# NMR Data Management

UKMRM Meeting 2016 -Oxford

### **Online survey – Academia vs Industry**

Academia – 27 Responses (87 %)

Industry – 4 Responses (13 %)

Does your NMR data store provide a search facility?







Does your institution use ELN (Electronic Lab Notebook)?



There has been no driving force to improve scientific data management in academia

### **Funder's Open Access Data Policies**

- All RCUK (EPSRC, STFC, BBSRC, etc.), EC (H2020), the Royal Society, the Wellcome Trust
- Mainly apply to published results
- Data Management Plan (DMP):
  - Describes the files that will be generated (formats, volume), how they will be stored, how compliance will be achieved, what will be shared; where data will be archived
- Data:
  - o should be made publicly available
  - should be retained at least <10 years> from last date of access
  - include a statement on how to access supporting data (where? On which conditions?)
- Costs:
  - in most cases they can be included in grant applications
- **FASTR** (Fair Access to Science and Technology Research Act of 2015)
  - Bill introduced to US senate last year
  - In summary every federal agency that spends more than \$100 million on grants for research would be required to adopt an open access policy.

# **Online survey – Open Access Data Policies**

#### How do you deal with the Open Access Data Policies?



- I leave this problem to PIs, institution management or IT services.
- I pro-actively try to change the way of archiving the NMR data to accommodate the policy.
- I have no idea what this is.
- I have been forced to accommodate the policy.

#### Summary

- We know about Open Access Data Policies .
- We mostly see it as another admin chore sent on our heads from above.
- We mostly think that it's not our business and we let other people to deal with that.

#### Let's try to see it from another perspective.

### **Golden Age of NMR Spectroscopy Workflow**



# **Online survey – Current NMR dataflow**

#### What type of chemistry does your lab support ?



Average NMR lab Number of instruments : 6 Number of sample changers : 4 Number of users : ~ 150 Number of experiments per month : ~ 4400



# **Online survey – Current NMR dataflow**

#### Where do you archive NMR data in your lab?



- Obvious trend towards Central File Storage systems
- Very likely impact of Open Access Data policies



#### More important is how we store and access the data

# **Online survey – Current NMR dataflow**

#### How do you archive NMR data in your lab?



- Archiving script can be seen unsecure by people maintaining central file store.
- It seems that no one in UK academia uses commercial database for NMR spectra like ACD/Labs Spectrus DB, Mnova DB plugin or Magic Angle LOGS.

#### Does your NMR data store provide a search facility?



## **NOMAD – NMR Online Management And Datastore**



- Web based "cloud computing" system that provides a resilient searchable data-store and centralised control of automation for Bruker NMR instruments.
- Current version 1.2 (prototype) 6 Bruker instruments with sample changers, 260 research users, 350 users in teaching labs
- Dataflow 6 000 to 10 000 Experiments (20 -30 GB) per month

#### **NOMAD – Submission Portal**

Always <b>REFRESH</b> this page before you decide/act, the list shown here is <b>00:00:34 (hh:mm:ss)</b> old and thus might be outdated           Submit Selected Jobs         BOOK NEW JOB         Cancel Selected Jobs																	
Holder	NMR Machi	ine l	Jsername	PIGrou	p Solvent	Sample Ide	entifier ExpNo	Experiment	Parameters	Title	Print	Email	Date	ті	me Sta	tus	Select
8	Alec		tl12	tl	CDCI3	09142015-8-	-tl-tl12-A 11	c13_deptq.c.and	ns,256	Demo		Y	2015-09-	14 12	:42 BOO	KED	
8	Alec		tl12	tl	CDCl3	09142015-8-	-tl-tl12-A 10	proton.c.and		Demo		Y	2015-09-	14 12	:42 BOO	KED	
Holder	NMR Machi	ine l	Jsername	PIGrou	p Solvent	Sample Ide	entifier ExpNo	Experiment	Parameters	Title	Print	Email	Date	Ті	me Sta	tus	Select
		Ale	c Felix	Hector	Marcus Noah	Robbin											
Holder N	IMR Machine U	Ale	c Felix PIGroup	Hector	Marcus Noah Automation - Running - Sample Identifier	Robbin Busy until : ExpNo	Mon 13:05 - Day Expe Experiment	eriments : 00:12 - Night Exp Parameters	periments : 00:00 - I Title	<b>1AX ALLOW</b> Pri	/ANCE : 00 nt Email	):15 Date	Time	Duration	Status	Holder	Sele
Holder Ni 11	IMR Machine L Marcus	Ale Username tb275	c Felix PIGroup doh	Hector Solvent CDCl3	Marcus Noah Automation - Running - Sample Identifier 09142015-11-doh-tb275-	Robbin Busy until : ExpNo M 10	Mon 13:05 - Day Expe Experiment ghsqc-ed.a.and	eriments : 00:12 - Night Exp Parameters	eriments : 00:00 - I Title tb150909	MAX ALLOW	YANCE : OC nt Email Y	D:15 Date 2015-09-14	<b>Time</b> 12:39	<b>Duration</b> 00:10:55	SUBMITTED	Holder 11	Selec
Holder NI 11 15	IMR Machine L Marcus Marcus	Ale Username tb275 na10	c Felix PIGroup doh doh	Hector Solvent CDCl3	Marcus Noah Automation - Running - Sample Identifier 09142015-11-doh-tb275- 09142015-15-doh-na10-f	Robbin Busy until : ExpNo M 10 12	Mon 13:05 - Day Expe Experiment ghsqc-ed.a.and f19cpd.a.and	eriments : 00:12 - Night Exp Parameters	Title tb 150909 VAphenylethanol	MAX ALLOW Pri	YANCE : OC nt Email Y Y	Date 2015-09-14 2015-09-14	<b>Time</b> 12:39 12:35	<b>Duration</b> 00:10:55 00:01:54	Status SUBMITTED RUNNING	Holder 11 15	Sele
Holder N 111 15 15	IMR Machine L Marcus Marcus Marcus	Ale Username tb275 na10 na10	c Feix	Hector Solvent CDCl3 0 CDCl3 1	Marcus         Noah           Automation - Running -           Sample Identifier           09142015-11-doh-tb275-           09142015-15-doh-na10-1           09142015-15-doh-na10-1           09142015-15-doh-na10-1	Robbin Busy until : ExpNo M 10 M 12 M 11	Mon 13:05 - Day Expe Experiment ghsqc-ed.a.and f19cpd.a.and pro19Fdec.a.and	eriments : 00:12 - Night Exp Parameters 1 1	tb150909 VAphenylethanol VAphenylethanol	MAX ALLOW Pri	VANCE : OC nt Email Y Y Y Y	Date 2015-09-14 2015-09-14 2015-09-14	Time 12:39 12:35 12:35	Duration 00:10:55 00:01:54 00:00:30	Status SUBMITTED RUNNING FINISHED	Holder 11 15 15	Sele
Holder Ni 11 11 15 15 15 15	IMR Machine L Marcus Marcus Marcus Marcus Marcus	Username tb275 na10 na10	c Felix PIGroup doh doh doh doh	Hector Solvent CDCI3 0 CDCI3 1 CDCI3 1	Marcus         Noah           Automation - Running -           Sample Identifier           09142015-11-doh-tb275-           09142015-15-doh-na10-1           09142015-15-doh-na10-1           09142015-15-doh-na10-1           09142015-15-doh-na10-1	Robbin Busy until : ExpNo M 10 M 12 M 12 M 11 M 10	Mon 13:05 - Day Expe Experiment ghsqc-ed.a.and f19cpd.a.and pro19Fdec.a.and proton_WIDE.a.and	eriments : 00:12 - Night Exp Parameters r r	title tb150909 VAphenylethanol VAphenylethanol VAphenylethanol	MAX ALLOW Pri	ANCE : OC nt Email Y Y Y Y	Date 2015-09-14 2015-09-14 2015-09-14 2015-09-14	Time 12:39 12:35 12:35 12:35	Duration 00:10:55 00:01:54 00:00:30 00:00:51	Status SUBMITTED RUNNING FINISHED FINISHED	Holder 11 15 15	Sele

	Username	tl12								
	Password	•••••			<ul> <li>Image: A set of the set of the</li></ul>					
	Name	Tomas Lebl			]					
	PI Group	tl								
	Available Machines	Alec 👻								
	No of Samples	2 *								
Holder	Sample Identifier	Solvent	Experiment	Parameters	Print	Email	Title	Duration	Total (D)	More
Holder 21	Sample Identifier	Solvent CDCl3 •	Experiment	Parameters	Print	Email	Title Demo 1	<b>Duration</b> 00:00:34	Total (D) 00:02:47	More Add Del
Holder 21	Sample Identifier 09142015-21-tl-tl12-A	Solvent CDCl3 •	Experiment proton.c.and • • p31cpd.c.and •	Parameters	Print	Email V	Title Demo 1 Demo 1	Duration 00:00:34 00:02:13	Total (D) 00:02:47	More Add Del
<b>Holder</b> 21 22	Sample Identifier 09142015-21-tl-tl12-A 09142015-22-tl-tl12-A	Solvent CDCI3 • DMSO •	Experiment proton.c.and • • p31cpd.c.and • proton.c.and • •	Parameters	Print	Email	Title Demo 1 Demo 1	Duration           00:00:34           00:02:13           00:48:33	Total (D) 00:02:47 00:48:33	More Add Del
Holder           21           22	Sample Identifier           09142015-21-tl-tl12-A           09142015-22-tl-tl12-A	Solvent CDCI3 • DMSO • Continue	Experiment proton.c.and • p31cpd.c.and • Start Over	Parameters	Print	Email V V t	Title Demo 1 Demo 1	Duration           00:00:34           00:02:13           00:48:33	Total (D)           00:02:47           00:48:33	More Add Del Add

#### **NOMAD – Admin Portal**

NOMAD v1.2.1			👗 User 🔻	👺 Manual 👻 🔛 Accounting	🏛 Teaching ▼ 🛛 📢 Feedbac	k 📴 Logout 👗 Tomas Lebi
🚯 Dashboard						
🔮 Admins & Users 🤟	Dashboard					
Litt Usage & Activity						
Machines	Alec Automation - Running Busy until - Tue 12:52	Felix NOT AVAILABLE	Hector	NOT AVAILABLE	Marcus	Automation - Running Busy until : No Jobel
幸 Experiments Settings 〈	Day Experiments : 00:00 Night Experiments : 00:00					Day Experiments : 00:00 Night Experiments : 00:00
Experiments Records	MAX ALLOWANCE : 00:20					MAX ALLOWANCE : 00:15
🛓 User Portal 🤇	Slots Available: 42 / 60	Slots Available: 60 / 60 Go to Oueue	Slots Available: 19 / 24 Go to Oueue		Slots Available: 98 / 120	
🖕 Manual Portal 🤟	TURN OFF Reset				Go to Queue	TURN OFF Reset
🏛 Teaching Portal 🤟	Noah Automation - Running	Robbin Automation - Running				
System Developers	Busy until : No Jobs! Day Experiments : 00:00	Busy until : No Jobs! Day Experiments : 00:00				
	Max ALLOWANCE : 00:15	MAX ALLOWANCE : 01:15				
	Slots Available: 47 / 60 Go to Queue TURN OFF Reset	Slots Available: 56 / 60 Go to Queue TURN OFF Reset				



#### **NOMAD – Admin Portal - user & experiment tables**

User Configuration								Ind Users Stats			
ID		A	UTO					Total Users (Resear	ch + Teaching)	1311	
Username					Total Active Users (R	Total Active Users (Research + Teaching)					
Real Name						Total Active Group	s	32			
	(Ophy if user is a	ast listed in the university lea			Very Active Users (last 30 days)	Research + Teaching)	136				
Special User		for instea in the university log						Total Research Use	ars	553	
Manual User	U (Only if user is a	allowed to do manual experim	ients)								
Inactive User	🔲 (User will not sh	now up on drop-down boxes)						Active Research User	S	264	
Category	*							Inactive Research Us	ers	289	
PI Group	ads 🔻							Special Research Use	rs	60	
	Save/Update	Clear						Manual Research Use	rs	19	
								Total Teaching Use	rs	757	
								Active Teaching User	5	351	
						Quick Filt	er				
Inclu	ude Inactive		▼ Category				pk	▼ Group	Days ≥ 90	120 ≤ Days	
User ID	Username	Real Name	Special	Manual	Inactive	Category	Group	Actions	Last Activity	Days of Inactivity (90d)	
60	bac8	Brian Chalmers	YES		Active	PostDoc	pk	Load Delete	2016-06-09 09:29:07.0	12	
1304	bjml	Bethany Lawson			Active	CDT	pk	Load Delete	2016-04-22 14:31:58.0	60	
776	cnb3	Claire Brodie			Active	MSci	pk	Load Delete	2016-04-05 12:07:22.0	77	
106	ljt8	Laurence Taylor			Active	PhD	pk	Load Delete	2016-06-10 16:39:56.0	11	
113	pk7	Petr Kilian			Active	PI	pk	Load Delete	2016-03-25 15:49:29.0	88	
53	pn46	Phillip Nejman			Active	PhD	pk	Load Delete	2016-06-03 15:10:41.0	18	
User ID	Username	Real Name	Special	Manual	Inactive	Category	Group	Actions	Last Activity	Days of Inactivity (90d)	

ALL Alec Felix Hector Marcus Noah Robbin Not Assigned Exp-ID Experiment Name Description Parameters Hidden Composite Machines Actions Portals 1H Observe ns = 8 ; d1 = 1 ; expt = 1m 9s ; ds = 2 ; td1 = 1 NMRF 1 proton.a.and ---R & T Load Delete 2 proton\_WIDE.a.and 1H Observe with Expanded Sweep Width [from 15 to -25 ppm] ns = 8; d1 = 1; expt = 51s; ds = 2; td1 = 1 NMRF Load Delete R-only ns = 8 ; d1 = 30 ; expt = 5m 59s ; ds = 2 ; td1 = 1 3 proton\_Quant.a.and 1H Observe - Quantitative ---NMRF Load Delete R-only 4 solventsup.a.and 1H Observe with Solvent Suppression ns = 8 ; d1 = 4 ; expt = 1m 49s ; ds = 4 ; td1 = 1 NMRF ------R-only Load Delete

### **NOMAD – Admin Portal - Accounting**

PI Gro	oup: dp 🗸 🔲 I	nactive Da	ite From:	01/09/20	015	1	Date	To: 14/09/2	2015		Viev	v Duratio	on Com	nplete Cl	ear					
age Data	(GB) Loading Status:							0	`omplete											
									iompioco											
			Alec			Felix Hector						Marcus			Noah		Robbin			
	Name	ЕхрТ	GB	Cost	ExpT	GB	Cost	ExpT	GB	Cost	ЕхрТ	GB	Cost	ЕхрТ	GB	Cost	ЕхрТ	GB	Cost	Total Co
Auto		-	-	-	-	-	-	-	-	-	11:17:21	0.066	£46.80	08:55:11	0.057	£37.12	-	-	-	£83.91
Manual		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	dvb - David Van Brussel	N/A	-	-	N/A	-	-	N/A	-	-	N/A	0.066	£46.80	N/A	0.057	£37.12	N/A	-	-	£83.91
Auto		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Manual		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	dp12 - Douglas Philp	N/A	-	-	N/A	-	-	N/A	-	-	N/A	-	-	N/A	-	-	N/A	-	-	-
Auto		00:40:40	0.009	£6.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	£6.00
Manual		-	-	-	2h	0.066	£12.65	-	-	-	-	-	-	-	-	-	-	-	-	£12.65
	jr74 - Josh Richards	N/A	0.009	£6.00	N/A	0.066	£12.65	N/A	-	-	N/A	-	-	N/A	-	-	N/A	-	-	£18.65
Auto		00:58:23	0.015	£8.65	00:31:52	0.006	£3.08	16:50:46	0.100	£170.97	00:12:57	0.002	£0.92	00:47:42	0.007	£3.36	00:55:56	0.006	£2.48	£189.4
Manual		-	-	-	-	-	-	1d - 10h	0.035	£220.88	-	-	-	-	-	-	-	-	-	£220.8
	lm216 - Leonardo Maugeri	N/A	0.015	£8.65	N/A	0.006	£3.08	N/A	0.135	£391.84	N/A	0.002	£0.92	N/A	0.007	£3.36	N/A	0.006	£2.48	£410.3
Auto		00:33:02	0.004	£4.78	00:07:06	0.002	£0.70	-	-	-	-	-	-	-	-	-	-	-	-	£5.48
Manual	door to war	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A	tk235 - Tamara Kosikova	N/A	0.004	£4.78	N/A	0.002	£0.70	N/A	-	-	N/A	-	-	N/A	-	-	N/A	-	-	£5.48
Auto		02:12:05	0.028	£19.42	00:38:58	0.008	£3./8	16:50:46	0.100	£1/0.97	11:30:18	0.068	£4/./1	09:42:53	0.065	£40.48	00:55:56	0.006	£2.48	£284.8
manuar	τοτοι	-	-		211	0.000	£12.05	10 - 10h	0.035	£220.88	-	-	-	-	-	C 40 49	-	-	-	£233.3
	Namo	IN/A	0.028	£19.42	IN/A	0.074	£10.43	N/A	U.135	£391.84	N/A	U.U08	£4/./1	N/A	Noah	£40.48	N/A	0.006	£2.48	£318.3
	name		Alec			i enx			nector			narcus			NUali		P		_	TUTALCO

#### **NOMAD – User Search Portal**

NOMAD v1.2 -	NMR Search P	ortal									Q User 📂 Ma	anual 👻 🕞 Logout	🛔 Tomas Lebl	
PI	Group jhn			Ŧ	User b	n22 - Brunello Nardone	Ŧ	Date From			Date To			
	NMRall			¥	Sample ID			Title	Patellamide D		Solventall		Ŧ	
Expe	erimentall						Ŧ	Parameters			Inactive			
						Q Sear	ch S							
	Always REFRESH this page before you decide/act, the list shown here is 00:25:23 (hh:mm:ss) old and thus might be outdated													
							Search Returned 3 Re	sults						
							1							
ID	Username	PlGroup	Holder	Solvent	Sample Ider	ntifier Exp No	Experiment	Parame	ters Title	Date	Bucket	Actio	IS	
402559	bn22	jhn	4	CDCI3	06202016-4-jhr	-bn22-H 12	ghsqc-ed.c.and	ns,2	Patellamide D	2016-06-20	H-20160211-105222	Download	View	
402558	bn22	jhn	4	CDCI3	06202016-4-jhr	i-bn22-H 11	gtocsy.a.and	ns,2	Patellamide D	2016-06-20	H-20160211-105222	Download	View	
402557	bn22	jhn	4	CDCI3	06202016-4-jhr	i-bn22-H 10	proton.c.and	ns,8	Patellamide D	2016-06-20	H-20160211-105222	Download	View	

# **NOMAD – PURE link & Open-Data**

- NOMAD stores the data securely and provides instant search facility but the data are not publicly available.
- Users have to download selected data and upload them to PURE (Research Information Management System).
- It's not clear what should be uploaded to PURE and capacity is limited.



### **Open Access Data Policies – Good or Bad or Ugly**

#### • Nowadays NMR spectroscopy produces vast quantity of data.

Is that data sufficiently utilised and shared?

Pictures of <sup>1</sup>H and <sup>13</sup>C spectra in Supporting Information



- Why academia does not use commercial NMR databases?
- Why NMR spectroscopy does not have an equivalent of Protein Data Bank or Cambridge Crystallographic Data Centre?

Likely there has been no driving force.

**Could Open Access Data Policies become that driving force?** 

Maybe

### **NOMAD - Current Status and Future Plan**

#### **Current status**

- Team of 6 people working part time
- We have setup infrastructure and processes for software development and production
- IAA funding available until March 2017

#### Short term future plan

- "PURE link" add-on function and "Open Data Portal"
- NOMAD 2.0 modular architecture, separate database and control function

#### Long term future plan



# **Discussion**

- Do you like the concept of NOMAD system? Any comments ?
- If yes how to propel further development and distribution?
  - a) University of St Andrews spin-off
  - b) EPSRC or other funding