Artificial intelligence as an instrument for smarter cataloguing: a prospective dialogue

Mathilde Koskas
Department of Legal deposit, National Library of France, Paris, France.

Mélanie Roche
Department of Legal deposit, National Library of France, Paris, France.

Abstract:

Empirical evidence suggests that the hiring, training and retaining of cataloguers is increasingly challenging for libraries. While funding and advocacy with upper management and government agencies that prioritise costs without understanding the intricacies of today's metadata ecosystem is a factor, a lower interest in the traditional competencies of cataloguers seems to be at play as well. Meanwhile, Artificial Intelligence (AI) and related technologies garner interest from decision-makers for their perceived modernity and economic promises, as well as from information professionals who have developed automated metadata processes for decades.
In times of economic constraints and a rapidly-changing workforce, apposite use of AI technology is a major factor, and it is incumbent on us to find how best to harness it as an efficient tool addressing the challenges we face. As far as metadata production is concerned, this could mean finding innovative ways of automating "traditional" tasks that production chains still need but cataloguers are less adept at, freeing them for the new tasks that have emerged as the consequence of the evolution of models and standards and where human intervention remains more than ever indispensable.
Drawing on French and international experience, this is a dialogue between two seasoned cataloguing and bibliographic models experts who take a prospective look at the recent technological developments and their potential contribution to the challenges faced by cataloguing as a profession. It is our belief that technological progress should be about freeing human intelligence

from repetitive and non-meaningful tasks, the better to bring about a world where cataloguers can focus on value-added, smarter tasks. This is how the cataloguing community has used automation in libraries in the past, and there is no reason why this should stop at the newest technologies.

Introduction

We are both cataloguing managers at a National library: we manage units whose mission is to provide widely reusable metadata for the national output.

While the question of AI has become unavoidable, it has not necessarily been addressed from a management point of view, which is our angle today. Not conceptual but the concrete situation in our library - in which hopefully you'll recognise something of your own.

What exactly is the issue for you from a managing point of view?

Libraries report an increased difficulty in hiring and retaining trained cataloguers. In France, some of it at least can be traced to policy (or lack thereof). Most library positions are public service jobs, accessed through entrance exams. The number of positions is determined by the ministries for Culture and Education. Successive governments have sought to contain or decrease the number of civil servants, including by limiting the number of positions opened and by hiring more people on private law contracts. The exams themselves have evolved with the profession, where direct service to the public and public-facing jobs tend to take precedence over collections management, including cataloguing.

And what incidence does it have on the profession?

As cataloguing becomes secondary in the exams, the LIS programmes that prepare for those exams devote less time to this aspect of librarianship. Thus, people come into the workforce with less knowledge of both the technical aspects of cataloguing and its role/importance (See the relative oblivion into which the concept of Universal Bibliographic Control had fallen, while the very practices that are constitutive of it are still considered fundamental). One could say it is a frame of mind. It makes training and retaining of qualified cataloguers more difficult.

So isn't AI the perfect answer to your problem?

I would argue that lack of knowledge and skill doesn't mean lack of interest or importance.

It's still a fundamental mission and tied with professional identity. As managers, we should seek workers' well-being. Using tools that make cataloguing easier but doesn't make it a meaningless task.

So how can we make cataloguing easier?

Let's look at our cataloguing rules.

Whereas ISBD was created at a time and for technologies that implied the mediation of a library professional, the movement that led to the creation of IFLA LRM was based in a user-centered approach. User tasks have also been embedded in the ICP since 2009. The cataloguing codes and standards derived from those principles offer possibilities that may be less cataloguer-friendly. For instance, in recording information in 4 possible ways (RDA), or getting rid of all abbreviations.

You say it like it's a bad thing. What's the problem with that?

It isn't a bad thing for the end users, but which cataloguer wants to type out "place of publication unknown" instead of "s.l."? While writing the standards, when such concerns are raised, heavy reliance is placed on technology to resolve potential difficulties in the production process. But reality is different: it may be difficult to implement a simple way of displaying the desired full phrase for end-users while sparing cataloguers from typing it out or to have a production software that manages hundreds of value vocabularies in ergonomical ways… But one thing is sure: ideally, resolution of conflict between user-friendly and cataloguer-friendly is where technology should intervene.

I agree. Technology's ideal contribution to human life is relieving us of tasks we don't want to do.

Which raises the question of what these tasks are: what is seen as not rewarding, not interesting, not cost-effective? The first two at least are subjective: some cataloguers revel in subject indexing while others favour the intricacies of a structured bibliographic description. Very few of them would like to stop cataloguing altogether.

Studies have shown that overall, cataloguers are deeply attached to the technical aspect of the profession (Philibert 2024).

But shouldn't we automate everything that we can since there is so much work to do in other areas?

As managers we are not only custodians of collections and metadata but also responsible for the well-being and professional development of the people who report to us. Our ambition should be to make sure the right skills are used where they are needed, rather than deprive our colleagues of work they value and skills they and the institution have heavily invested in acquiring.

Even efficiency is relative to context and objectives. A National Bibliographic Agency whose mission is to create the metadata of record for the national output will deem it efficient to invest in processes that a public library whose job is not to preserve collections in the long run but to make them available quickly doesn't.

One thing is sure: today, talk about technology that could replace some human tasks implies AI.

Let me stop you right here and clarify what it is we are talking about.

Or, Ask not what you can do with AI but what AI can do for you.

Let me remind you that the term "AI" is usually used as a catch-all for lots of different concepts. To quote but two, LLM and Generative AI are the two that have been quoted the most this congress: LLM, Large Language Models that need to be trained with an enormous amount of data before showing results. Generative AI that creates new content from existing data. And many many more.

Applying this to cataloguing, does it mean that there are several possible approaches? Are we talking about one program that applies a set of targeted actions to a batch of records? Or one that would generate full individual records?

Certainly, and the first one seems easier not only from a technical point of view but also from a change management point of view, meaning how we introduce it to our teams and also how we wrap our own heads around it. Let me take an example. Apart from the decision-makers, the idea of an indexing bot scares everyone. Whereas if you say "we are developing a tool that will help with indexing by suggesting concepts", it is not hard to get everyone onboard. Because we are used to working with automated tools that make our work easier. It's always been about the right tools and the best tools, not the trendy tools. Whether or not it says "AI" on the package is irrelevant – or at least should be.

So what makes it the right tool or not?

Well, that's where the question of cost comes in: whether we buy it from an external vendor or we develop it in-house, the economic and environmental costs of using AI (especially generative AI) have long been emphasised. Large institutions such as wealthy universities are developing pilots whose cost would be prohibitive even as final projects for smaller institutions. If we use an external tool, it can disappear overnight if the company ceases to exist or no longer supports it, or even more radically if legislation intervenes and makes some aspects of AI illegal.

In that respect, IFLA recently identified 14 possible risks linked to AI, among which the first one was the potential exaggeration of its added value (Cox, De Brasdefer 2025). As the document points out, « There remains limited use cases in the library field with demonstrated cost benefits ». And of course cost-efficiency is only one criterion among others, because efficiency cannot be measured solely in terms of financial resources invested and created.

Absolutely!

We have strong frameworks for our work and missions, like UBC ("To provide free and open access to trustworthy information to library users to meet their information needs, via appropriate current technologies"), ICP.

Libraries perform a public service. Bibliographic and authority metadata's value is measured in the ability of users to meet their information needs. Recreation and relaxation (imagination), research and professional use, health, informed decision-making (citizen, reliability of information).

We also have an ethical responsibility to act as responsible managers.

As such, we cannot make the value of the work of our colleagues completely secondary to blunt efficiency.

In other words, do not provide workers with tools they hate.

Indeed! We should never become the drone workers of AI or any other technology. Libraries as employers have a responsibility in modelling ethical managerial practices around use of AI.

We've all been saying that AI must be developed and used in an ethical way. And I think we all agree that smart cataloguing is all good and well, but in the end human intelligence must prevail.

In conclusion

We hope to have shown a few actual ways that can be manifested: working together in our institutions to have an objective definition of repetitive tasks that can be automated and consensus on what should stay within the remit of

human operators, both for the quality and interest of the work. Which implies, among others, not forgetting about fundamental principles and values, why we have the rules we enforce, etc.

We derive easy optimism from the fact that previous technological revolutions haven't replaced us. Let us not forget that happened through hard work, trial and error, adapting and demonstrating the value of our work. And that some parameters are out of our hands, such as the political and economical context.

Not to end on such a heavy note, I'll leave you with this thought: the global cataloguing ecosystem is and will remain a Lego set: we'll find ways of implementing AI that are adapted to our specific contexts while keeping our sights on our fundamental missions.

References

Cox, Andrew M; De Brasdefer, Maria. *IFLA Entry Point to Libraries and AI*. IFLA, 2025.
https://repository.ifla.org/handle/20.500.14598/4034

Philibert, Coralie. *Catégorie B et catalogage en bibliothèque : histoire d'un statut et d'un cœur de métier, état des lieux actuel et perspectives d'avenir*. Mémoire d'étude pour le diplôme de conservateur de bibliothèque. Villeurbanne : ENSSIB, 2024.