

# Developing Web Applications for Data Analysis

Mathieu Doucet / ORNL

## Why?

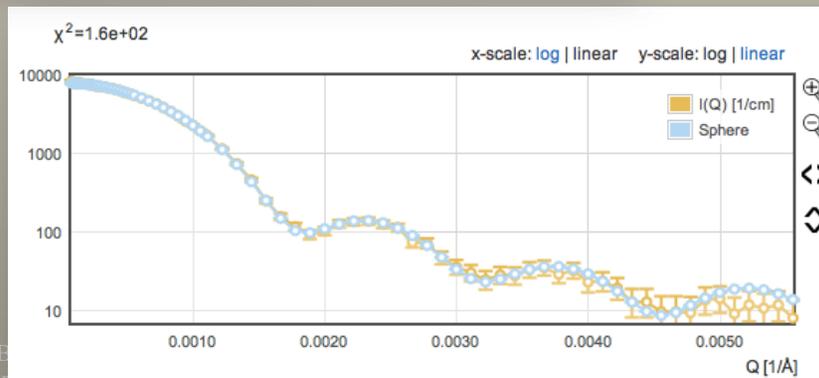
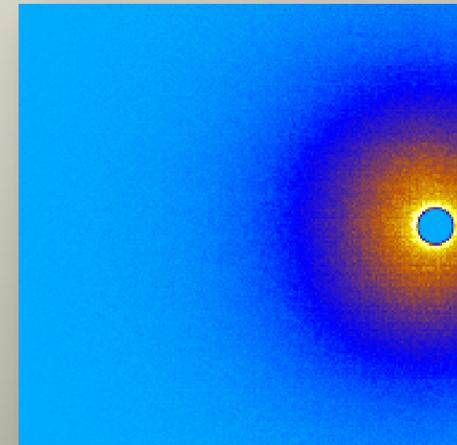
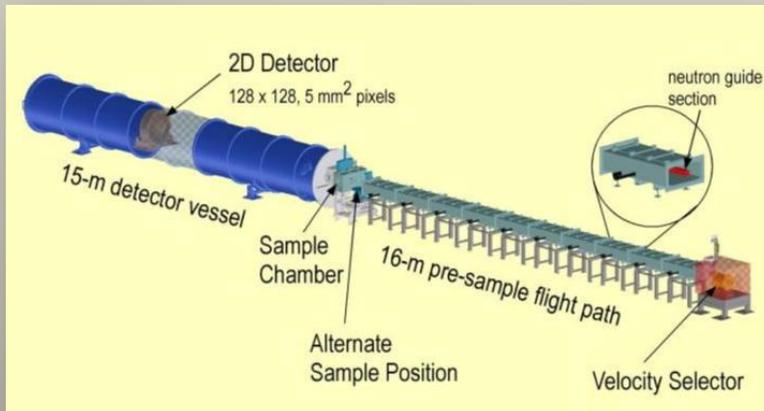
... mostly curiosity... the business world is now mostly web-based, so why not us?

## Questions:

- How hard is it, really?
- What do we gain?
- What are the limitations?
- Can we leverage this to enhance collaboration?

# Start with something known: SANS analysis

- Focus on the architecture
- Focus on answering the questions
- Re-used knowledge and computational modules of DANSE's SansView and built a web application.



Reduction of raw 2D data

Radial averaging + fit

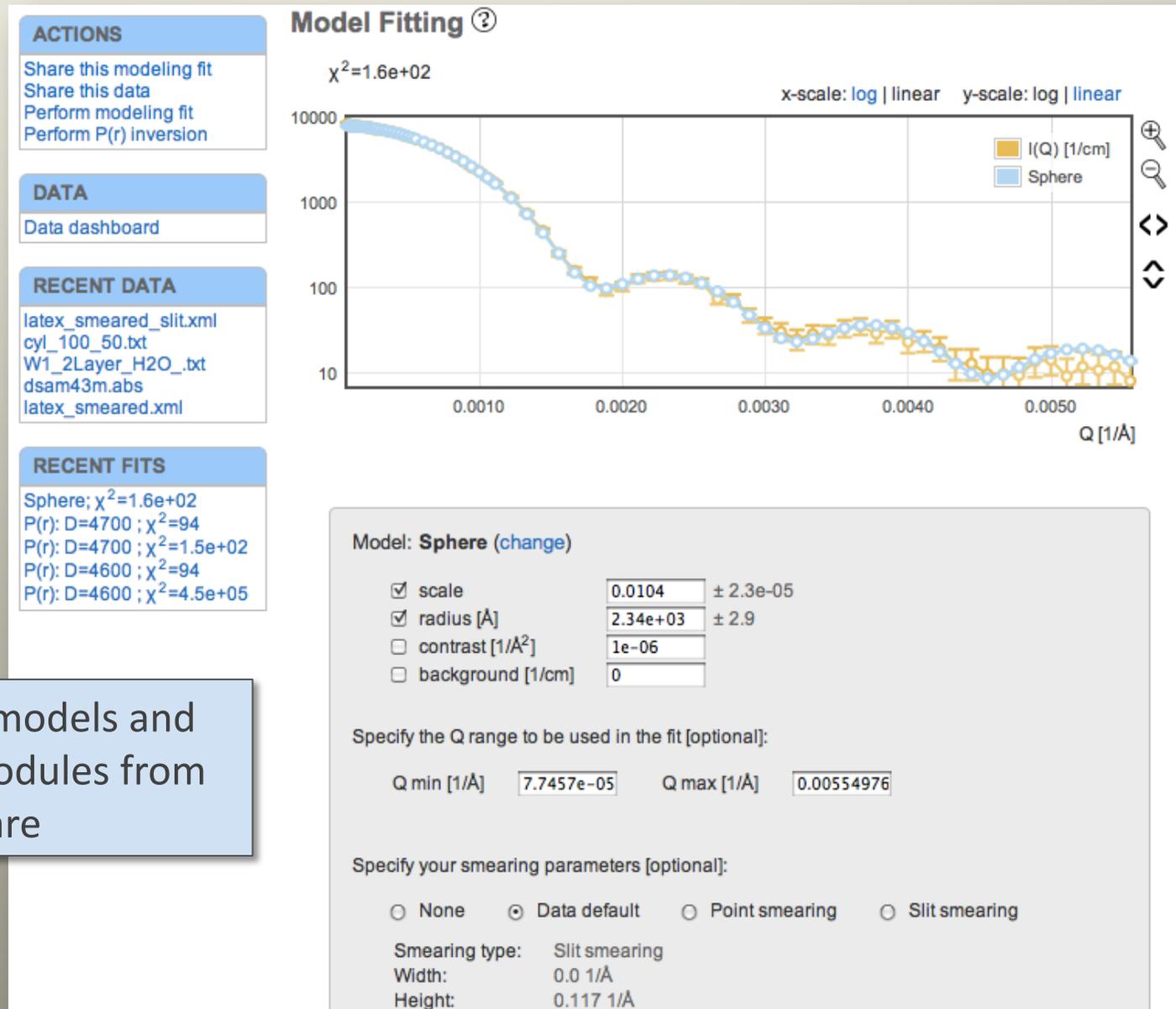
The screenshot shows a web browser window titled "SANS Data Analysis" with the URL "http://www.sansanalysis.com/". The browser's address bar includes a search engine dropdown set to "Google". Below the address bar, there are navigation links for "Apple", "Yahoo!", "Google Maps", "Wikipedia", "Popular", and "Note in Reader".

The website header features the "SANS Data Analysis" logo with a "beta" tag and a user profile for "mdoucet" with links for "logout" and "suggestions?". A "Home" link is also present.

The main content area is divided into several sections:

- DATA**: A "Data dashboard" link.
- RECENT DATA**: A list of files including "cyl\_100\_50.txt", "dsam43m.abs", "sample\_data.txt", and "W1\_2Layer\_H2O.txt".
- RECENT FITS**: A list of fit results for the same files, showing parameters like  $I(q)$  and  $P(r)$ .
- SANS Data Analysis**: A central heading with a descriptive paragraph: "The SANS Analysis site is a data analysis application that allows users to upload, analyze, and share small angle scattering data. At the moment, this application is focused on  $P(r)$  inversion calculations, but more modeling utilities are underway."
- 1 Log-in using your OpenID**: A section with a large green number "1" containing text about OpenID authentication and a link to log in.
- Upload and fit your data**: A section with a large blue and yellow striped number "2" containing text about supported data formats.
- 3 Share your results with colleagues**: A section with a large purple and green number "3" containing text about sharing results.

# Model fitting



Re-use SANS models and calculation modules from DANSE software

# Web-style navigation model

Each data and fit object has a mangled URL for sharing

data-centric actions, including sharing

quick access to recently viewed data sets

quick access to recent fits on that data set

## ACTIONS

Share this modeling fit  
Share this data  
Perform modeling fit  
Perform P(r) inversion

## DATA

Data dashboard

## RECENT DATA

latex\_smeared\_slit.xml  
cyl\_100\_50.txt  
W1\_2Layer\_H2O\_.txt  
dsam43m.abs  
latex\_smeared.xml

## RECENT FITS

P(r): D=4700 ;  $\chi^2=94$   
Sphere;  $\chi^2=1.6e+02$   
P(r): D=4700 ;  $\chi^2=1.5e+02$   
P(r): D=4600 ;  $\chi^2=94$   
P(r): D=4600 ;  $\chi^2=4.5e+05$

# Online data and fit management

SANS Data Analysis

http://www.sansanalysis.com/analysis/dashboard/ Google

Apple Yahoo! Google Maps Wikipedia Popular Note in Reader

**SANS Data Analysis** beta LOCAL USER | [logout](#)  
[suggestions?](#)

Home > Dashboard

**DATA**

Data dashboard

**RECENT DATA**

latex\_smeared\_slit.xml  
cyl\_100\_50.txt  
W1\_2Layer\_H2O\_.txt  
dsam43m.abs  
latex\_smeared.xml

**RECENT FITS**

P(r): latex\_smeared\_slit.xml  
I(q): latex\_smeared\_slit.xml  
P(r): cyl\_100\_50.txt  
I(q): W1\_2Layer\_H2O\_.txt  
P(r): dsam43m.abs

**Data sets**

File name	Owner	Modified on	
sample_data.txt	sansanalysis	Oct 24, 2009, 9:53 p.m.	✕
latex_smeared.xml	sansanalysis	Oct 25, 2009, 10:26 a.m.	✕
cyl_100_50.txt	sansanalysis	Oct 25, 2009, 10:26 a.m.	✕
W1_2Layer_H2O_.txt	sansanalysis	Feb 02, 2010, 6:19 p.m.	✕
W1_2Layer_D2O.txt	sansanalysis	Feb 02, 2010, 6:19 p.m.	✕

more...

Upload a file:  no file selected

**P(r) Inversion Fits**

File name	Owner	D <sub>max</sub> [Å]	Q <sub>min</sub> [1/Å]	Q <sub>max</sub> [1/Å]	Chi <sup>2</sup>	Created on	
sample_data.txt	System	150.0	0.0010101	0.1	6.7e-06	Oct 24, 2009, 9:53 p.m.	✕
sample_data.txt	System	150.0	0.0010101	0.1	6.7e-06	Oct 25, 2009, 10:15 a.m.	✕
cyl_100_50.txt	System	150.0	0.0010101	0.1	9.4e-06	Oct 25, 2009, 10:26 a.m.	✕
latex_smeared.xml	System	150.0	0.003797	0.401	5.6e+05	Oct 25, 2009, 4:50 p.m.	✕
latex_smeared.xml	System	150.0	0.003797	0.401	5.9e+05	Oct 25, 2009, 4:50 p.m.	✕

more...

**I(q) Modeling Fits**

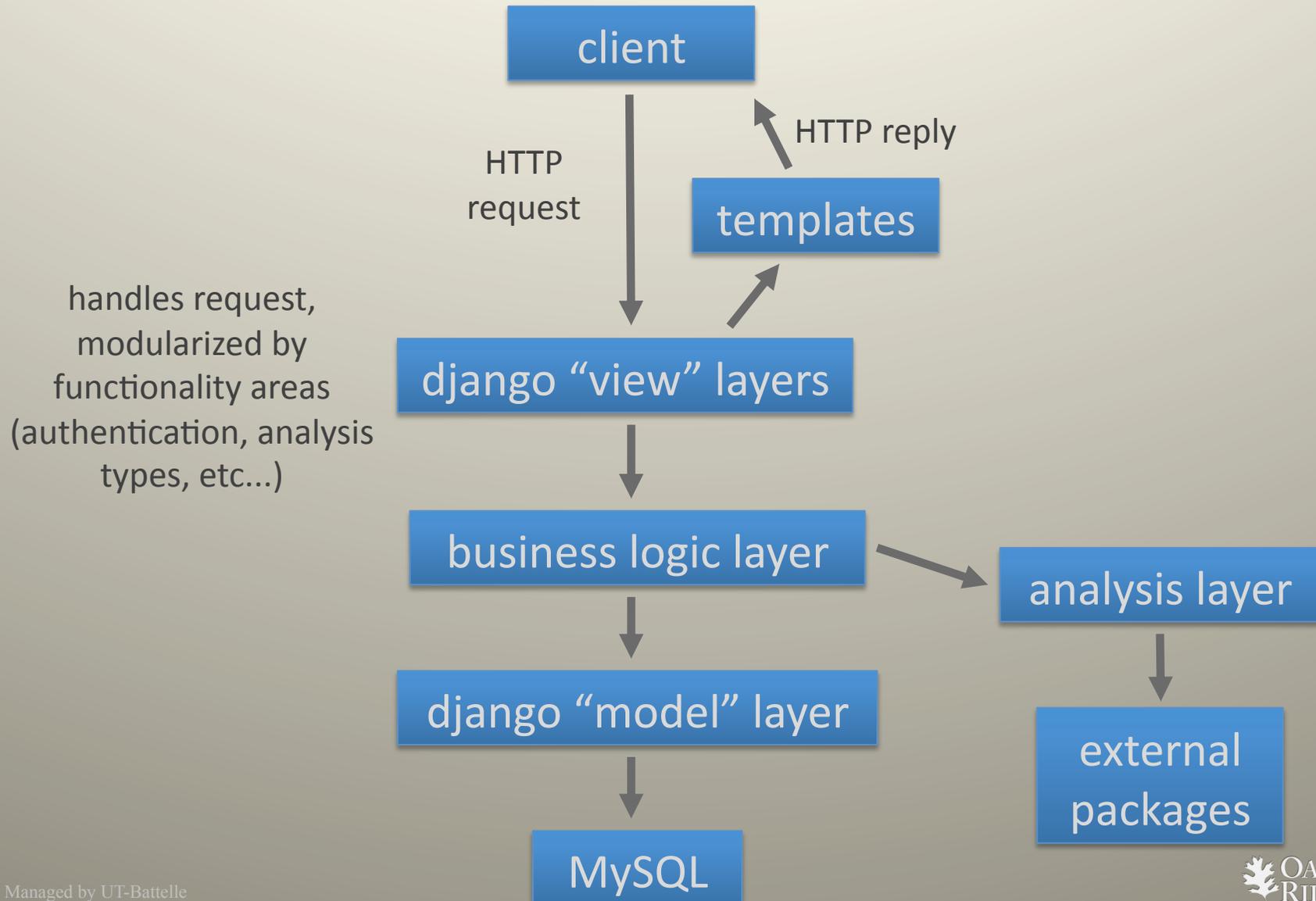
File name	Model name	Smearing	Owner	Q <sub>min</sub> [1/Å]	Q <sub>max</sub> [1/Å]	Chi <sup>2</sup>	Created on	
cyl_100_50.txt	Cylinder	None	sansanalysis	0.001	0.2	1.9e+05	Feb 01, 2010, 6:50 p.m.	✕
cyl_100_50.txt	Cylinder	None	sansanalysis	0.001	0.2	1.9e+05	Feb 01, 2010, 6:58 p.m.	✕

# Under the hood

- Started out as dumbed down web app version of SansView
- Computations all rely on existing software packages
- Build on Django / MySQL / apache
- Uses OpenID for easy authentication



# Software Architecture



# Why is this a good idea?

## For the users:

- Users always have the latest version
- Relatively platform independent
- Data management can be done server-side (like Gmail, it follows you everywhere)
- Access to larger computing power
- Good for collaboration (think Facebook for science)

## For the developers:

- Easy to deploy
- Audit tools can be built in:
  - know HOW your users use the application
  - know HOW they break it
- Forces you to design a better architecture
- With the right tools, quick to get going (tedious to get it right)

# Why is this not a good idea?

## For the users:

- Little flexibility: you can't play around with the code
- You'll still want your scripting language
- You'll still want your plotting software

## For the developers:

- Little room to let users contribute code.

# Conclusions

For the facility user who only wants to hit “GO”, web apps are perfect.

Perfect for things like data reduction, a little less for data analysis.

They are unlikely to be a great solution for the scientist who likes to play around with the code... but in reality there are few of those, and they'll learn to run locally.

At the very least, it's interesting to see what the others are doing and see if we can learn something...!

Go see the application at

[www.sansanalysis.com](http://www.sansanalysis.com)

...but not all at the same time!