

Destroying Data and Technology in Support of Students

Kelly McElroy

I'm here to talk with you today about destroying data and technology in support of students. It's a funny title -- it actually comes from someone else's presentation at another conference, which my friend misread. She shared it with me, and I immediately wanted to give *that* talk, so thank you for giving me the opportunity to, today.

Over the last year, I've been a part of the first cohort of the Library Freedom Institute, a training program making privacy advocates of librarians. So that is the spirit behind this talk -- inviting you to consider the privacy implications of our overall topic today of online-only, which of course involves lots of data and technology. The ways that libraries collect and use data, we are often complicit, purposefully or not, in data regimes that specifically undermine other values that we supposedly hold dear, including equity, democracy, and privacy.

I want to frame this around another group of people who focused on destroying certain technology. The term "Luddite" gets used these days to describe anyone who is resistant to technology in general, but the original Luddites were a group of radicals in the 19th century who were concerned about how certain machinery was being used to cut jobs and oppress workers. They weren't against ALL technology, they were just really skeptical of whose interests were being served -- in one statement, they called to "put down all machinery hurtful to commonality." This is the sentiment I invite us bring to our work: to consider what data and technology is, or could be, hurtful. And to destroy it, or better yet, prevent it from being created or captured in the first place.

And you don't have to go far to see examples of data being mishandled, if not outright abused. Just in the last week or so, data breaches have been identified from library vendors Elsevier and Kanopy. While it may not seem so egregious for the lists of what movies have been watched to be released, this directly violates our values around protecting what people read, learn, and search for.

And data accumulates and can be connected across systems -- a group of researchers in LIS refer to this as creating a "data double" -- the digital shadow left across those systems. Given that all data is vulnerable to breaches, let alone misuse by whoever collected it in the first place, any commitment to privacy should give us pause over the ways we contribute to that data double.

And certainly, we ask for a great deal of data from our students. This quotation comes from Unizin, a consortium of public universities, including Oregon State University, with a focus on data-driven solutions to problems including student retention and graduation rates. A common premise is that this vast array of data points can be used to make things better. But, better for who?

I took a few minutes to brainstorm just some of the data points of that vast array. You can probably think of other ways that students create data or metadata as they interact with systems on your campus, or simply use the campus network or spaces. Some of these data are probably maintained locally, others through third-party systems; some of the data are required, others are opt-in; some are controlled by students, but many are created without explicit student consent.

Certainly, there are ways to use and connect these data that can help identify and solve problems, or as we are all increasingly pushed to do, to demonstrate effectiveness of programs. However, channeling the Luddites, I have to ask: how does this vast array of data points get used? By whom, and to what ends?

While goals of student retention and graduation may feel supportive, data may not be applied equitably. Scholar Virginia Eubanks has written about what she calls the digital poorhouse -- investigating how supposedly data-driven solutions often exacerbate and formalize inequality. Data can be used to create barriers to access to services -- and then to justify those decisions. On a college campus, we might imagine how something like intrusive advising, informed by data-driven alerts, might have the effect of pushing underprepared students out of STEM majors, without consideration of why some students may be more likely to be underprepared. All of this, in the light of data breaches, and considering the interconnectedness of data, I have to say, I feel a little, well, destructive.

Libraries can choose to destroy data within our direct control: ensuring that our in-house systems collect only what is absolutely necessary, maintain it securely, and delete it after it is no longer needed. We can negotiate agreements with vendors to protect user data, and we can ensure that our users are informed about how their data is collected and used. But given our claims to moral ground, we also have an obligation to advocate in our broader communities. We may not always have the ability to destroy all the data, but we have to at the very least speak up.