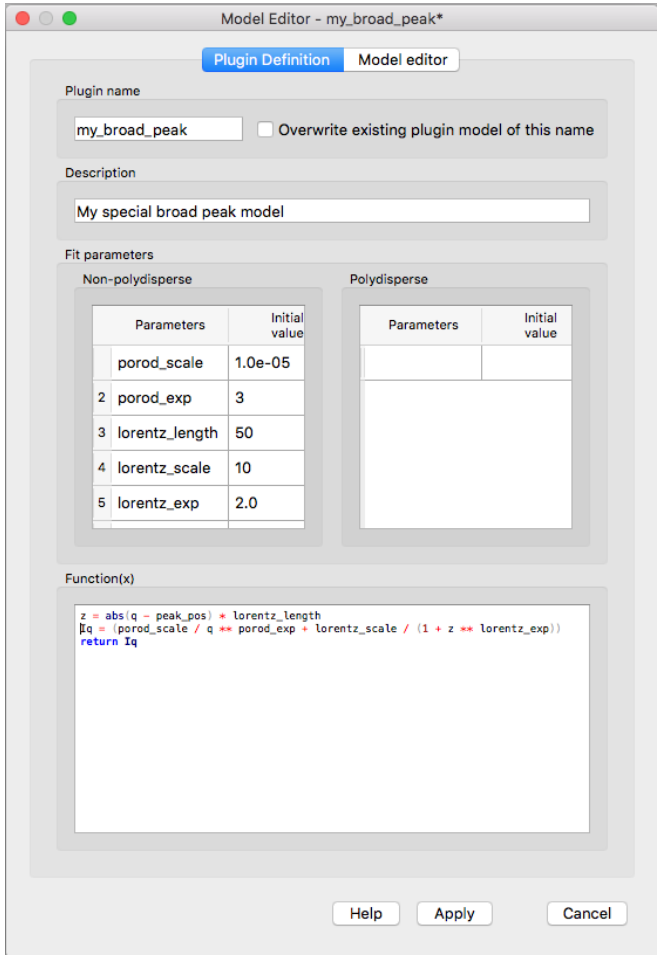


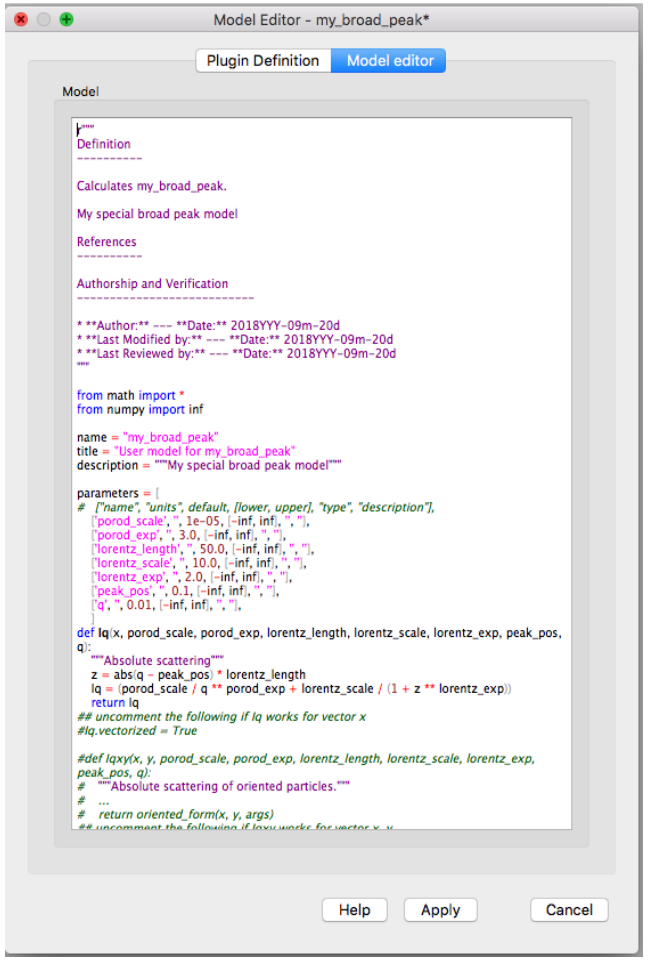
Pluging Models

SasView provides tools and infrastructure for custom/plugin models

- Dedicated editor
- Syntax and performance testing
- Directly available in SasView ecosystem
- Community developed models can be deposited to marketplace: https://marketplace.sasview.org/







Beyond the GUI

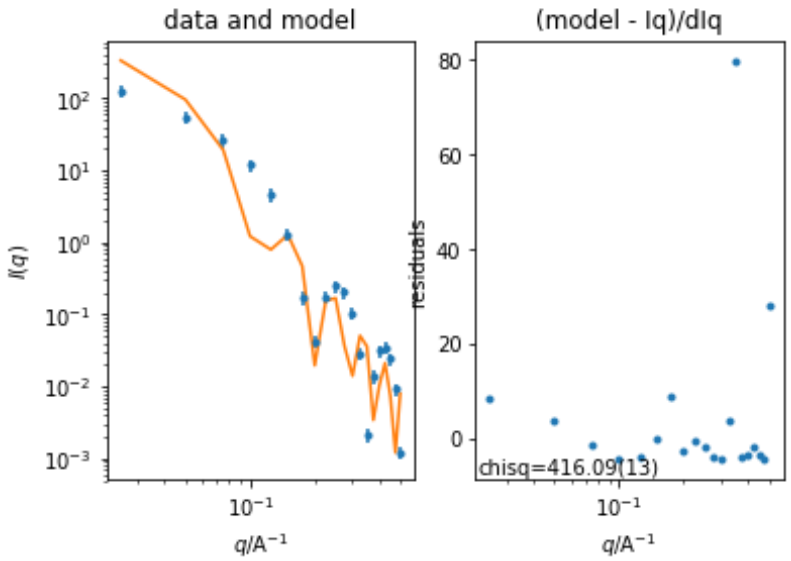
Running SasView from scripts

- Useful for batch jobs and reproducibility
- Scripts can be run on computer cluster

```
import pylab
from bumps.names import *
from sasmodels.core import load_model
from sasmodels.bumps_model import Model, Experiment
from sasmodels.data import load_data

from bumps.fitters import fit
from bumps.formatnum import format_uncertainty

test_data = load_data('cyl_400_20.txt')
kernel = load_model('cylinder')

test_data.dy = 0.2*test_data.y

pars = dict(radius=35,
            length=350,
            background=0.0,
            scale=1.0,
            sld=4.0,
            sld_solvent=1.0)
model = Model(kernel, **pars)

# SET THE FITTING PARAMETERS
model.radius.range(1, 50)
model.length.range(1, 500)

M = Experiment(data=test_data, model=model)
problem = FitProblem(M)
print("Initial chisq", problem.chisq_str())
problem.plot()
pylab.show()

result = fit(problem, method='amoeba')
print("Final chisq", problem.chisq_str())
for k, v, dv in zip(problem.labels(), result.x, result.dx):
    print(k, ":", format_uncertainty(v, dv))
problem.plot()
pylab.show()
```

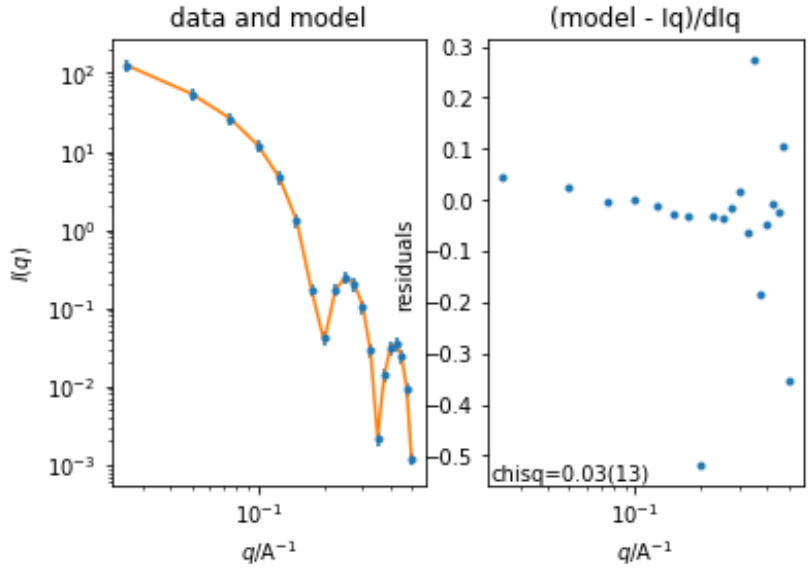
Before fit

```
Initial chisq 416.09(13)
```



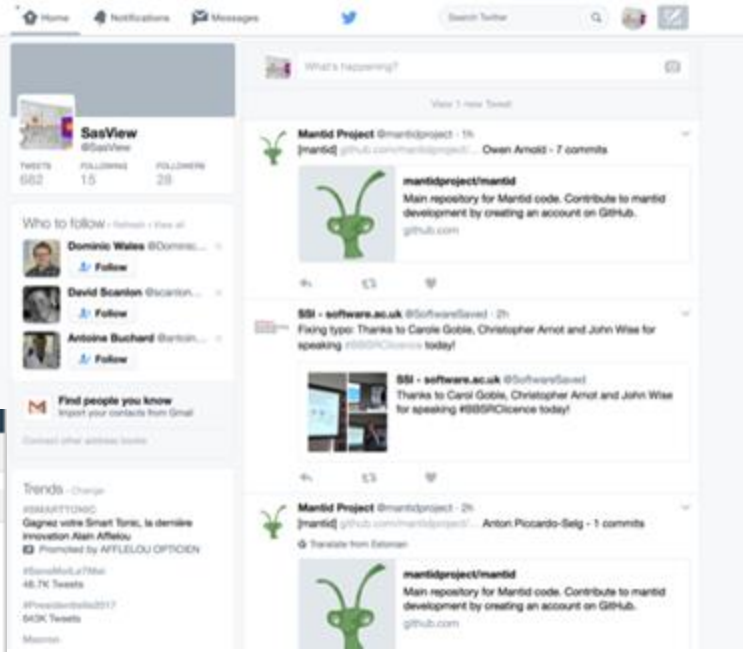
After fit

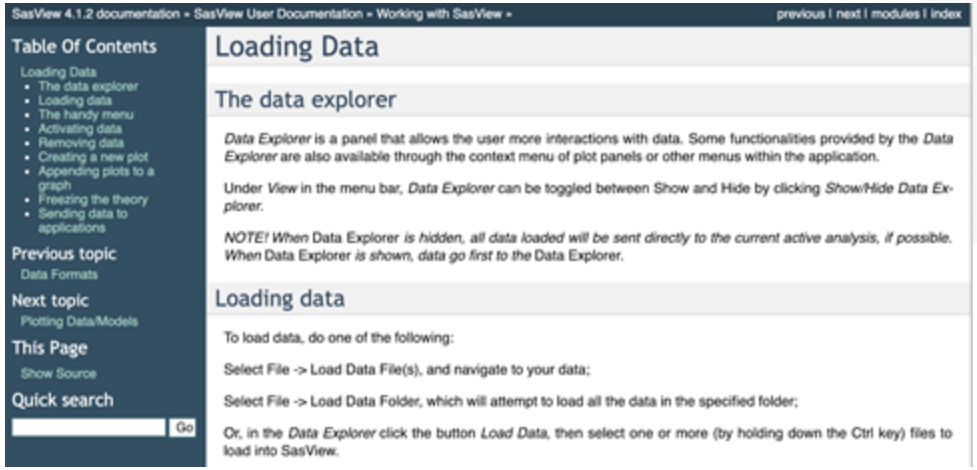
```
Final chisq 0.03(13)
length : 464.9(55)
radius : 19.977(64)
```



Resources, Education & Outreach

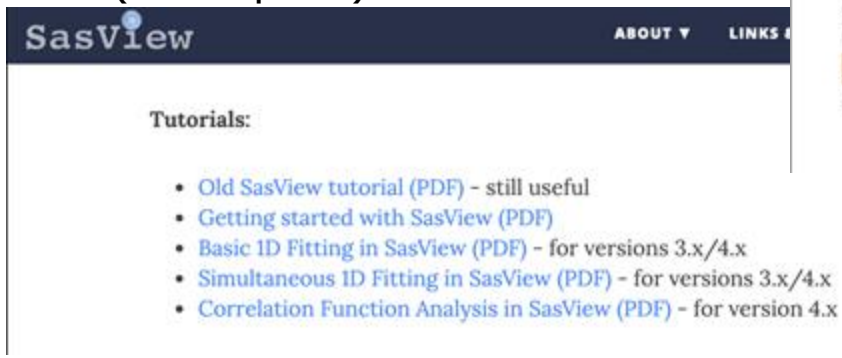
- Website
- Documentation
- Written Tutorials
- Video Tutorials (YouTube)
- Taught courses
  - Scattering schools
  - University courses
- E-learning
- Twitter
- Slack
- Mailing Lists
- Bootcamps & Regional Workshops
- (Marketplace)

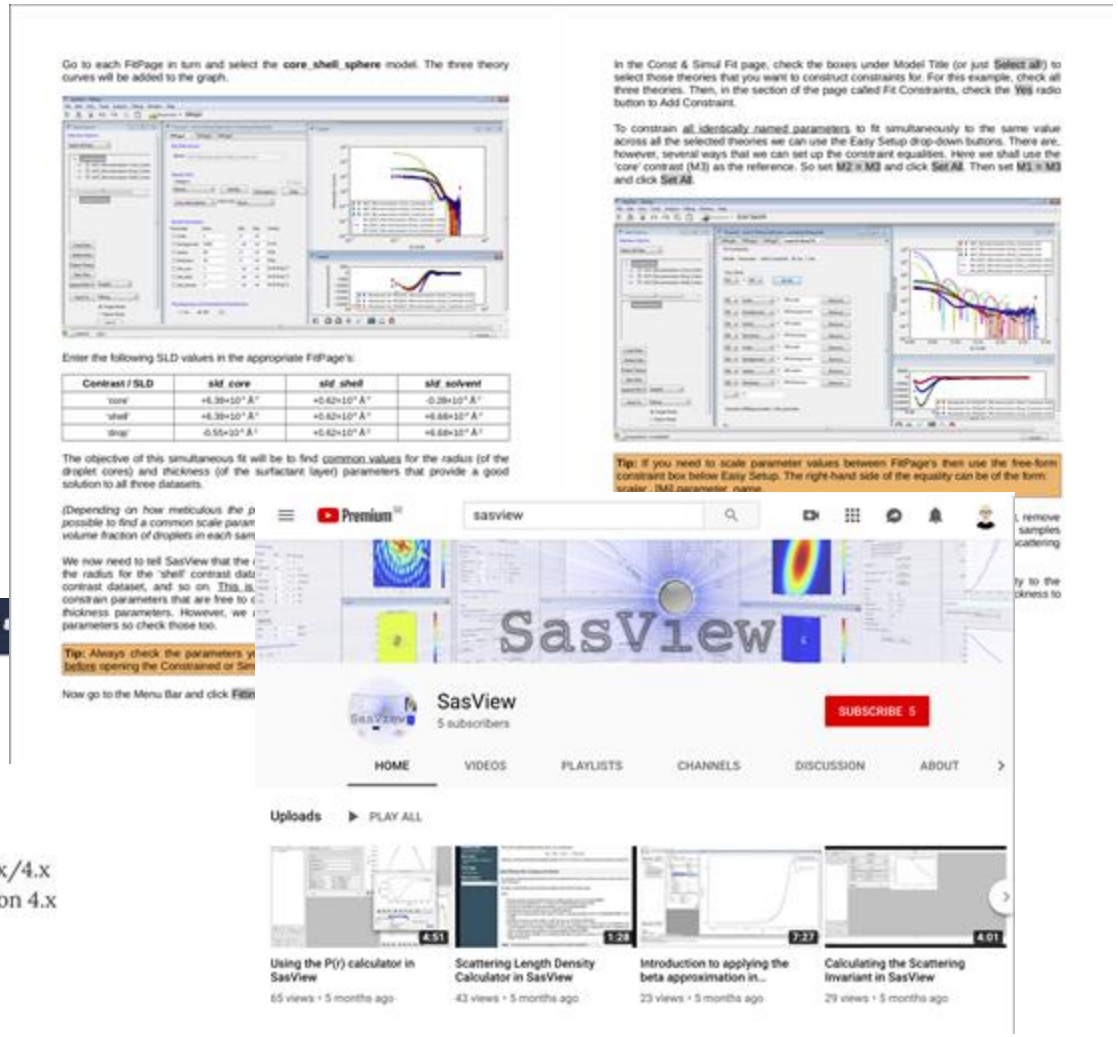




Resources, Education & Outreach

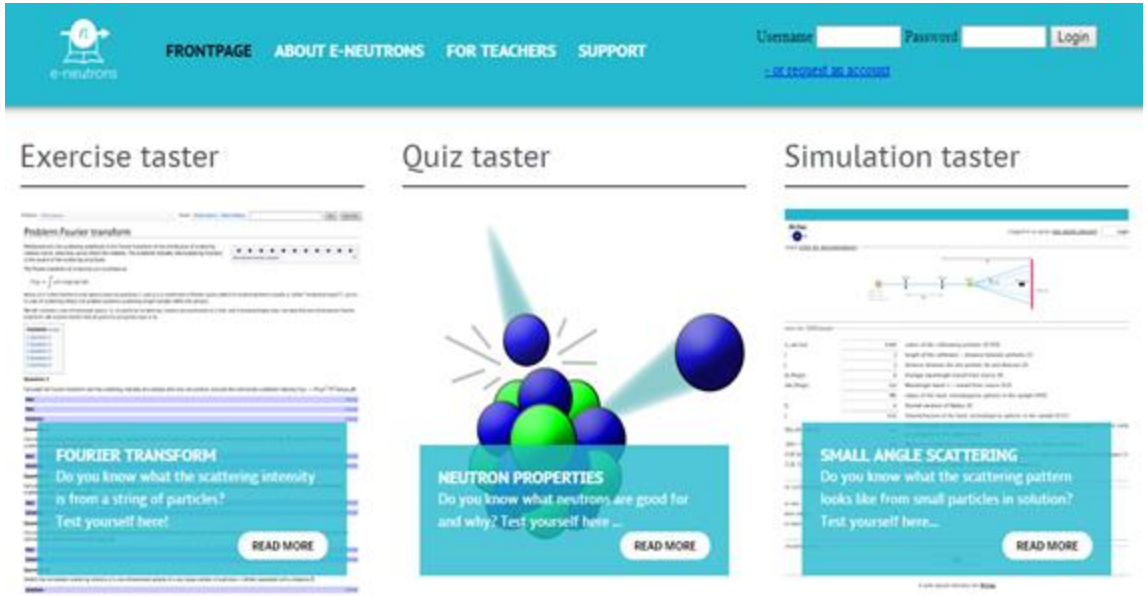
- Website
- Documentation
- Written Tutorials
- Video Tutorials (YouTube)
- Taught courses
  - Scattering schools
  - University courses
- E-learning
- Twitter
- Slack
- Mailing Lists
- Bootcamps & Regional Workshops
- (Marketplace)





Resources, Education & Outreach

- Website
- Documentation
- Written Tutorials
- Video Tutorials (YouTube)
- Taught courses
  - Scattering schools
  - University courses
- E-learning
- Twitter
- Slack
- Mailing Lists
- Bootcamps & Regional Workshops
- (Marketplace)

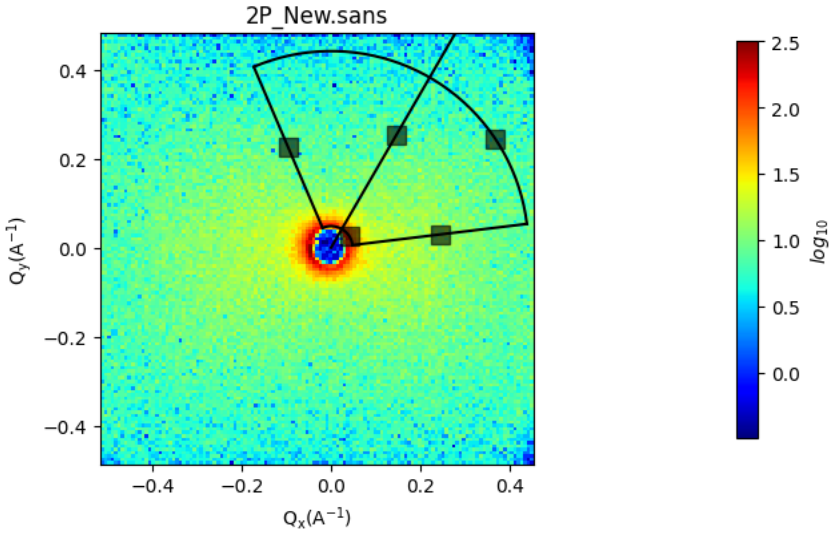


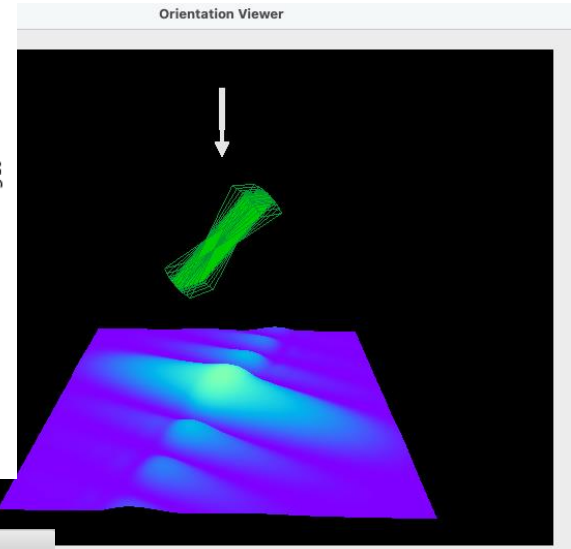
All the work of ISIS Sandwich Student Michael Oakley

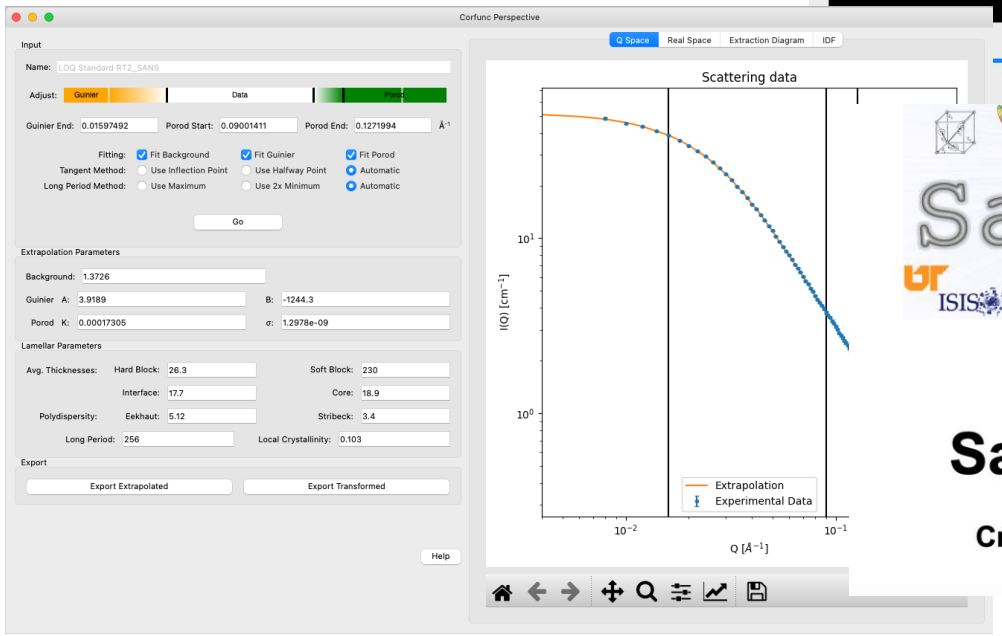
A taste of things to come – SasView 6.0.0-alpha

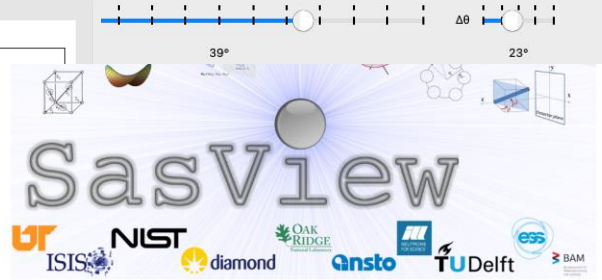
Use at your own risk…

- Orientation viewer available
- Corfunc perspective refactored
- Simultaneous fitting allows for a weighting scheme
- Preferences panel with display and plotting options
- Improved label handling on plots
- Residuals plots refactored
- PDB reader refactored
- Wedge slicer added
- Sasdata package separated
- Custom Model writing tutorial









SasView Tutorials

Creating Custom Fitting Models in SasView Version 5.x

Towards Generic Resolution Functions

```
<Idata>
<Q unit="1/A">0.00714</Q>
<I unit="1/cm">226.539</I>
<Idev unit="1/cm">1.93973</Idev>
<Qdev unit="1/A">0.00144239</Qdev>
</Idata>
```

Data reduction (resolution values)

↓

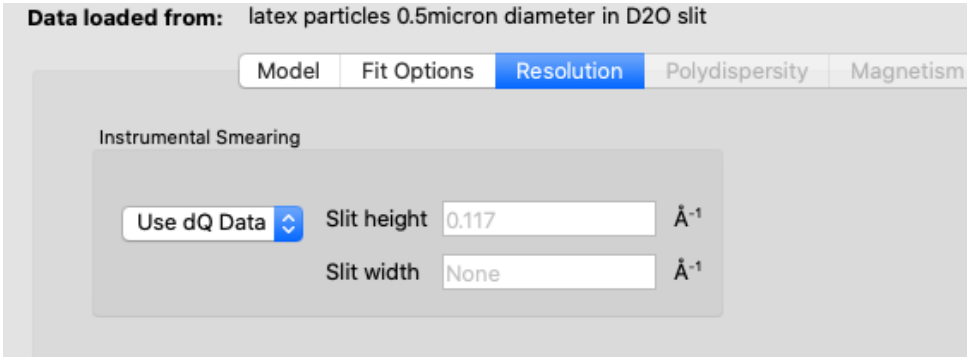
File with q, I(q), dI, dQ

↓

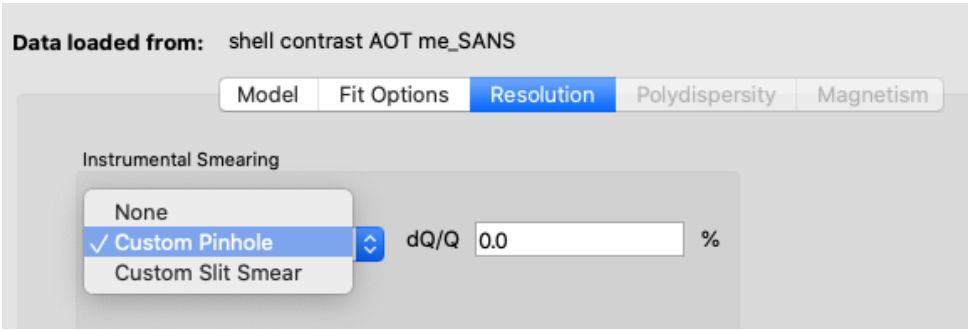
Data Analysis (reads in dQ from file)

In SasView:

File with dQ



File without dQ



Towards Generic Resolution Functions

Working with canSAS to extend NXCanSAS format with defined labels

	ResName	ResData
Analytic Gaussian	GaussFn	Qdev
General, equally spaced	CurveEq	Ndata, V0, V1,V2, ..., Vmid, ... Vend
General, arbitrary spacing	CurveGen	X0, V0, X1, V1, X2, ..., Xj, Vj, ..., Xend, Vend
User supplied	UserRes	FnName, FnData

NXCanSAS placeholders:

/entry/data/q/resolutions
Dataset name corresponding to Q resolution data

/entry/data/q/resolutions_description
Can be a simple description (i.e. Gaussian) or metadata on Q resolution data

SasView Model: Some Current Challenges

- The tragedy of the commons: somebody else will take care of the problem – New features are more interesting than fixing fragility issues
- Hard to get non-coders to believe they can contribute equally – it is unfortunately in the name "code camp" … words matter → "Contributor camp"
- Building community, especially during a pandemic is HARD WORK
- Lack of funds for small things … but beware too much money?
- Cost of writing sustainable/maintainable code is surprisingly high and can be a deterrent to new contributors

- Hard to keep up with increasing security issues with the resources so far.

- Heavily geared towards colloid science
  - NEW: Magnetic SANS effort

The price of success:

- Many people view the project as a well-funded group of professionals → barrier to contribution
- Large project with many moving pieces is a barrier to new volunteer coders

Contributing to SasView

- *You don't have to be coding ninja to help out!*

- Respond to queries posted to help@sasview.org or github
- Teach others how to get the best from SasView
- Write and improve documentation
- Write and record tutorials
- Test SasView (over and over!) and write bug reports
- Provide new plugin models
- Deploy and improve automated testing
- Review the code contributions of others
- Develop code in Python3/C and PySide2/Matplotlib
- Server and Github admin task
- Develop/improve databases (e.g. marketplace) etc.

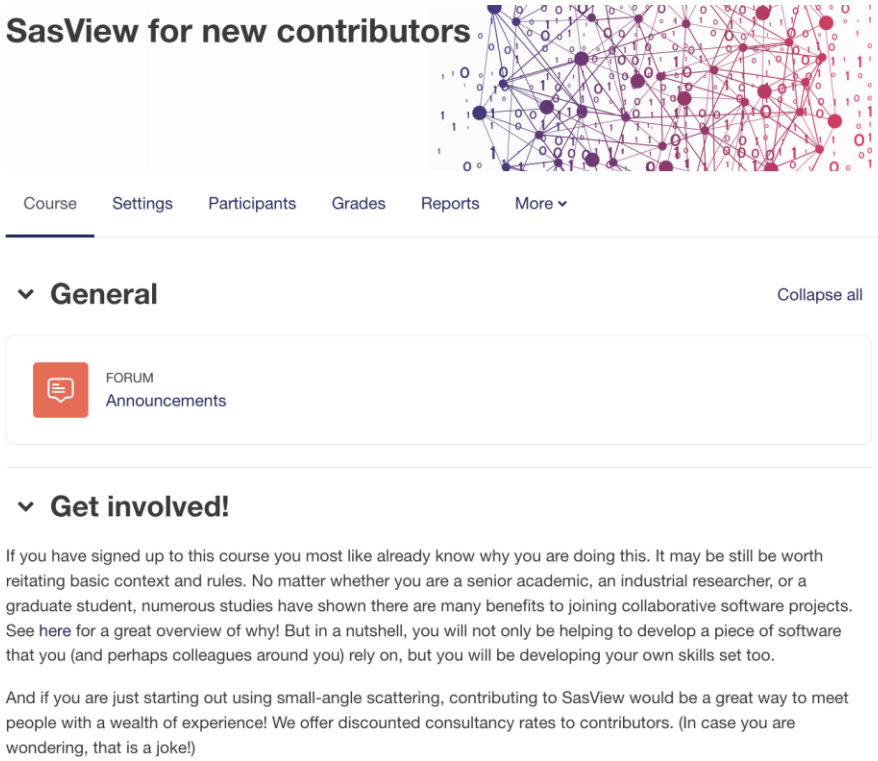
www.sasview.org/contribute

Contributing
Student Projects – Summer 2022 and 2023

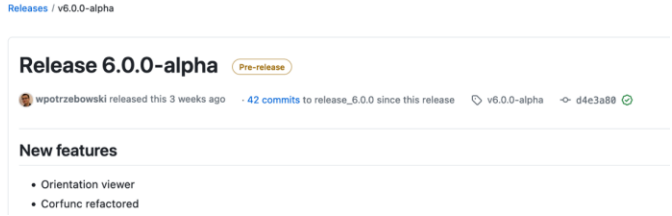
- Simultaneous SAXS and SANS fitting- **Iestyn Cadwallader-Jones**, supervisor: Miguel Gonzales (ILL)
- Plotting improvements, benchmarking for hypergeometric functions**- Nouhalia Agouzal** and **Dorian Lozano**, supervisor: Miguel Gonzales (ILL)
- Wedge slicer and improving slicer code - **Ellis Hewis**, supervisor Dirk Honecker (ISIS)
- Extending scattering calculator for magnetic systems – **Ruben Lopes**, supervisor Dirk Honecker (ISIS)
- Extending scattering calculator for Interacting bio-molecules – **Alex Zheng**, supervisor Yun Liu (NIST)
- Local documentation generator and editor (plugin models) – **Brayden Miller**, supervisor Paul Butler (NIST)
- Web-based API, **Xael Shan**, supervisor Jeff Kryzwon (NIST)

Resources for Contributors

- Reducing entry barier and improving release cycles frequency
  - Faster release cycles
  - E-learning course
  - Contributor Camp



https://e-learning.pan-training.eu/course/view.php?id=52



Save the date!
Jan 16th – 22nd 2024
University of Delaware, US

What will SasView do in the future?

Whatever the community contributes.

The resource Problem – The SasView Solution

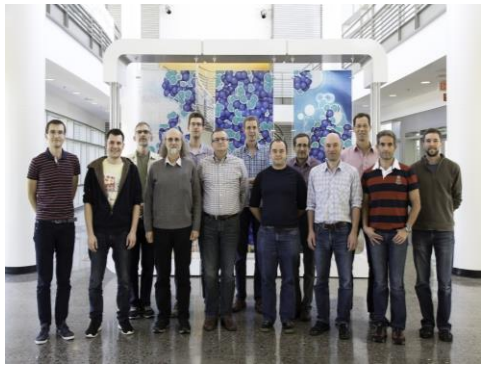
vertical and horizontal **COLLABORATIONS**

- **Facilities** provide foundational support through participation of data and instrument scientists
- **Grants and other projects** provide "bold new functionality"
- **Community**, writ large, helps provide support and functionality
- New ideas tested and developed as before by individuals or larger groups (the community)
- Once validated and deemed ready for the larger community these groups provide resources (their labor) to integrate while active developer community helps with training on where things go and on parts of interest to them (collaborative)

Current Status of SasView

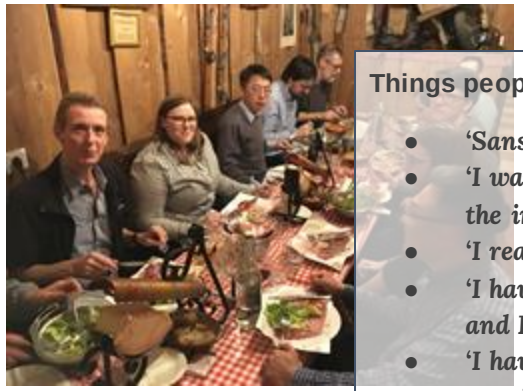
- 2006; originates in NSF *DANSE* project
- 2013; transitions into a community project
- 2016; Sine2020 project funded
- 2023; Essentially a “volunteer army”

- ~40 contributors from 9 organizations so far (~10-15 active at any one time)
- 1 to 2 releases/year (5.0.5 JUST RELEASED)
- Documentation/tutorial projects ongoing
- Usage? Seems to be "everywhere?"
- Lots of "complements" (a bit scary)
- Publications? > 100/year



Come and Join the Fun!

Contributor Camp XII Jan 16-22 University of Delaware



Things people are saying about SansView/SasView

- *'SansView is a very helpful tool, very complete and easy to use'* - Niki
- *'I want to thank you for this amazing software. It's UI and options make the interpretation of spectra easier and faster'* - Philippe
- *'I really like the SasView software'* - Martin
- *'I have been using SasView as my software of choice for fitting SANS data, and I have been very happy with the software'* - Greg
- *'I have found SasView very easy to use and the batch fit function is a wonderful time saving tool. I can finally stop making painful excel macros!'* - Andrew
- *'I am a new user of SasView and I think it is a very useful and practical tool'* - Arnaud
- *'Within 30 seconds...I am completely converted to SasView!'* - Mike
- *'Thank you for creating and maintaining SasView. It is an incredibly helpful tool, and I use it regularly'* - Pasha
- *'All the best and thank you again to carry on such a good job on SasView'* - Niki
- *'Ooooh NICE PROGRAMME!! Hours of fun!'* - Stuart
- *'I love such amazing software so much. It help our researches a lot.'* - Po-Wei