



Neutrality and Algorithms in Libraries

Author _ T. J. Lamanna (professionalirritant@riseup.net)

A recently published book, Safiya Noble's 2018's *Algorithms of Oppression*, has become an extremely popular read in our field as of late. While the book highlights some very important information about how our digital architecture de facto marginalizes people, it offers few remedies, other than expressing concerns about humans' control of how the algorithm is built, thus influencing how it works. The book details how we must admit that our algorithms are human-generated, but does little to explain how this situation can be remedied beyond "fixing the algorithms." Algorithms cannot be neutral, nor should they be; they are created by people and thus inherit the biases, conscious or unconscious, of their creators. No human has the capacity to be unbiased, so no algorithm can be. If they were, they could easily be gamed by malicious actors who would try to skew results. They need to be constantly worked and massaged to make sure they are behaving in a positive and progressive direction.

So, where do librarians fit into this? We don't see ourselves as algorithm-heavy, but they pop up beyond the usual places we'd look, and they impact us in ways we may not be aware of. I'd like to focus on the most obvious and damaging place, which is our online public access catalogs (OPACs). An OPAC search engine is usually designed to default to keywords unless the patron specifically chooses something else. This means that patrons may find what they are looking for, but they are not encouraged to wander through the stacks and stumble upon things they may not otherwise find.

We can counteract this challenge in a variety of ways. First, through the use of displays. We can highlight books

that are important but may not circulate because people aren't aware of them. Using your website to promote lesser-known books is also another solution. Don't let the search engine algorithm dictate how you create displays. This is part of what we do as librarians, and since people rarely have an issue finding James Patterson, use the opportunity to highlight other authors, specifically marginalized ones your OPAC may not be finding. From speaking with colleagues, I've realized that most of our search engines don't do a fantastic job with discoverability and this has been my first-hand experience as well.

To improve the searchability of OPACs, I suggest working with whomever is running your OPAC and



talking to them about search parameters. Be aware how they work and, if at all possible, find an open-source OPAC so you can collaborate and make the edits you need to make. There are a lot of librarians with the technical skills and passion for working on these types of projects. With open source software, they are able to share and build a better system for all. We are, and should be, responsible for how our OPACs return results. Like many other conversations we have with vendors, we need to make sure we are pushing them forward to meet our needs, not the needs they believe we have. Too often we fail to push our vendors to create the dynamic services we need. A coalition of institutions should advocate for this change. Collaboration and solidarity should be the rallying call of the current library zeitgeist.

We also allow algorithms to dictate what we purchase, which limits our scope and may eat through precious resources; then it may cost staff time and funding to rectify problems with the acquisitions. As Noble points out, we can create algorithms that work for libraries, and though it is impossible to remove bias, libraries can work much harder on being aware and counteracting it. We pull information from our results and use that to choose how to purchase books, which usually puts us in a vicious, not virtuous, circle. We buy more of what circulates and less of what doesn't. This binds us in a Catch-22 because if an item isn't discoverable, then fewer of them will be found, fewer will circulate, and fewer will be purchased as a result.

This hurts marginalized authors: people don't read their books because libraries don't own them and we don't own them because 'people don't read them,' which is really a problem of discoverability rather than lack of interest. We shouldn't be taking circulation statistics at face value. While they are absolutely useful, they only tell part of the story. Learn from your community, highlight it, make it pop. Your community is invested in your library taking their requests seriously. A book requested by your community that only circulates a few times may be far more valuable than a best-seller that goes out a hundred times. Our algorithms fail to take that into account; it's a bias we as librarians can and should be aware of as we work to turn our collections into a more inclusive and diverse selection.

Algorithms are good, and can be powerful tools for libraries, but we need to understand them and handle them with care. We need to remember that we control them, not the other way around. Libraries have a mission to serve their patrons and communities, and the algorithms we use have an obligation to serve us so we can make the best decisions. We cannot take them for granted and must be vigilant about how they work and how we make use of them. If we've learned anything from Google's mishandling of their own algorithm (as Safiya Noble expertly shows), it's that carelessness and crassness can cause actual harm to our communities, and that harm is almost always focused on the most marginalized of communities.