



It's Academic: Data Portability and ELNs

The Life Sciences are slowly but surely making the move into Electronic Lab Notebooks (ELNs), and while much attention has been focused on the barriers to entry *into* ELNs, so far little has been spent on the flip side of that coin - getting data back *out* of an ELN.

When a lab (or department, or college) adopts an ELN, they are making a long term commitment and a serious investment of time and effort - they will be trusting whatever system they ultimately implement with their life's blood, the end result of millions of dollars of R&D money. And yet the transient nature of science is often overlooked - especially in academia. We all know the drill - technicians, rotating students, graduate students, postdocs - they all come and go, some more frequently than others, but over the course of a single decade even a small lab can have several dozen people pass through it. It can be nightmarish in several ways - tracking who was on which project, trying to extract the relevant and vital knowledge from each of those brains (without resorting to freezing, grinding, and lyophilizing) - let alone trying to store all of that information somewhere where it can be retrieved later.

A well-designed ELN with sophisticated search capabilities can readily solve this problem. But what about the seemingly simple problem - who gets to keep the notebooks?

We've all been there - if you spend several months to several years on a project, it's nice to have something to show for it. Especially if you intend to pursue the same (or similar) research in the future! So it's common for graduate students and postdocs to "liberate" their notebooks when they leave the lab. Who can blame them, other than the PI, who *also* poured their life blood into the project? Other than the grant-managing institution, which technically owns the data?

Of course notebooks can be photocopied. But the amount of labor and resources involved often prevents this from happening. And what about data that simply isn't amenable to duplication? Much of my data from my grad school days lived in the form of large autorads (exposed x-ray films) of polyacrylamide sequencing gels. These autorads were on the order of 20" x 50" in size - *not* something I could just tuck into any old notebook, and *definitely* not something easily duplicated. And that was a geological epoch ago in terms of scientific data - nowadays our data comes off of high-speed sequencers and microarray machines and in-situ imagers that can all be pre-loaded with days worth of runs, resulting in an absolute profusion of data - you just can't take it all with you.

Enter the well-designed ELN. All data, all formats, all sizes. No copying and pasting (not in the literal sense, with scissors and tape and glue, anyway). Nightly incremental backups of *everything* - lab data, journal articles, and notebooks. And - if the ELN is truly well designed - data portability. PDF printouts of the entire notebook (or just select portions, if that's what you need) - beautifully rendered, containing everything the scientist put into the experiment, and best of all, legible. Exported, portable archives of notebooks, ready to import into the *next* ELN system that the researcher works with. Exported, portable archives of all the actual datafiles - text files, spreadsheets, instrument data, *everything*!

A beautiful dream?

We're *almost* there *today*. Most ELNs will generate PDF outputs of experiments. Some let you export entire notebooks as PDF and the aforementioned archives. One (The Rescentris CERF system) even lets you reimport those archives back into a different instance of that same ELN system and get right back to work (say, if you move to a different institute or lab that also uses CERF). CERF also supports [Iron Mountain](#) Digital archiving which allows you to remotely store your finalized records as PDFs, then access those records from a new instance of CERF at a new location.

None of them, however, will export the data into an archival format that can be read in *other* notebook systems produced by other vendors - for that you would need a common archival format supported by all vendors, and sadly, despite the efforts of organizations like IC3 (<http://www.i3c.org/>), this level of cooperation has still not appeared in the industry.

Despite this lack of interoperability between *brands* of ELN, this new software tool does allow relocating researchers to take copies of all notes, and raw data with them in *some* form, without denying the PI or the institute that same data, and this is clearly a huge step forward. The researcher, the lab, the PI and the institution *all* retain access to the data, so in this particular struggle for ownership, everyone wins.