

Digital Preservation at Oxford and Cambridge

A collaborative research project to evaluate and provide sustainable recommendations for our digital preservation programmes

Electronic lab notebooks and digital preservation: part II

Posted on [12 September, 2018](#) by [Sarah](#)

In her [previous blog post](#) on electronic lab notebooks (ELNs), Sarah outlined a series of research questions that she wanted to pursue to see what could be preserved from an ELN. Here are some of her results.

In my last post, I had a number of questions that I wanted to answering regarding the use of ELNs at Oxford, since IT Services is currently running a pilot with [LabArchives](#).

Those questions were:

1. **Authenticity of research** – are timestamps and IP addresses retained when the ELN is exported from LabArchives?
2. **Version/revision history** – Can users export all previous versions of data? If not users, then can IT Services? Can the information on revision history be exported, even if not the data?
3. **Commenting on the ELN** – are comments on the ELN exported? Are they retained if deleted in revision history?
4. **Export** – What exactly can be exported by a user? What does it look like? What functionality do you have with the data? What is lost?

What did I find out?

I started first with looking at the IT Services' [webpage](#) on ELNs. It mentions what you can download (HTML or PDF), but it doesn't offer much more about the long-term retention of it. There's a lot of useful advice on getting started with ELNs though and how to use the notebook.

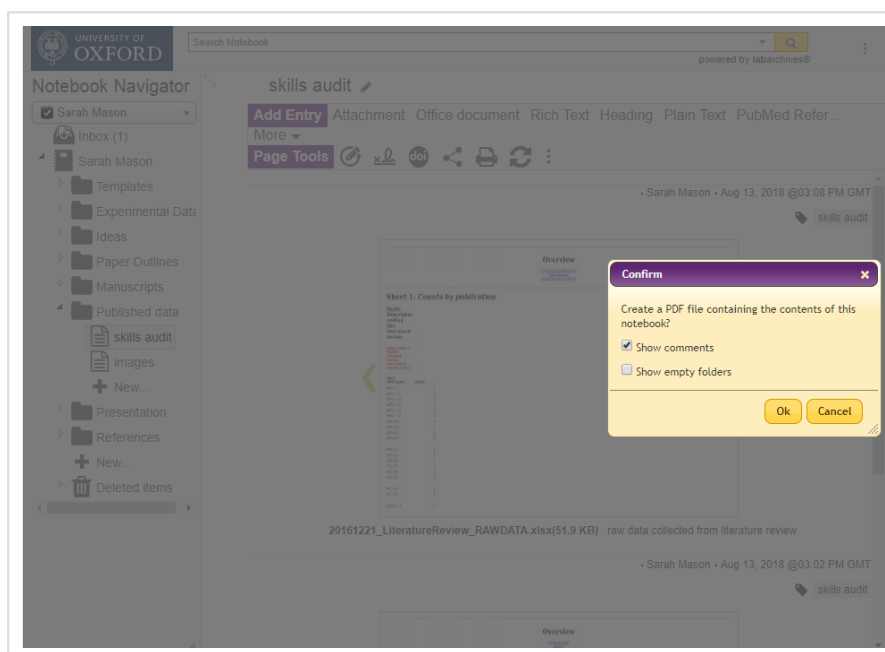
In the Professional version that staff and academics can use offers two modes of export:

- Notebook to PDF
- Offline Notebook – HTML

When you request one of these functions, LabArchives will email it to the email address associated with your work. It should happen within 60 minutes. Then you will have 24 hours to download the file. So, the question is: what do you get with each?

PDF

There are two options when you go to download your PDF: 1) including comments and 2) including empty folders.



So, this means that comments are retained in the PDF and they look something like this:



GLAMStaffEvent25-01-18byIanWallman-5247.jpg(2 MB) - [download](#)

Comments

Sarah Mason

Sep 07, 2018 @03:10 PM GMT

Nice photo. By Ian Wallman

It also means that where possible, previews of images and documents show up in the PDF. As do the latest timestamps.

What you lose is:

- previous versions and revision history
- the ability to use files – these will have to be downloaded and saved separately (but this was expected from a PDF)

What you get:

- a tidy, printable version of a lab notebook in its most recent iteration (including information on who generated the PDF and when)

Sarah Mason
Complete Notebook
PDF Version generated by
Sarah Mason (bodl2766@ox.ac.uk)
on
Sep 07, 2018 @03:11 PM GMT

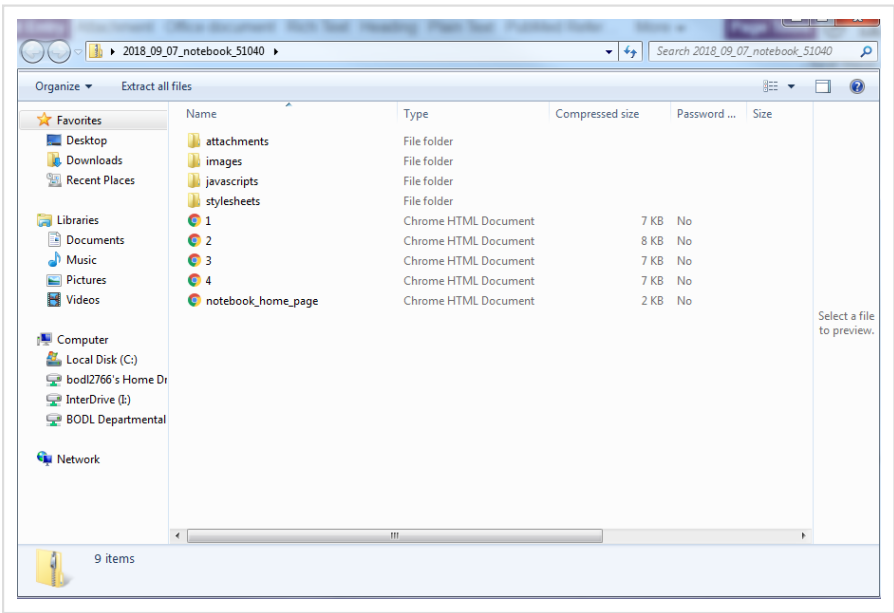
Table of Contents

Templates	2
Experimental Data	3
Ideas	4
Paper Outlines	5
Manuscripts	6
Published data	7
skills audit	7
images	8
Presentation	9
iPres 2017	9
References	10
Literature Review	10

— What the PDF cover of a lab notebook looks like.

Offline HTML version

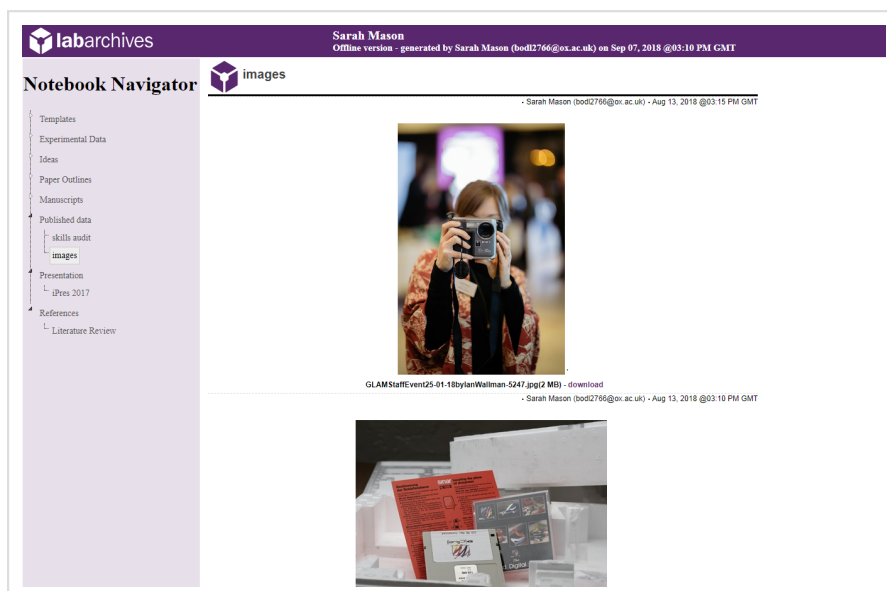
In this version, you are delivered a zip file which contains a number of folders and documents.



All of the attachments are stored under the attachments folder, both as original and thumbnails (which are just low res jpegs used by

LabArchives).

How does the HTML offline version stack up? Overall, the functionality for browsing is pretty good and latest timestamps are retained. You can also directly download the attachments on each page.



In this version, you do not get the comments. You also do not get any previous versions, only the latest files, updates and timestamps. But unlike the PDF, it is easy to navigate and the uploaded attachments can be opened, which have not been compressed or visibly changed.

I would recommend taking a copy of both versions, since each one offers some different functions. However, neither offer a comprehensive export. Still, the most recent timestamps are useful for authenticity, though checksums for files generated on upload and given you to in an HTML export in a manifest file would be even better.

Site-wide backup

Neither export option open to academics or staff allows a comprehensive version of the ELN. Something is lost in the export. But, what LabArchives does offer is an annual site-wide back up to local IT Services as part of their Enterprise agreement. That includes: all timestamps, comments and versions. The copy contains everything. This is promising, so all academics should be aware of this because they can then request a copy from IT Services. And they should be able to get a full comprehensive

backup of their ELN. This also means that IT Services is also preserving a copy of the ELNs, like LabArchives.

So, we are going to follow up with IT Services, to talk about how they will preserve and provide access to these ELN backups as part of the pilot. Many of you will have similar conversations with your own IT departments over time, as you will need to work closely with them to ensure good digital preservation practices.

And these are some of the questions you may want to consider asking when talking with your IT department about the preservation of ELNs:

- How many backups? Where are the backups stored? What mediums are being used? Are backups checked and restored as part of testing and maintenance? How often is the media refreshed?
- What about fixity?
- What about the primary storage? Is it checked or refreshed regularly? Is there any redundancy if that primary storage is online? If it is offline, how can it be requested by staff?
- What metadata is being kept and created about the different notebooks?
- What file formats are being retained? Is any data being stored on the different file formats? Presumably with research data, there would be a large variety of data.
- How long are these annual backups being retained?
- Is your IT department actively going to share the ELNs with staff?
- If it is currently the PI and department's responsibility to store physical notebooks, what will be the arrangement with electronic ones?

Got anything else you would ask your IT department when looking into preserving ELNs? Share in the comments below.

SHARE THIS:



This entry was posted in [case study](#), [digital preservation](#), [research data management](#) by [Sarah](#). Bookmark the [permalink](http://www.dpoc.ac.uk/2018/09/12/elns-part-2/) [<http://www.dpoc.ac.uk/2018/09/12/elns-part-2/>].

About Sarah

Digital Preservation Specialist - Outreach and Training:
Bodleian Libraries, Oxford University

[View all posts by Sarah →](#)

2 THOUGHTS ON “ELECTRONIC LAB NOTEBOOKS AND DIGITAL PRESERVATION:

PART II”

somaya

on **23 November, 2018 at 10:54** said:

Hi Jon,

I don't think Sarah got a chance to reply to you as she headed off on maternity leave just after she wrote this blog post. Personally, I think several user stories should be investigated, and while we didn't get to either of the Jisc research notebooks events over the past fortnight, Valerie McCutcheon in her email (<https://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=RESEARCH-DATAMAN;530af50b.1810>) has mentioned that there will be a write-up and potentially followup discussions about next steps.

I do also think is something that needs to be better taken up collaboratively between digital preservation practitioners, research data managers and IT departments together – to devise ways forward together (like so many digipres concerns, cross-discipline input is so valuable). While Oxford is trialling LabArchives, other higher education institutions will be looking at different options. And (I suspect) different disciplines are leaning towards different mechanisms and tools for the electronic lab notebooks.

There's an article in Nature here:

<https://www.nature.com/articles/d41586-018-05895-3> that references the work that Harvard did to put together an Electronic Lab Notebook Matrix – <https://datamanagement.hms.harvard.edu/electronic-lab-notebooks>

I expect you're aware of all of this already, so hopefully the Jisc event has some info on the priority user stories that have come out of the events over the past two weeks.

Thanks
Somaya

Jonathan Tilbury

on **13 September, 2018 at 09:30** said:

Hi

Would it make more sense to write a harvester to use the API to extract the material from the Lab Notebook into a fully fledged DP system? This would extract the files, metadata and structure according to specific criteria and recreate this in the DP system so it can be re-used. This is what we are working on with SharePoint etc and it should also work with lab notebooks such as LabArchives

Jon

This site uses Akismet to reduce spam. [Learn how your comment data is processed.](#)