



July 20, 2018

Conversations with Chemists Redux

Preliminary Results from an Interview-Based Study on the Information Needs
and Habits of Chemistry Faculty

David Flaxbart

Lydia Fletcher

University of Texas at Austin



Background



- ▶ Purpose: Inform next generation of research support services
- ▶ Ithaka S+R Research Support Services – model
- ▶ UT Austin's Ithaka replication projects
- ▶ UK Chemistry Faculty study, 2013
- ▶ "Redux" – refers to local 2003 interview project
- ▶ Final chemistry report expected in fall 2018



Methods



- ▶ Develop “semi-structured” interview questionnaire
- ▶ IRB approval
- ▶ Target: 15 interviews (Feb-Apr 2018)
- ▶ One-on-one interviews conducted in faculty offices
- ▶ Recorded and transcribed; average 30 minutes each
- ▶ Analyzed and annotated transcripts




Demographics



- ▶ Chemistry Department Faculty: 28 members
- ▶ 19 professors invited via email: 15 accepted, 4 did not respond
- ▶ Rank: 5 Assistant, 2 Associate, 8 Full
- ▶ 4 Divisions: Organic (3), Physical (5), Analytical (5), Inorganic (2)
 - ▶ Biochemistry in separate department since 2013



Major Themes of Questions

- 
- ▶ Research Focus; Funding; Collaborations
 - ▶ Instrumentation; Data Output; Recordkeeping
 - ▶ "Primary" and "Secondary" Information Sources: Discovery, Challenges, Management
 - ▶ External Data: Need, Discovery
 - ▶ Keeping up with trends and new research
 - ▶ Publication practices; Open publication and data sharing
 - ▶ Research Data Management: archiving, preservation
 - ▶ "Magic Wand"



In the Lab

- ▶ **Instrumentation:**

- ▶ MS; NMR; IR; XPS; EPR; Microscopy; Lasers; X-ray crystallography; Chromatography; Computers (primary)
- ▶ In-lab vs Shared Facilities

- ▶ **Data Handling:**

- ▶ File outputs and sizes of every description
- ▶ Storage and sharing

- ▶ **Lab Notebooks**

- ▶ All paper; no ELNs
- ▶ ELN awareness but no uptake: only 1 indicated past consideration



Literature Formats

- ▶ **Journals are #1**
 - ▶ 13 of 13 who answered the question
 - ▶ No surprise here
- ▶ **Other “primary” formats mentioned as sources of ideas and information:**
 - ▶ Personal communications (4)
 - ▶ Conferences (attendance and networking, not published) (3)
 - ▶ Patents (2)
 - ▶ Dissertations (1)
 - ▶ Other sources: CSD (1)
- ▶ **Ambivalent attitudes towards books/monographs**
- ▶ **Need for external data is minimal**



Discovery Tools of Choice

	Top Choice	Secondary
Web of Science	5	2
Google	3	2
SciFinder	2	5
Google Scholar	2	2
PubMed/MEDLINE	1	2
Reaxys	0	1
Inspec	0	1
USPTO	0	1
None specified	2	

Publishing Choices

- ▶ **Stated preferences for non-profits, esp. ACS and RSC portfolios**
 - ▶ Decisions based on reputation, audience, likelihood of acceptance
 - ▶ Reviewing choices reflected too
 - ▶ 2 noted dislike of “cascade” model of publisher portfolios
- ▶ **Web of Science analysis of 15 interviewees’ articles, 2013-18:**
 - ▶ 66 journals with 2+ articles: ACS or RSC = 52%

ACS	21
RSC	13
WILEY	10
ELSEVIER	9
AIP	3
NATURE GROUP	3
OTHER	7



Beyond Publication

- ▶ Tradition trumps trendiness
 - ▶ Speaking at conferences or seminars (9)
 - ▶ Only 1 noted social media; 3 indicated “No social media”
 - ▶ Research group web sites are valued platform



Open publication

- ▶ Open is good, “but not on my dime”
- ▶ **Bias against Gold OA and APCs**
 - ▶ Hostility or indifference
 - ▶ 2 indicated they post published versions on personal web sites
 - ▶ “Publication pachinko” – journal cascades derided
- ▶ **WOS analysis of 579 articles authored by interviewees (2013-18):**
 - ▶ 193 have open versions (33%)
 - ▶ 111 of 193 (58%) are “gold or bronze” status; 82 (42%) are “green”
 - ▶ Most green OA is probably compliance-related deposit, not deliberate
 - ▶ Only 1 indicated deliberate green deposit (arXiv)

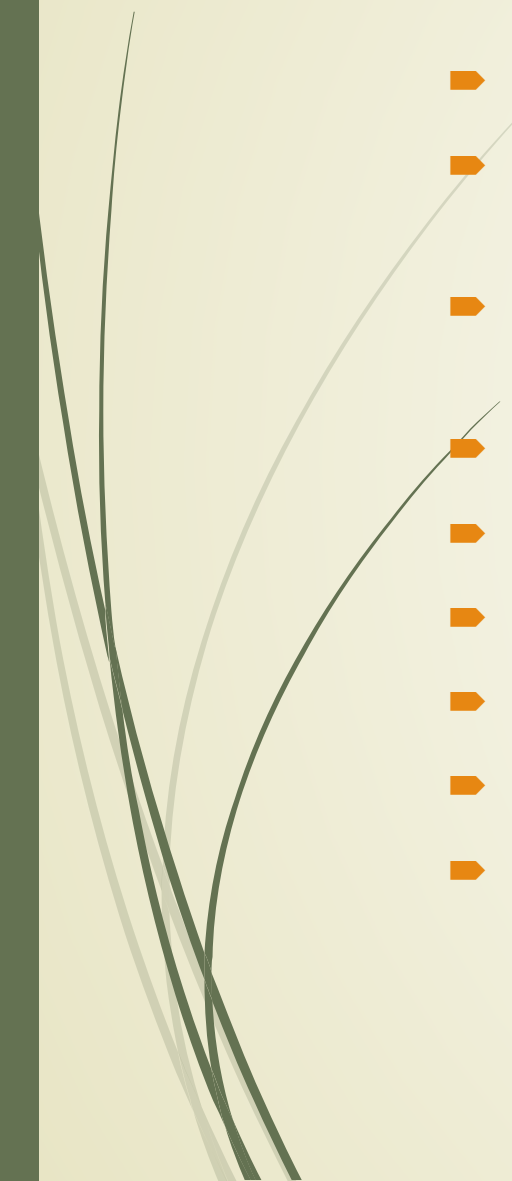


Quoted on Open publication

- ▶ "I'll send it somewhere else, thank you, and save my money."
- ▶ "It's just not an issue at all. My audience are people at places like UT. I'm not worried about people who don't have access to those journals."
- ▶ "I think [paying APCs] is a waste of money.... I am opposed to paying a fee if there's something that I can upload to another site myself or if it just has to wait a year, then so be it."
- ▶ "The presence or absence of a journal's open access policy basically has no bearing on whether or not we would choose to submit...to that journal."
- ▶ "I'm a firm believer in Open Access, but I guess not to the extent of paying a thousand bucks per article."
- ▶ "Personally I don't think authors should have to pay to publisher their stuff. Publishers make plenty of money."
- ▶ "I think it's a neat idea that you can just let anyone have access to your results."




Data – It's complicated...

- ▶ Understanding of data concepts varies widely
 - ▶ Ad-hoc procedures and solutions: Box backup most common for local storage and sharing within group
 - ▶ Headaches everywhere: Hardware, software, file formats and sizes, networking, backups
 - ▶ PIs: hands-on vs hands-off; Varying levels of confidence
 - ▶ Compliance vs practical needs
 - ▶ Reliance on published supporting information in journals vs repositories
 - ▶ Sharing on request is acceptable; open deposit generally not
 - ▶ Need for long term preservation and archiving unclear to some
 - ▶ Value of raw/unpublished data not evident to some
- 



Quoted on Data Management

- ▶ "We might not do so well."
 - ▶ "I know my students have some archival data. I personally don't manage any of that. ... Who knows where the data is. It's probably on their computers."
 - ▶ "There's nothing that we do that can't be recreated if we need to."
 - ▶ "Up to now, we've kind of patched it along using these externals hard drives or what-not."
 - ▶ "I might know we have to keep it. We're doing our best. But years go by and nobody requests that data."
 - ▶ "I've been a little remiss."
 - ▶ "I need to think about that."
- 



Key takeaways

- ▶ Dependence on peer-reviewed journals is universal
- ▶ Use of other formats is low
- ▶ Information-seeking strategies vary, within limits
- ▶ External data needs are minimal
- ▶ Open is understood but not a priority; APCs are unpopular
- ▶ Keeping up with literature is very difficult; confidence varies
- ▶ Data management strategies ad-hoc, underdeveloped or nonexistent
- ▶ Little uptake of support technology (e.g. ELNs, ref mgt, data archiving)
- ▶ Low awareness of library support services (e.g. repositories)
- ▶ Approach is traditional rather than innovative, constrained by time, resources, career and disciplinary norms
- ▶ Never enough time in the day!



Potential targets for research support

- ▶ **RDM training, best practices, assistance**
 - ▶ Tailored to PI and lab requirements – one size doesn't fit all
 - ▶ Uptake will vary too – not always a priority
 - ▶ Local repositories not always the answer
- ▶ **Awareness of and help with support technology**
 - ▶ Reference managers
 - ▶ Better alerting services needed
- ▶ **Proceed with Caution**
 - ▶ Open Access attitudes aren't easily changed (\$\$)
 - ▶ "Sharing" means different things to different people
- ▶ **Focus on saving their time, not changing their world**



Questions?

- ▶ David Flaxbart flaxbart@austin.utexas.edu
 - ▶ Lydia Fletcher l.fletcher@austin.utexas.edu
- 