On indexing and the discoverability of tea bags and inclusive resources

UKSG eNews 456

1 November 2019

Regina Everitt, University of East London

The shelf items at my local supermarket are periodically shifted around for one reason or another, for example to create space for new stock or to keep frequent shoppers guessing. It has become a bit of a game for me to work out the rationale for how items are grouped. Popping corn next to dry lentils and other pulses, but separated from microwave popcorn, which is next to crisps and other savoury snacks. But when the tea was moved across from the cat food alarm bells rang. Was this categorisation making some sort of judgement about cat owners? If someone needed to grab a box of tea in haste, surely they would look for it near the other hot drinks and perhaps biscuits. Cats would not come into the equation, would they?

As an information professional, I am interested in how people look for and find information. Of course, this discoverability must be as quick and painless as possible. When I was a technical author, I took great pride in creating detailed indexes for the user and reference documentation that I wrote. In those days, most of the documentation was in hard copy, so a good index was needed to enable navigation of the granola-dry, dense manuals. At one company, I wrote documentation for software used on the manufacturing plant floor. The pages were laminated to protect against dirt and grease, similar to the laminated pages of Argos catalogs still found in my local supermarket. The index, in theory at least, spared the users from randomly flipping through pages to locate information. In reality, however, the plant workers probably did what many of the Argos shoppers continue to do despite the prevalence of search tablets. They flipped through the pages until something caught their eye. This serendipitous approach was not particularly efficient for the plant floor workers but must be a great way for Argos to potentially up-sell as shoppers go in to buy one thing and leave with five of something else!

For the record, I always started with the index in the Argos shops just to test the precision of the work. The speed with which I found items or whether I found the items at all hinged on whether the terms used to categorise information were those that were meaningful to me, the user. The same was true for the plant floor workers using my documentation. Oh, the responsibility and power!

Today application documentation is now generally embedded in the tool under the Help button. Users can enter natural language and keywords are mined from the text. Information retrieval is faster, easier and with a good success rate. Reference manuals, if online, tend to be pdf with 'old school' indexing in the document itself. Retrieval of the pdf object depends on the algorithm used by the search engine, but searching within the pdf itself is a simple text-mining function.

For web searching, however, an unassuming keyword used to retrieve information can unearth some unsavoury results. Information professionals, such as Safiya Umoje Noble, have been critical about the assumptions behind web search algorithms. In Algorithms of Oppression, Noble argues that the Google search engine results perpetuated sexualised stereotypes about African American women, particularly. The search results that she describes in the book are no longer replicable for a range of reasons (eg addition of new pages, re-indexing of pages, changes to algorithm assumptions). The fact that she retrieved the information at all highlights how the categorisation of information and/or the language used to describe information can make social judgments (perhaps unwittingly) that are downright racist and offensive. I have to say that I was not particularly impressed to find 'angry black woman' appear on the first page of results when I simply entered 'black woman' (on 19 August 2019). Unfortunately, the fact that this tired stereotype was at the top of the search ranks enhanced its credibility to some impressionable soul.

As part of the 'decolonisation of the curriculum' movement, UEL Library teams, in collaboration with colleagues at other HEIs, have been collating a collection of inclusive resources for academics to use in their teaching, learning and research practices. I was keen for the staff members to tag the collection in some consistent way so that the documents were easily retrieved by staff and students. Staff members have already grouped physical and virtual titles into collections under the decolonisation and/or inclusive resources headings. That is, physical items are grouped together on shelves and virtual items (links to e-books and bibliographic information) are grouped as a collection online. But what would be the underlying description (or metadata) to bind the collection? I am less interested in the technical solution than the language that one would use to categorise the collection. Would it be decolonisation?

If I am honest, I struggle with the decolonisation heading. The Random House online dictionary defines decolonisation as 'allowing a colony to become self-governing or independent' and 'to free a colony to become self-governing or independent'. The definitions feel too passive to me as if independence was simply handed over without struggle or cost. And to group titles by non-white authors under this heading feels like 'othering' their knowledge, expertise, and work into a sub or tertiary group rather than part of a wider landscape of critical thought.

If we were to persist with the approach to tag works by, say, non-white authors under 'decolonisation', would anyone search for items under that heading regardless of the subject? We would certainly create a lot of spurious results. Boolean searching would enable some granularity but there will be many results about decoloniality rather than non-white authors specifically. Of course, non-white authors and a subject would yield some results to start, but this language still suggests white (western or European) as default and everything else as 'other. '

I now have visions of the manufacturing plant workers flipping through my manuals thinking "I can't find anything" because of my poor assumptions about how they would access information. For all the UX folk out there screaming "you could have just asked the end user what they thought", usability tests in those days were around use of the software and not the documentation.

Ah, I've over-thought enough. I just need to grab some haimisha pickles on the way home for a nice salad. Wait, so is that the world food aisle or condiments?