

From Bias to Transparency

Ethical Imperatives in AI-Based Library cataloging

Hannes Lowagie
WLIC 2023 Satellite Meeting

*Universal Bibliographic Control at the Crossroads :
The Challenges of Unifying IFLA Bibliographic Standards*

Universal Bibliographic Control

- Standardized approach: output, cataloging rules
- Challenge due to different cultural perspectives and diverse interpretations, especially in subject indexing
- Proposition via AI and linked subject indexing

Mauro Guerrini, 2022

- *“Who knows what the future will bring us? Perhaps, we are still at the beginning of the digital revolution. Precisely in the field of metadata and authority control, we could expect developments and surprises from alternative technologies on machine learning or artificial intelligence, a tool that promises to be very useful; a tool that takes nothing away from the cataloguer’s judgment, which remains a fundamental intellectual activity.”*
- Yes, tools available
- Presentation of Power Platform Wednesday 23/08 at Rotterdam
- Subject indexing → ‘tagging’ texts

Subject indexing using AI

- Database with two columns: texts and 'tags'

	A	B
1	Public administration	L'histoire des institutions administratives françaises s'inscrit dans une remarquable continuité que peu d'auteurs ont, jusqu'à présent, mise en valeur. Des travaux récents montrent pourtant que notre administration plonge ses racines les plus profondes au Moyen Âge. Les ministres d'aujourd'hui sont les successeurs des secrétaires d'État de l'Ancien Régime, eux-mêmes descendants des notaires-secrétaires qui entouraient le roi à l'époque féodale.
2	Political science	Ce livre de petit format réunit quelques-unes des réflexions les plus pertinentes et les plus actuelles de Paul Valéry (1871-1945) sur la politique. Lui qui ne fut d'aucun parti, ni de droite ni de gauche, et qui voulut, toute sa vie, conserver son indépendance, a rapidement pris les mesures du système démocratique auquel il se résignait plutôt qu'il ne l'approuvait...
3	Human rights	Peu connu en France, Thomas Paine fut pourtant député de la Révolution française et ardent défenseur des droits de l'homme. Après avoir été parmi les organisateurs de l'Indépendance américaine en 1776, il a rejoint Paris pour défendre, par la plume et par le verbe, les valeurs fondamentales de liberté, d'égalité et de fraternité, et leur inscription juridique dans la Déclaration des droits de l'homme et du citoyen....
4	Wellbeing	Des rituels de respiration pour apaiser ses émotions, booster son bien-être et se transformer ! La respiration est au cœur de la vie ! C'est elle qui apporte l'oxygène et l'énergie de bouger, c'est elle qui équilibre le système nerveux et, par-là, les émotions. Elle agit même sur le sommeil et sur la digestion. Lorsqu'on respire, on revient à soi. La respiration permet de se reconnecter à son corps, à soi et à ses envies profondes....
5	Psychology	Les bases fondamentales et les grands concepts de la statistique en psychologie explicités et mis en scène dans un manuel clair et visuel, illustré par 150 photos, graphiques et tableaux. Dans chaque chapitre : zoom sur les notions essentielles ; mise en avant de définitions et de renvois vers des sites web ; développement d'applications et de documents interactifs (exercices corrigés, QCM, webographies). Cette 3e édition bénéficie aussi de compléments en ligne (vidéos).
6		
7		

Potential Pitfalls: fictitious example

<p>The integration of AI and Internet of Things has transformed environmental monitoring. Through sensor networks and data analytics, AI can predict natural disasters, monitor air quality, and manage resources efficiently, ushering in a new era of smart environmental management.</p>	<p><u>Attributed:</u> Environmental Monitoring Internet of Things</p>
<p>The widespread adoption of Internet of Things devices has raised concerns about data privacy. AI-driven analysis of personal data collected by these devices necessitates stringent regulations to protect user privacy and prevent unauthorized access.</p>	<p><u>Attributed:</u> Data Privacy Internet of Things</p>
<p>In the realm of modern technology data analytics has paved the way for remarkable advancements. This synergy enables us to predict and mitigate the impact of natural disasters with unprecedented accuracy. By harnessing data from diverse sources, we gain crucial insights into the changing environment. However, this integration also raises concerns about data privacy. As we delve into the intricacies of environmental data collection, the safeguarding of personal data privacy becomes a paramount consideration.</p>	<p><u>Suggested:</u> Internet of Things</p>

Subject indexing using AI: solutions to avoid wrong interpretations

- Diverse Training Data
- Fine-tuning and customization
- Human-in-the-Loop approach
- Evaluation and Monitoring
- Transparency and Documentation

International collaboration

- Do we want future where each NBA will train its own model, or can we collaborate?
- Can we create one model and thus harmonize ‘input’ (instead of focusing on a standardized ‘output’)
- First idea: one worldwide ‘model’ for subject indexing
- However...

Embracing different interpretations

- One world but different interpretations, even on basic concepts like ‘poverty’, ‘youth’,

Books about people that are...	NBA A	NBA B	NBA C
15 years old	Childhood	Childhood	Childhood
16 years old		Young Adulthood	
17 years old			Adolescence
18 years old	Adulthood		
19 years old		Adulthood	
20 years old	Adulthood		
21 years old		Adulthood	
22 years old	Adulthood		
23 years old		Adulthood	Adulthood

Embracing different interpretations

- Avoid imposing a singular interpretation as the ‘true’ one
- KBR: based on data.bnf.fr:
 - export summaries of books related to subject x
 - Use data to train data for that concept
 - Result: dataset KBR is a French, European model, mostly based on manual efforts French cataloguers. We have to be aware of that

Embracing different interpretations

- **IFLA National Bibliographic Register**
 - 20% local classifications,
 - 60% local subject vocabulary
- **Value of local classifications and subject vocabularies**

Main question

- *“How can we achieve a universal way of organizing and accessing information with respect to each culture’s different interpretations, resulting in an improved ‘universal’ discoverability of nation’s publications?”*

Proposition

- **Unified interconnected knowledge graph by combining and mapping local classifications and subject vocabularies**
 - Theoretical, challenging, complex, maybe impossible but worth investigating
 - Embracing diversity of cultural perspectives

Mapping

Books about people that are...	NBA A	NBA B	NBA C
15 years old	Childhood	Childhood	Childhood
16 years old		Young Adulthood	Adolescence
17 years old			
18 years old	Adulthood	Adulthood	Adulthood
19 years old			
20 years old			
21 years old			
22 years old			
23 years old			

Proposition

- **Links can be created by comparing classifications of same 'works' (ifla-irm)**
- **Research Project BELTRANS**
 - translated works from KBR, BnF, BL, KBNL
 - same works, different subject indexing
- **Establish meaningful relationships between the diverse classification schemes**
- **Create one dataset**



Challenges

- **Conflicts in interpretations**
- **Ambiguities**
- **Scalability**

- **--> shared ontology**

Advantages

- **Researchers**
 - Extensive and interconnected Knowledge Base
- **Librarians**
 - Streamline cataloging processes
 - Environment of collaboration and knowledge exchange
 - Better serve their users
 - Part of a global knowledge-sharing network
- **Users**
 - Access a wider array of resources beyond their local library
 - Enriching understanding of diverse cultures, ideas and perspectives

Advantages

- **Potential to foster cross-cultural information retrieval**
- **Users can navigate through interconnected web of concepts**
 - Uncovering related materials
 - Deeper understanding of diverse viewpoints
 - Access a wide range of resources
 - Navigate and explore information in a manner that aligns with their unique context and preferences
- **Users can access materials relevant to their interests, regardless of the originating library's classification**
 - Ethical considerations

Ethical considerations

- **Cultural sensitivity**
- **Diversity and Inclusion**
- **Bias and Fair Representation**

Zwarte Piet



Het Boek van Sinterklaas en Zwarte Piet



ISBN 978-9490470029

NBA A = 'Saint Nicholas', 'Children', 'Saints'

NBA B = "Racial representation"

- Meaningful to create a link between both subject terms
- Citizens of NBA A can see that their indexations is linked to 'racial representation'
- When one searches 'racial representation' → also that books as search result

Piet



Final ethical considerations

- **Artificial Intelligence**

- Do not trust AI
- Create and maintain your own training set (and publish it)
- Opportunity: train a ‘universal’ brain to interpret texts (‘input’)
- Can only using proposed ‘mapping’ between local classification schemes and vocabularies (respect local interpretations)

Final ethical considerations

- **Classification and subject indexation**
 - Cultural relativist
 - Finding a balance between respecting cultural differences and ‘correcting’ a cultural viewpoint (for example ‘Zwarte Piet’).
 - Foundational framework
 - Universal Human Rights
 - Inclusivity
 - Proposed mapping does not correct, but can expose possible harmful terms and offer diverse insights

Conclusion

- **Unified and interconnected knowledge graph by combining different local classifications and subject vocabularies through skos links.**
 - Diversity of cultural perspectives
 - Harmonious network of interconnected concepts
- **AI-algorithms can play a pivotal role in**
 - Recognizing patterns and relationships
 - Automatic subject indexing
- **Ethical considerations**

Open questions

- Is it possible to collaborate internationally to create a universal model for subject indexing using artificial intelligence?
- Is it possible to 'map' different classifications and subject vocabularies?
- If a NBA uses AI for the automated subject indexing, should it publish its training data? Should the NBA indicate that subject indexing was made by AI?
- If a NBA identifies a harmful term or potentially offensive language in another CA's classification scheme or subject vocabulary, is there a mechanism for communication? If yes, how can this communication take place? Furthermore, should a NBA reflect on such observations and remarks?