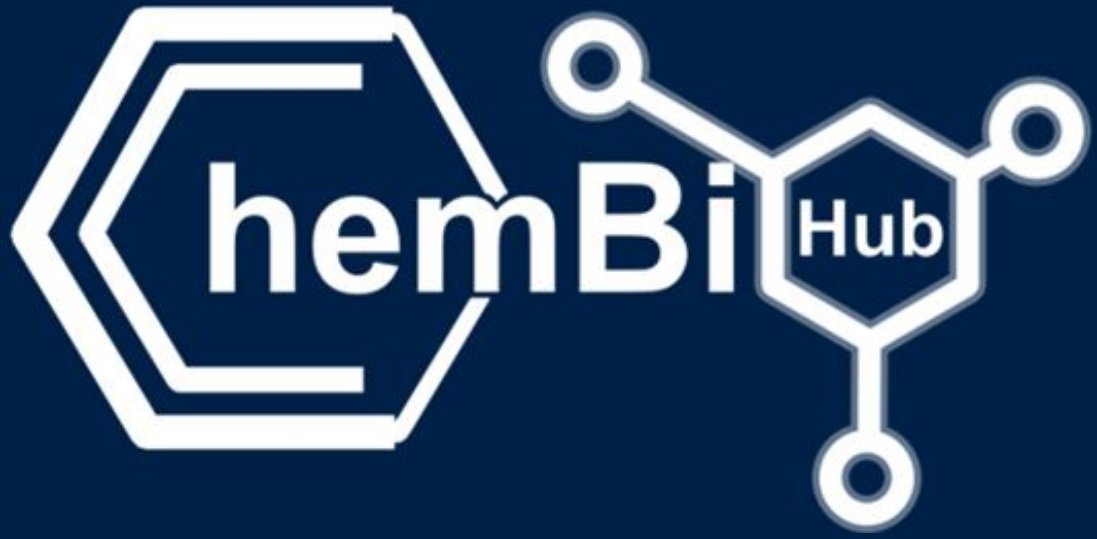
ChemBio Hub - capturing and sharing chemical biology information and knowhow

Brian Marsden
SGC
University of Oxford



http://chembiohub.ox.ac.uk

Helmholtz Open Science Webinars on Research Data
Webinar 38 – 28 September / 6 October 2016

(cc) BY

UNIVERSITY OF OXFORD

ChemBio Hub

Goals of ChemBio Hub

A place to share expertise, equipment, reagents, techniques and current areas of investigation associated with Chemical Biology.

- Improve internal communication
- Improve industry communication
- Encourage investment.
- Re-use/re-cycle pre-existing tools
- Make all available as Open Source

HELMHOLTZ GEMEINSCHAFT

Open Science

What is Chemical Biology ?

Human PIM1

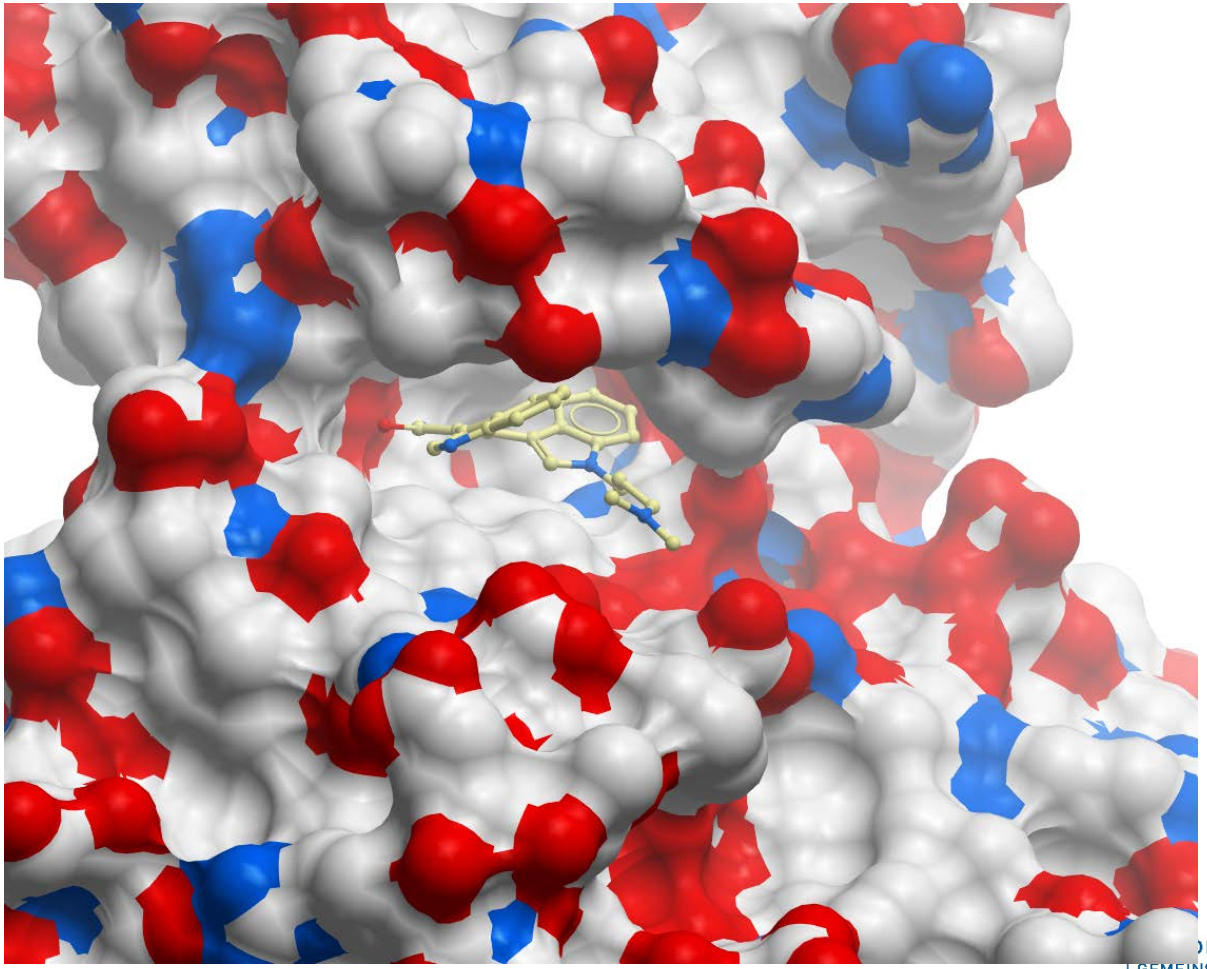
A protein kinase.

It puts a phosphate on other proteins as a means to cascade important signals within the cell.

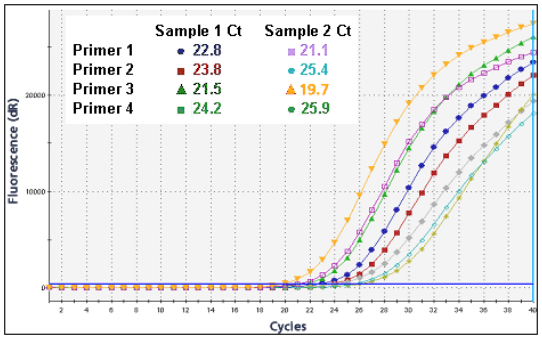
When this goes wrong, cancer results.

A small-molecule can help to prevent this by inhibiting the protein's default action.

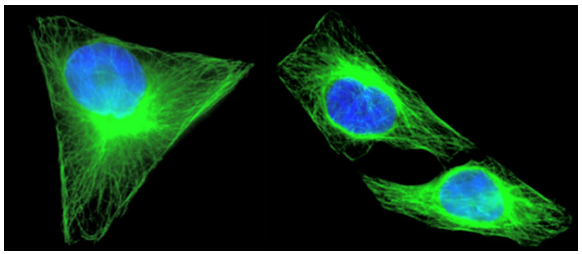
But how do we interrogate this?



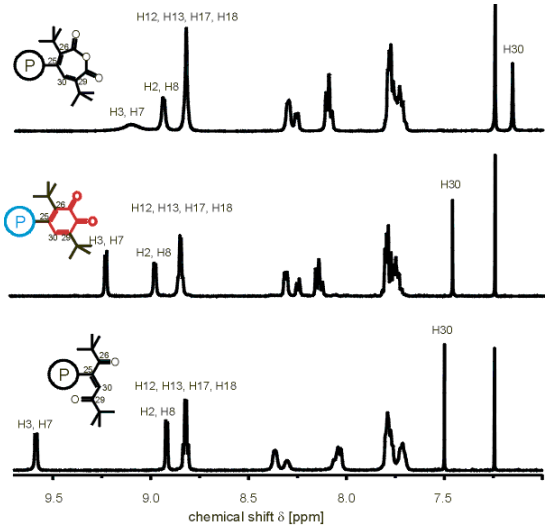
Generic assay (meta) data capture is very challenging



qPCR data

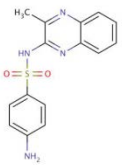


Immunostaining



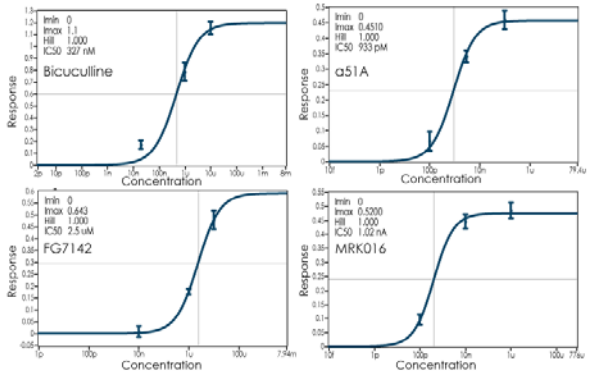
NMR spectra

There is an incredibly diverse variety of assay data types

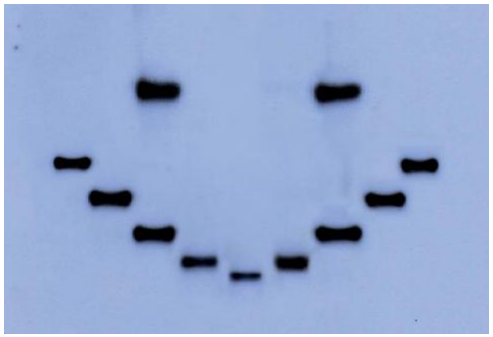


NCI 13156 C15H14N4O2S MW: 314.36

Phenotypic data



IC_{50} data



Western Blot

UNIVERSITY OF OXFORD

chemBio Hub

ChemBio Hub

Challenges for University Researchers

- **Good science**

  i.e. data management

- **Efficiency**

  duplication of effort

- **Fulfilling Grant Conditions**

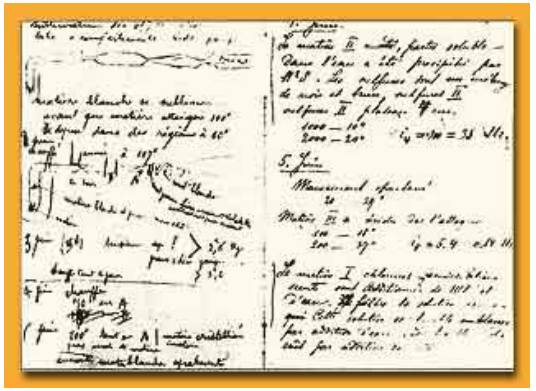
  open data

- **Personal progression**

  finding effective collaborators

- **Outcomes**

  partnerships with industry to bridge the gap



wellcome trust

NHS National Institute for Health Research

MRC Medical Research Council

HELMHOLTZ GEMEINSCHAFT

Open Science

UNIVERSITY OF OXFORD

ChemBio Hub

ChemBio Hub – What the researchers say

We approached chemists and biologists direct and asked them about:

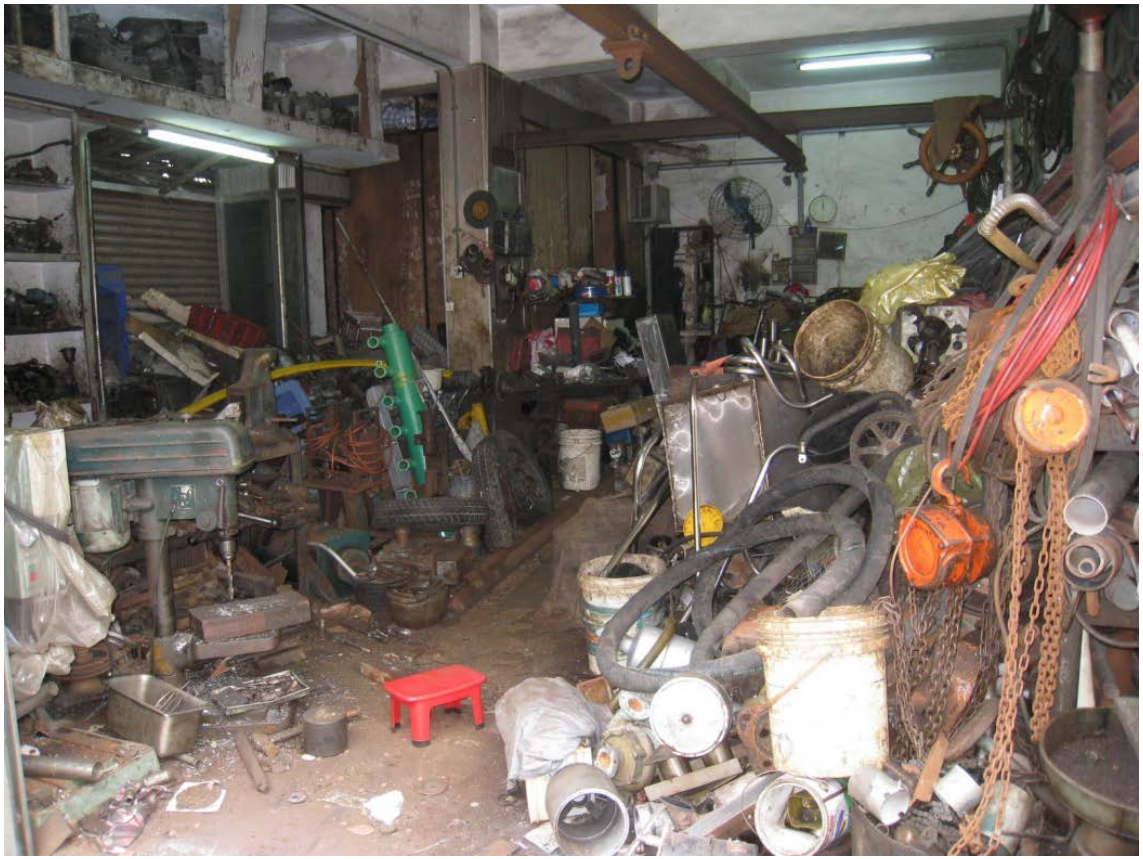
- What they were using to store data (if anything)
- What they disliked about their software
- Whether sharing data was something they currently did

Lots don't use anything at all -
paper and pen and filed in a "database"...

Sharing was not widespread but scientists could see the benefit
Searching and better data processing were high priorities

Hasn't this already been fixed?

Sure, there are lots of tools available.



UNIVERSITY OF OXFORD

ChemBio Hub

The solution – ChemBio Hub

Capture data, reagents/compounds, expertise –

- With assistance and curation
- In a central repository

Controlled levels of access –

- Within group, department, Oxford or externally

Example outcomes:

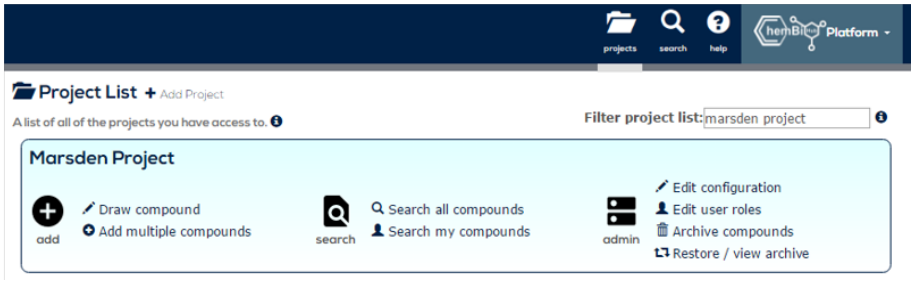
- The ‘go-to’ location for all aspects of University Chemical Biology
- Ability to discover tool compounds against target/protein of interest
- Identification of possible translational routes
- Pushing data externally, attracting pharma funding towards novel targets

UNIVERSITY OF OXFORD

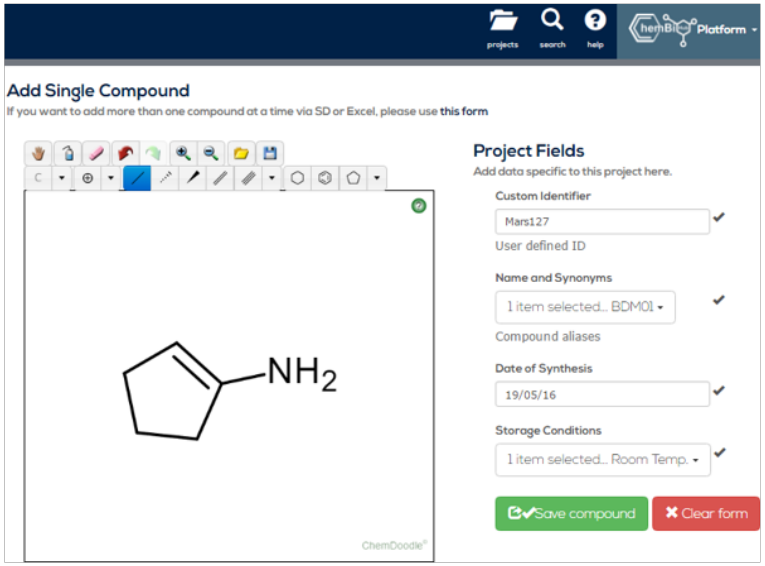
ChemBio Hub

The Web Platform

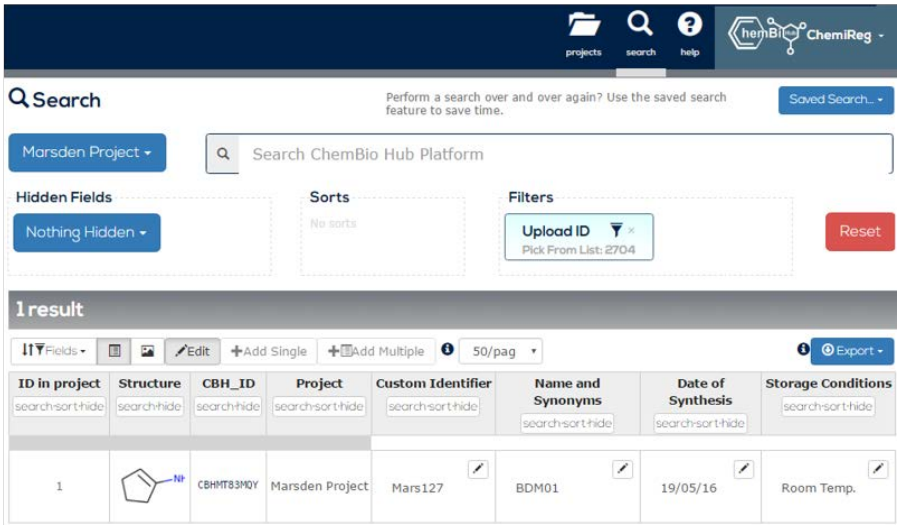
1.



2.



3.



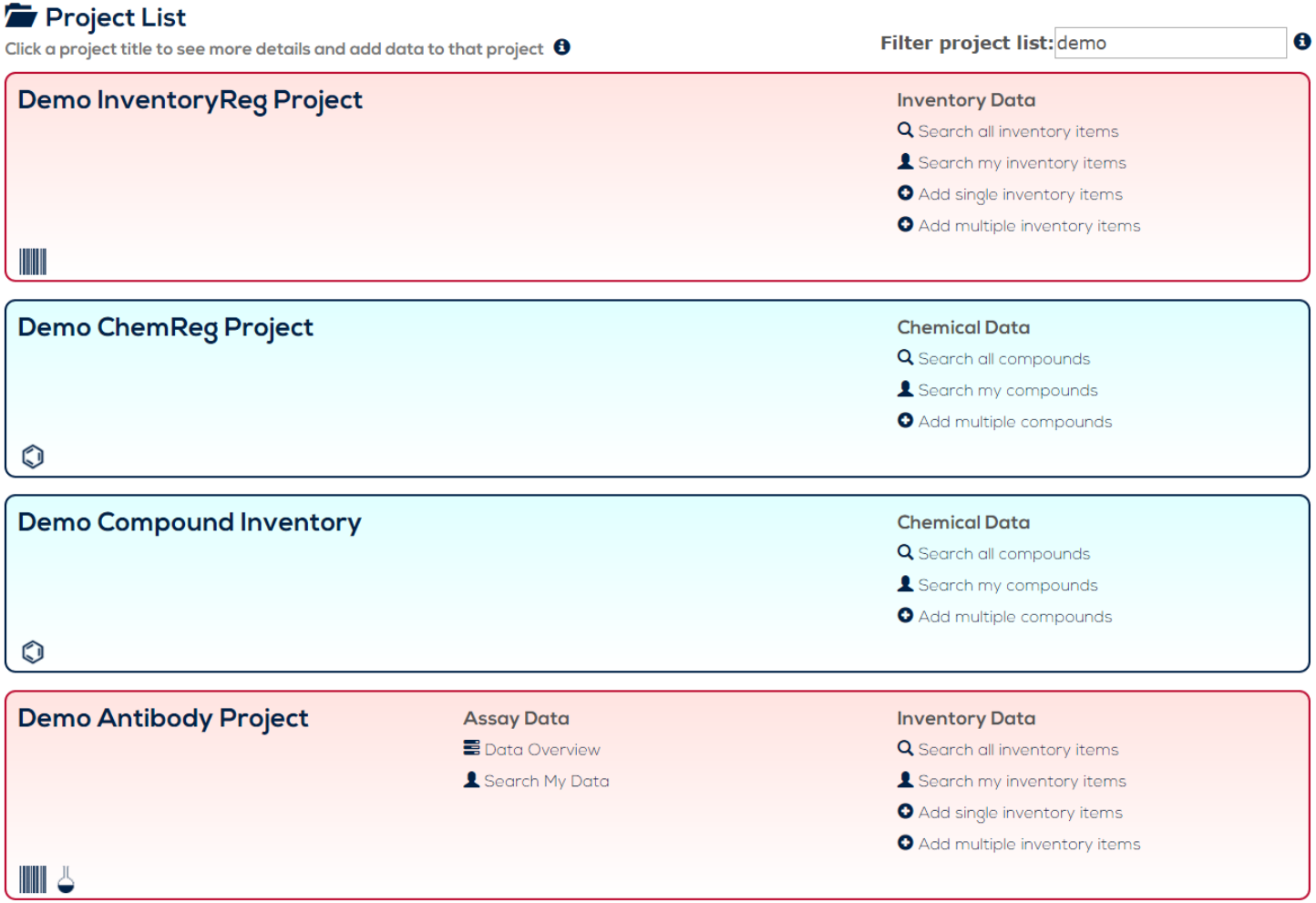
HELMHOLTZ GEMEINSCHAFT

Open Science

UNIVERSITY OF OXFORD

ChemBio Hub

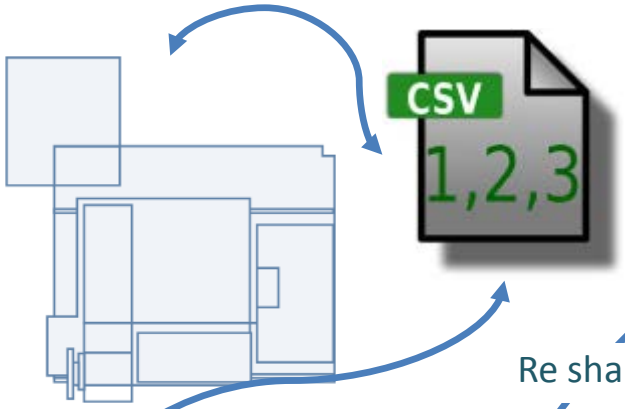
Not just for chemistry…



HELMHOLTZ GEMEINSCHAFT

Open Science

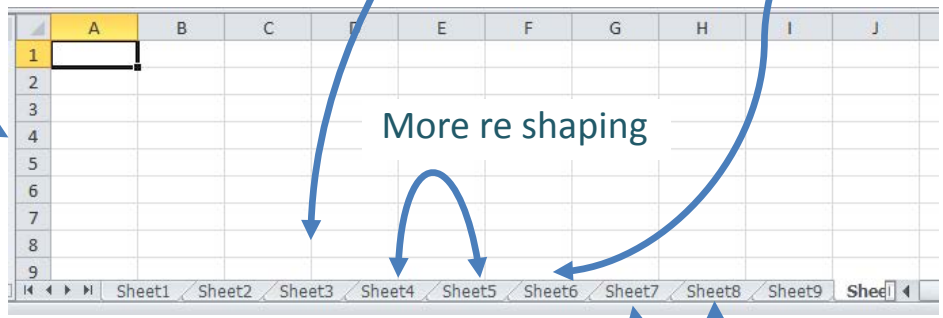
Example Alpha Screen data flow



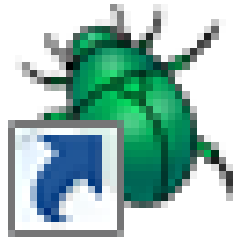
Raw input

Re shaping

Dose-response curve fitting +
Data cleansing



More re shaping

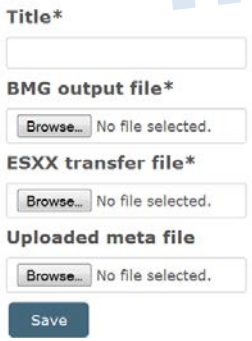


Saved to ELN

ChemBio Hub

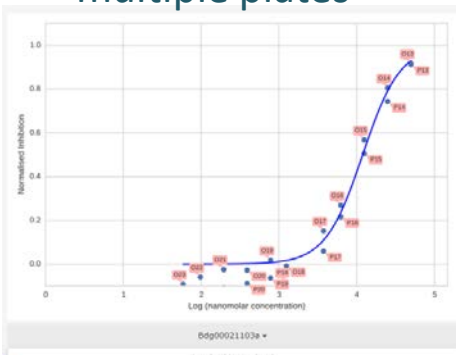
The solution: ChemBio Crunch

1) Upload raw data



2) Validate plates for systematic errors

3) Calculate IC_{50} for multiple plates



4) Mark poor fits where observed

	A	B	C	D	E	F	graph
1	plate	coupound	logIC50	ic50 (nM)	system_comments	user	
2	2	BDG00021060a	4.25927396	18166.6127	Low total inhibition, values could be inaccurate	no	
3	2	BDG00021344a	4.8345796	68324.9943	Low total inhibition, values could be inaccurate	no	
4	2	BDG00019351a	4.41049146	25733.0618	Low total inhibition, values could be inaccurate	no	

5) Export and deposit in ELN (Comments automatically generated where IC_{50} may be inaccurate)

ChemBio Hub

The technology stack

What you see

What that runs on

What we use to deploy and test it

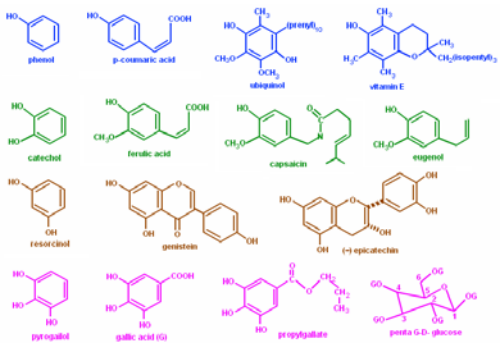
HELMHOLTZ GEMEINSCHAFT

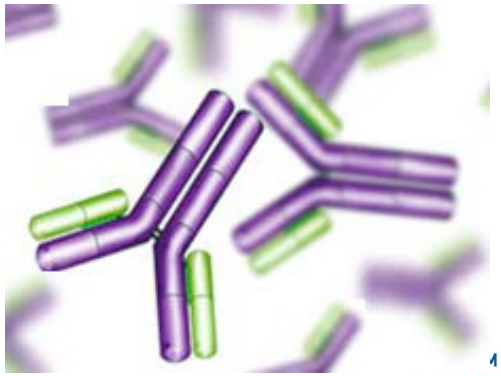
Open Science

Who would use it?

Anyone who manages an inventory through a spreadsheet

- Compounds
- Antibodies
- Cells
- Plasmids
- Lab orders
- Mice
- Unicorns





UNIVERSITY OF OXFORD

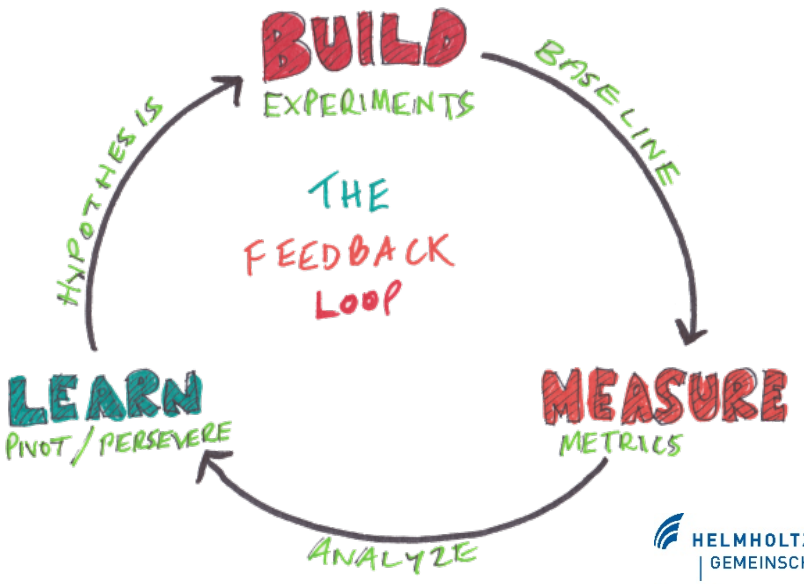
hemBi Hub

ChemBio Hub

Why use it?

- Online
- Easy to use
- Search for things instantly
- Reduce duplication
- Free

- Academic project
- Evolving project
- Me!



HELMHOLTZ GEMEINSCHAFT

Open Science

Weaknesses of an evolving project

– Innate resistance to IT solutions

– System not quite up to scratch

– Unclear message

– Concerns about future

– Group engagement

Advantages of a data management system

For the PI -

Avoid the Post-Doc leaving panic

At the Bench -

Quickly search your data

Across the Department -

Save money – Sharing of resources can reduce departmental spending

Engagement

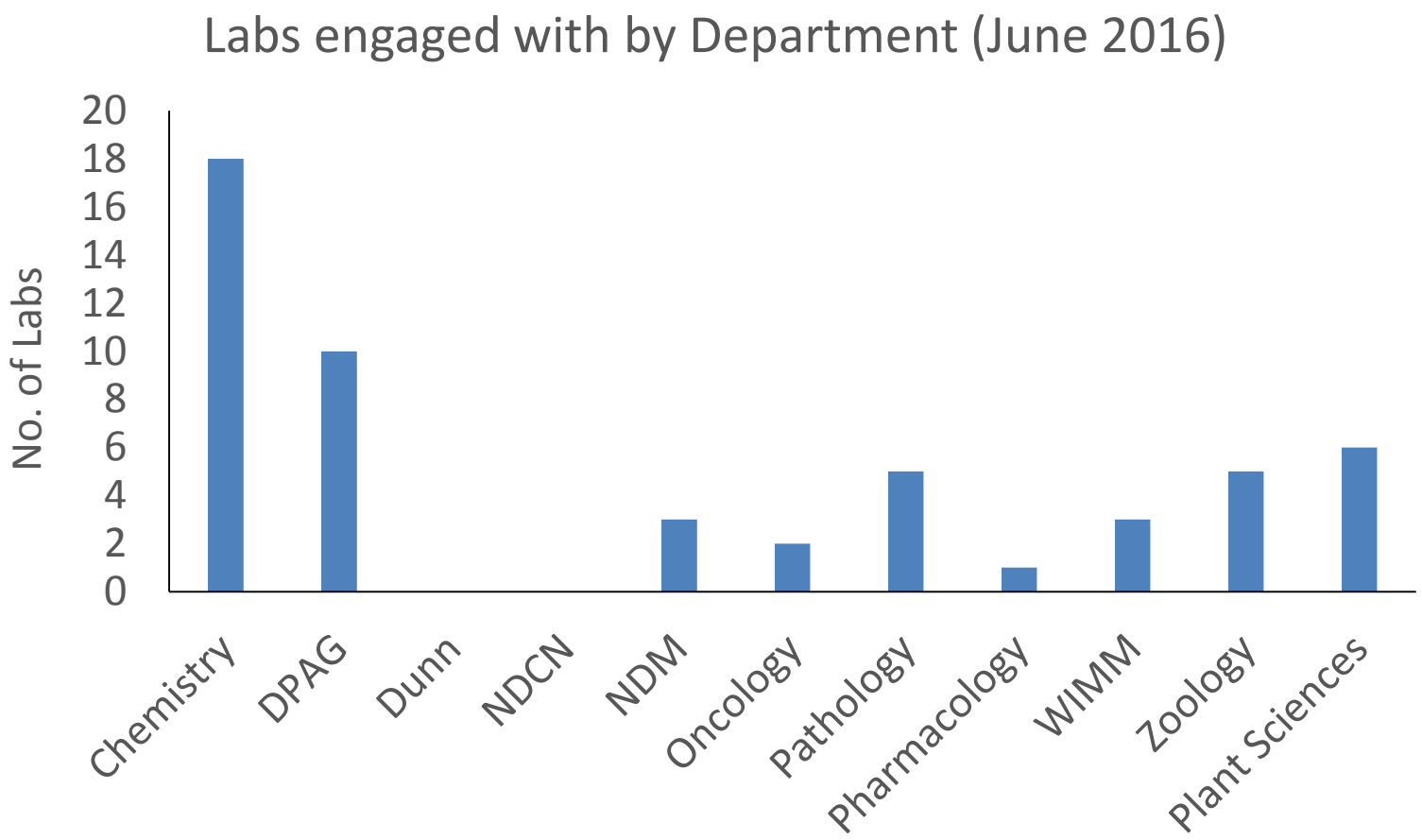
Engaged with 80 + labs across 10 departments:

- Chemistry
- Dunn school
- WIMM
- Pharmacology
- DPAG
- Plant Sciences
- Zoology
- Biochemistry
- Oncology
- TDI / SGC

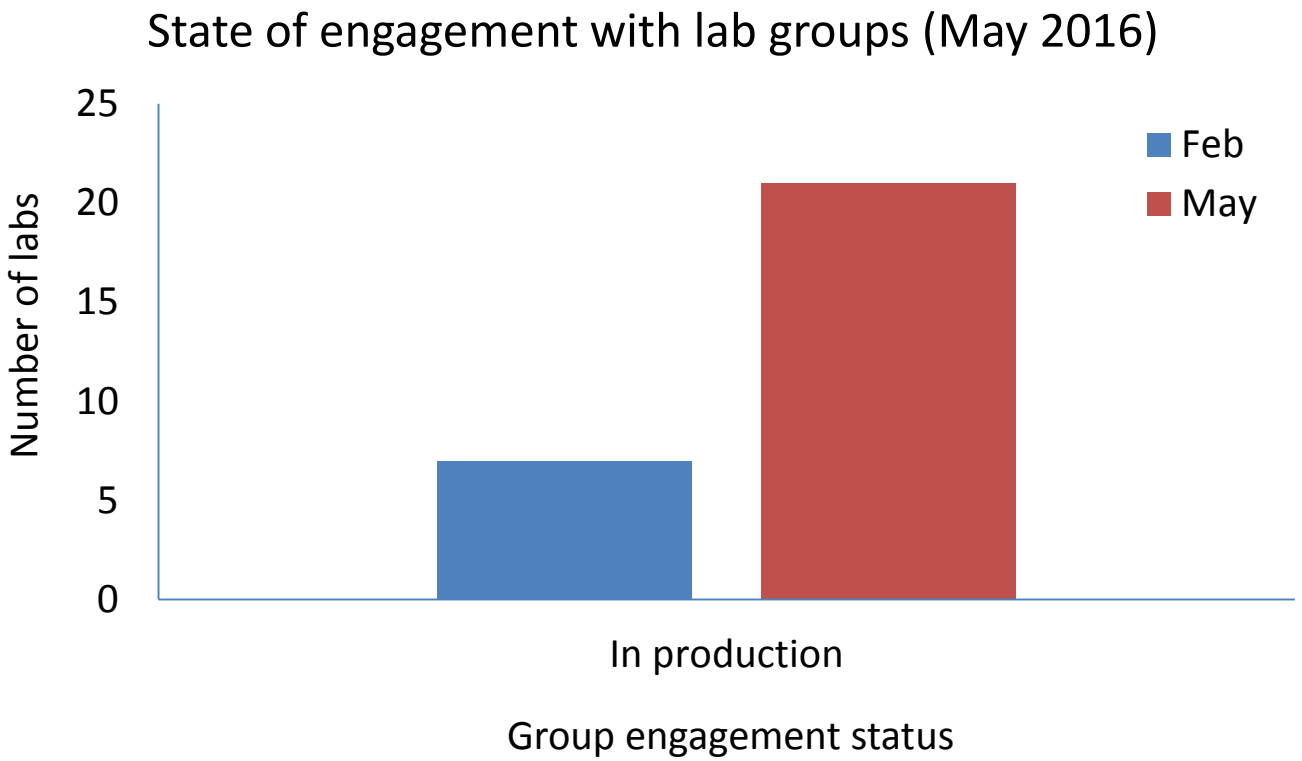
Engagement tools

- Emails / phone calls etc.
- Friends of friends
- ‘Dropping in’
- Departmental meetings
- ChemBio Hub symposium
- Trade stands

Engagement over time



Engagement over time



ChemiReg is in demand

- Clear systemic need for an Inventory manager
- People don’t realise they have a problem
- IT support is not working as it should
- People bemused by computers

Clear University requirement for IT solutions

- Dunn School of Pathology stores inventory
- Chemistry departmental reagent inventory
- Chemistry Mass spec compound registration system
- Plasmid database system

Implementing change – Things to look out for

- Politics
- Novel solutions require an open mind
- Who is the lab authority?
- Lab morale can play a key factor

Things to keep on top of:

- Managing expectations is key
- You can’t please everyone
- User aftercare
- Herding cats

ChemBio Hub

ChemBio Hub - capturing and sharing chemical biology information and knowhow

http://chembiohub.ox.ac.uk

Brian Marsden – Principal Investigator
Karen Porter – Project manager
Michael O'Hagan – KE coordinator
Andy Stretton – Developer
Paul Barrett – Developer
Adam Hendry – Project scientist

Prof. Alastair Buchan
Prof. Chris Schofield

Nuffield Department of Medicine
Medical Sciences Division

HELMHOLTZ GEMEINSCHAFT
Open Science

UNIVERSITY OF OXFORD

HIGHER EDUCATION FUNDING COUNCIL FOR ENGLAND hefce

Supported by wellcome trust

The John Fell OUP Research Fund