ADVANCES IN KNOWLEDGE ORGANIZATION
Proceedings of International ISKO Conference - ISSN: 0938-5495 - 2026 - São Paulo

Model Cards: AI Regulation and the Renaissance of Structured Metadata

D. Grant Campbell

Faculty of Information and Media Studies, University of Western Ontario, Canada

Abstract

Amid the growing concern about the regulation and containment of general artificial intelligence tools, one solution has emerged within the very industry that is developing these tools: the use of structured metadata in the form of model cards, which require the developers to assign values to a pre-set collection of descriptive attributes. Model cards outwardly resemble the practices of resource description that stretch back to the nineteenth century. The unit of description, however, is not a fixed, published artifact, but a more amorphous entity: a project, defined by its use, metrics and evaluation, and enhanced by explicit treatment of its ethical considerations. The model card practice has been greeted with fully justified skepticism by commentators who argue that this practice of cards, especially developed and maintained internally within organizations such as Google and Open AI, will not provide adequate safeguards against the social, economic and existential dangers posed by generative AI systems. Their value, I argue, lies in their use of the practice of structured description in the form of a standardized vocabulary and areas of description. These emerging standards will hopefully enhance the awareness and the influence of organizations and individuals who seek to engage critically with generative AI tools.

1.0 Introduction

As generative Artificial Intelligence systems develop with increasing speed, the governance and regulation of these systems have become a cause for deep concern. According to Mustafa Suleyman (2025), former head of DeepMind and now the CEO of Microsoft AI, while the existential threat posed by generative AI has created an urgent need for some means of regulation and restraint, such regulation is all but impossible:

> Governments should, on the face of it, be better primed for managing novel risks and technologies than ever before. National budgets for such things are generally at record levels. Truth is, though, novel threats are just exceptionally difficult for any government to navigate. That's not a flaw with the idea of government; it's an assessment of the scale of the challenge before us. ... Governments fight the last war, the last pandemic, regulate the last wave. Regulators regulate for things they can anticipate. This, meanwhile, is an age of surprises (Suleyman 2025, 228).

As the technology sector develops ever more powerful systems, their inner workings elude our understanding and our capacity to control their evolution and development. And while

governments may offer the most important and immediate means of such regulation, they hold only a slim chance of successfully reining in the potential harms of these systems.

Within this bleak scenario, one small but astonishing development is worth noting: the renaissance of descriptive cataloguing in the form of model cards. As presented by a team from Google in 2019, model cards offer a new form of documentation for trained machine learning models, strongly resembling resource description principles that stretch back at least to the mid nineteenth century, with Sir Anthony Panizzi in Great Britain and Charles Cutter in the United States. While the descriptive elements and presentation format differ vastly from those that filled the vast card catalogues of early twentieth-century libraries, the underlying principle appears to be the same: to create a set of structured metadata elements which address those attributes of a tool identified as relevant for specific purposes. And across the gulf between Cutter and Google, we can recognize a common need: to link the practice of documentation—be it a book or an AI tool—to an articulated set of objectives. But is this resemblance a mere passing coincidence, or a source of insight that could lead to the regulation we need? This paper will argue cautiously for the latter. Bibliographic description and model cards do have an important connection: not because of their shared success, but because they share a rich and productive failure.

2.0 From Cutter to Google: The Objects of a Model Card Documentation System for Large Language Model AI Tools

In their introduction of the model card system in 2019, the Google team argues that AI model designers should provide documentation that covers "model performance characteristics, intended use cases, potential pitfalls, or other information to help users evaluate the suitability of these systems to their context" (Mitchell et al. 2019, 1). From the outset, they realized that model cards had to address the needs of multiple stakeholders, including not just the practitioners and developers of AI systems and software, but also policymakers, "ML-knowledgeable individuals" who might want to adjust systems for their own use, and above all, "impacted individuals who may experience effects from a model" (Mitchell et al. 2019, 3). To meet these varied needs, the team developed nine areas of description, as follows:

- model details: including the developer, model date, version and type, information about features, sources of further information, citation details, license, and directions for questions and comments;
- intended use: including primary uses, primary users, and cases where the tool might be used outside its original scope;
- factors: including demographic or phenotypic factors that may be relevant, either for using or evaluating the tool;
- metrics of impact: including performance measures, decision thresholds and variation approaches;
- evaluation data: including the datasets used for quantitative analysis;
- training data;
- quantitative analyses;

- ethical considerations;
- caveats and recommendations.

While the content of these areas are related to AI projects and not bibliographic resources, the practice of organizing the key features of description in a clear order based on anticipated relevance should be familiar to anyone who has worked with the International Standard Bibliographic Description.

Responses to the model card approach have centered around two primary questions:

2.1 Who is creating the metadata?

So far, AI organizations are creating model cards for their own AI models and services, and they have strong motivations to do so: early research shows that resources with model cards have significantly higher download rates (Liang et al. 2024). Such research, however, also shows that developers do not approach all the areas of description with equal precision and care. Areas that deal with environmental impact, limitations and evaluation (Liang et al. 2024) frequently lack detail. The areas dealing with ethics are also subject to wide variation (Mehraj et al. 2025), with greater emphasis placed on the model's capabilities and reliability and less on the model's "explainability, user autonomy and fairness" (Puhlfürß et al. 2025, 1).

For this reason, we are naturally seeing interest in model card development within communities of use, rather than within the organizations that developed the models. In the future, we may see the use of model card authoring kits designed for creation and use by voices within communities affected by AI services, to promote discussion and comprehension among those experiencing the impact of such services (Shen et al. 2022).

2.2 Who is Reading the Metadata?

Some discussions envision this metadata within a machine-readable environment, similar to the Web scenario in which search engine crawlers respond to web-embedded metadata that directs them toward sites to crawl and away from sites deemed private. In AI environments, the analogous figure would be the "service provider," as defined by Anthropic's Claude Constitution (Anthropic 2026), the frequently automated intermediary between the AI model and the end user. Such environments would favour conformity to established standards, to facilitate the automated assessment, selection and configuration of certain AI models according to generally-accepted criteria (Mehraj et al. 2025).

Others argue that the primary work of model cards is to bridge the complexity of AI systems with the human comprehension needed to make important human decisions. In such cases, the emphasis falls less on conformity and more on the need to explain advanced and sophisticated systems to decision-makers who don't possess deep expertise in machine learning or natural language processing (Crisan et al. 2022).

While none of the literature to date mentions any ties to bibliographic description, the questions posed by researchers strongly resemble the questions that have dominated the development of bibliographic metadata. From the beginning, libraries have striven to bridge the gap between experienced cataloguers and the diverse and less-experienced user communities they serve. And the history of the Dublin Core Metadata Initiative has been marked by a similar tension between metadata designed for agents and crawlers that favour consistency and interoperability, and metadata designed for human communities that are anything but consistent and interchangeable.

3.0 Model Cards and Traditional Bibliographic Control

The merest glance at these areas shows how far they diverge from those of the International Standard Bibliographic Description (ISBD) used by most libraries. Model cards are not designed to describe published books: artifacts that have been replicated from a single master copy and widely distributed in a state of typographical fixity that does not acknowledge further changes. Model cards are designed to describe *projects*: coordinated activities surrounding the creation, development, deployment and evaluation of highly flexible and changeable tools that are used for multiple purposes and in multiple environments.

Nonetheless, by looking at the primary entity of concern in model cards as a project, we can detect a more subtle connection to the descriptive cataloguing tradition, by considering how that library tradition has dealt with what Google would call “Out-of-scope use cases”: cases which “the model might easily be confused with, or related contexts that users could try to apply the model to” (Mitchell et al. 2019, 4). As anyone who trained on either the Anglo-American Cataloguing Rules or their successor, Resource Description and Access, knows, descriptive standards intended for situations of typographical fixity have frequently been stretched and extended to embrace resources that are anything but fixed.

From the beginning, serials cataloguers have been forced to adapt the rules designed for self-contained, published books to describe resources that are not books, and are anything but self-contained. Serials are, by their nature, group efforts, drawing on an ever-changing multitude of authors, whose contributions are drawn into discrete sets of contributions which follow in chronological sequence for an undetermined period. The result was an endless process of barely resisting chaos. Cataloguing the articles of journals has been outsourced to indexing services. Cataloguing the journals requires complex strategies for selecting access points, determining main entry, and noting details such as numbering, the presence of cumulative index issues, and parallel editions in different languages and formats. Cataloguing serials is an endless process of careful effort and improvisation on the part of trained cataloguers: the system generally works, but not with effortless and inevitable ease.

The advent of Internet resources in the 1990s capped a decades-long process of adapting to a widespread proliferation of different formats, each of which had different principles of publication, authorship and access. With Internet resources came the problem of ever-

shifting content: instead of receiving regular supplements, Internet resources were subject to change and revision throughout the entire content, with little or no advance notice. Authorship was not merely widely dispersed; it was frequently impossible to determine. The paradigm of publication made very little sense on the World Wide Web. As the Web moved from being a "library" to being a "platform," multitudes of other types of publication—music, books, motion pictures, computer programs—migrated onto it.

Moving with the painful slowness of a consensus-building body, the Steering Committee of the cataloguing standard, *Resource Description and Access*, developed new paradigms to deal with the fresh challenges, including the concept of Integrating Resources for websites. These paradigms are under constant revision, as the standard struggles to keep up, not just with new information formats such as streaming sites, social media and games, but also with IFLA's evolving *Library Reference Model*, which seeks to redefine all library cataloguing according to the principles of linked data.

The point I want to make here is not that resource description standards failed: merely that they met with incomplete success, and that the incompleteness was inevitable. Information production has always proved resistant to bibliographic control, and the dream of complete control has typically been the topic of satire, as in Jonathan Swift's *The Battel of the Books*, or fantasy, as in the short fiction of Borges. Libraries have never achieved complete control of their collections. The best they have done is to create a bridge that extends partially over the chaos, creating small areas of order, surrounded by bordering areas of limited order, and other information domains that might as well be labeled "here be dragons." If we're fortunate, our users take what they can learn from the orderly oases and then engage in some effort to battle the chaos: such battles frequently involve contemptuously discarding our tools, but not before the tools have provided the user with fresh ideas.

If we look at Model Cards through the prism of the vexed history of cataloguing, we can see grounds for similar skepticism, for a very similar reason. The model card scenario looks very satisfactory when leveled at an application which aligns with a tool's intended primary use. But just as libraries have painfully outgrown their reliance on the book model, much of the concern around AI system lies in their potential drift into out-of-scope uses, and the need to administrate and regulate not just those systems in the safe centre of a domain but also those on the fringes. And the development and commentary on model cards so far suggests that familiar predicaments will ensue.

4.0 Relation to Bibliographic Description

If we hold these questions up against the history of bibliographic description since the beginning of the Twentieth Century, we can perceive the outlines of long-lasting and enduring problems faced by cataloguers since the days of Charles Cutter and Charles Jewett, and which remain stubbornly relevant to AI model cards. Three familiar axes of conflict have appeared (see Figure 1):

- **Internal or External Authorship:** Which is more useful and reliable? Metadata created by the original authors or publishers of the document? Or metadata created by trained intermediaries such as professional cataloguers? We could ask the same of model cards. Who should create and maintaining these model cards? The developers themselves, or bodies outside the development environment, may not know as much about the tools, but may be in a better position to assess their social and ethical implications?
- **Embedded or Separate Metadata:** Should the metadata be an intrinsic part of the document, such as a header in a Web page or the page numbering and title page of a book? Or should the metadata be stored separately in a resource such as an external catalogue or index? Likewise, should model cards be embedded into the tool itself as a means of identifying its properties to automated agents that could assess the tool's utility? Or should they be maintained separately as a means of facilitating an assessment of a range of tools?
- **High or Low Consistency:** Should the metadata rigidly conform to standards of consistency to enable the aggregation of data in a union catalogue? Or should it sacrifice consistency in order to meet the varied and diverse needs of users? This echoes the more current question: is AI regulation best achieved through the rigid enforcement of a central set of agreed-upon standards and principles? Or should it be designed to investigate and assess individual tools in individual contexts, giving full scope to the social, economic, and ethical specificities of each tool?

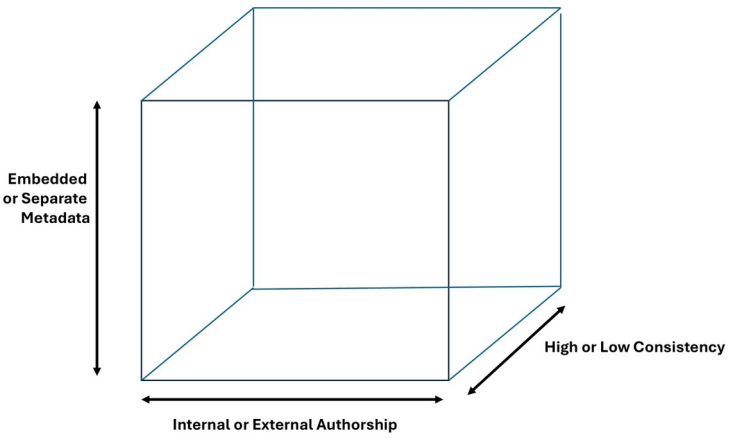


Figure 1. Axes of Conflict

In addition to these familiar complexities, another presents itself: the shifting nature of standards, both internal and external. Since the days of Cutter and Panizzi, bibliographic control has had to adjust to constant revisions to its own standards, with the *Anglo-American Cataloguing Rules* replacing earlier standards, and then yielding to *Resource Description and Access,* which has endured ongoing, frequently drastic revisions. Similarly, the initial standards proposed for model cards are already undergoing pressure to revise, as new categories and taxonomies emerge in response to the growing study of "ethical AI" (Kennedy-Mayo and Gord 2024).

Bibliographic control has also been forced to revise in accordance with the changing demands of external standards such as the International Standard Bibliographic Description, the Functional Requirements for Bibliographic Records, the IFLA Library Reference Model and the IFLA's International Cataloguing Principles. AI Model Cards will likewise need to adapt to changing standards of safe and ethical AI, as more and more standards emerge alongside the European Union's taxonomy and definition of safe AI practices.

Finally, both communities grapple with the sheer pace of change. The breathtaking difference in the pace of AI development, compared to the pace of publication developments in the Twentieth Century, threatens to blind us to a shared reality in both fields. It is widely recognized that regulatory practices cannot keep up with the pace of AI innovation and change. We must, however, recognize in turn that we could never keep up. The history of cataloguing has been a history of trying to embrace constant change within a stable infrastructure of metadata creation, encoding and use, often with wildly uneven results.

5.0 Conclusion

So, then: will structured metadata in the form of model cards make us all safe from the existential threats posed by Artificial Intelligence?

Judging by the history of structured metadata in bibliographic description, probably not. Structured metadata did not protect us completely from the complexities of knowledge organization in real life. We did our best, and we continued to do our best. But information, even at the comparatively glacial pace of conventional publishing, is a many-headed monster, and every time you cut off one head, two more grow up in its place. And if structured metadata cannot completely control library collections, it is unlikely to provide effective and complete regulatory control over generative AI.

Nonetheless, the history of bibliographic control in libraries suggests more than this grim outlook. For all its imperfections, structured metadata gave us a fresh set of tools that at least got us started: tools in the form of:

- **vocabulary**: terms like edition, title, extent, mode of issuance, which enhanced our awareness, both as intermediaries and as end users, of the processes involved in information production;
- **relationships**: classes, facets, taxonomies and ontologies, which for all their limitations and problems have formed the basis of many information architecture environments;
- **purpose**: recognition that we need to define, and refine, the purposes of our metadata systems, as a means of directing our choices and our practices.

Structured metadata did not "save us" from the complexity of knowledge organization, but structured metadata helped us negotiate that complexity. And sometimes, we did rather well at it.

Structured metadata may not save us from the threat of AI, but once again, the practice of reasoning out the model card paradigm provides us with:

- **vocabulary**: we have terms we can use to examine the existence and the importance of phenomena such as training data, out-of-scope use, decision thresholds, intersectional results.
- **relationships**: we are getting used to classification of risk factors (Kennedy-Mayo and Gord 2024) and taxonomies of danger, as specified in the European Union Artificial Intelligence Act (European Union 2024).
- **purpose**: we have terms that articulate, and remind us of, the underlying objectives of AI systems, including use, factors, metrics, evaluation, analysis and ethics.

Whatever the metadata does with these terms, relationships and purposes, and whatever we do with the metadata, the model card system has the potential to educate us by articulating these things with greater precision. In so doing, the system might turn our apocalyptic nightmares of human extinction into tangible concepts and paradigms that frame threats and challenges in the daylight.

Many of us, when young, had parents who tried to help us. Their advice and experience helped us part of the way; we then left them behind and forged our own way forward. But every now and then, we are astonished to hear our parents speaking out of our own mouths. AI systems have gone far beyond the systems that produced our history and traditions of bibliographic control and knowledge organization. But every now and then, we hear knowledge organization coming out of the mouths of systems that are leaving us far behind. Structured metadata. On cards, no less. Who would have thought?

AI Usage Statement: No Artificial Intelligence tool was used in the research and writing of this paper.

References

Anthropic. 2026. Claude's Constitution. Anthropic.

Crisan, Anamaria, Margaret Drouhard, Jesse Vig, and Nazneen Rajani. 2022. "Interactive Model Cards: A Human-Centered Approach to Model Documentation." In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, 427-39. New York: Association for Computing Machinery. DOI: 10.1145/3531146.3533108.

European Union. 2024. Artificial Intelligence Act (Regulation (EU) 2024/1689). Official Journal of the European Union L, July 12, 2024.

Kennedy-Mayo, DeBrae, and Jake Gord. 2024. "'Model Cards for Model Reporting' in 2024: Reclassifying Category of Ethical Considerations in Terms of Trustworthiness and Risk Management." arXiv preprint arXiv:2403.15394. DOI: 10.48550/arXiv.2403.15394.

Liang, Weixin, Nazneen Rajani, Xinyu Yang, Ezinwanne Ozoani, Eric Wu, Yiqun Chen, Daniel Scott Smith, and James Zou. 2024. “What’s Documented in AI? Systematic Analysis of 32K AI Model Cards.” arXiv preprint arXiv:2402.05160. DOI: 10.48550/arXiv.2402.05160.

Mehraj, Ali, An Cao, Kari Systä, Tommi Mikkonen, Pyry Kotilainen, David Hästbacka, and Niko Mäkitalo. 2025. “AI Model Cards: State of the Art and Path to Automated Use.” In Proceedings of the 21st International Conference on Web Information Systems and Technologies, vol. 1, edited by Karl Aberer, Josep Domingo-Ferrer, and Massimo Marchiori, 71-82. SciTePress. DOI: 10.5220/0013706600003985.

Mitchell, Margaret, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. “Model Cards for Model Reporting.” In Proceedings of the Conference on Fairness, Accountability, and Transparency, 220-29. New York: Association for Computing Machinery. DOI: 10.1145/3287560.3287596.

Puhlfürß, Tim, Julia Butzke, and Walid Maalej. 2025. “Model Cards Revisited: Bridging the Gap Between Theory and Practice for Ethical AI Requirements.” arXiv preprint arXiv:2507.06014. DOI: 10.48550/arXiv.2507.06014.

Shen, Hong, Leijie Wang, Wesley H. Deng, Ciell Brusse, Ronald Velgersdijk, and Haiyi Zhu. 2022. “The Model Card Authoring Toolkit: Toward Community-Centered, Deliberation-Driven AI Design.” In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, 440-51. New York: Association for Computing Machinery. DOI: 10.1145/3531146.3533110.

Suleyman, Mustafa, with Michael Bhaskar. 2025. The Coming Wave: AI, Power, and Our Future. Crown Trade Paperback ed. New York: Crown. Originally published 2023.