

# Think big – act smart: thoughts on the future of the Art Discovery Group Catalogue

Jan Simane

The Art Discovery Group Catalogue is a virtual bibliographic resource based on the catalogues of over 70 art and museum libraries. Its future is closely linked to the question of technically and topically innovative data concepts on the one hand and to new retrieval and analysis procedures on the other. Will it be possible to interpret both the quantity and the diversity of the data in a more future-oriented and efficient way for scientific questions than traditional catalogues do?

The intention of this article is not to tell the story of the network operating under the name 'artdiscovery.net' and merely append a few closing sentences with an eye to the future, but quite the opposite. Almost 25 years since the founding of the 'Virtual Catalogue of Art History'<sup>1</sup> and after ten years of cooperation with OCLC<sup>2</sup>, it is time to take stock and ask ourselves where do we go from here? In doing so, one can note with satisfaction that the initiative has been a great success and that the largest international network of art and museum libraries ever has been assembled. Let us start with a view of its main characteristics. This operational and collaborative network comprises 72 art and museum libraries, as well as art departments of larger university libraries, from 20 countries spread across four continents. On an operational level, these libraries run a collective virtual discipline-oriented catalogue, the Art Discovery Group Catalogue.<sup>3</sup> As cooperation partners, they convene at regularly organised conferences and maintain informal communications. It is noteworthy that this network has no institutional format, that there is neither an administrative structure nor a financial basis, and that all basic decisions are made democratically. The only management body is a Steering Committee, consisting of nine volunteers from six countries. Membership of artdiscovery.net is voluntary and informal, and it entails no financial obligation. The only costs are those that relate to the operation of the common catalogue, which brings us to the essential element of the initiative, the Art Discovery Group Catalogue.

The network is primarily perceived in the form of this jointly operated bibliographic research platform. OCLC WorldCat has played the lead role technically and also logistically since 2012. Although WorldCat is designed to process large amounts of data, it also has a high degree of functional flexibility. This includes the possibility of filtering out specific data and displaying it separately without reducing the operational potential. And this is the basis for the Art Discovery Group Catalogue. It is a filtered view in WorldCat of the holdings of the art and museum libraries in the network. What are the benefits? First of all, a search in the group catalogue provides a concentrated, one could even say 'undisturbed', view of the holdings of the libraries united in the network with their subject-specific, and in part very special, catalogue data. They refer, in addition to 'normal' bibliographic data, to special collections, remains, image archives and the like. The latter is due to the fact that the group catalogue covers holdings of numerous libraries that otherwise would not have sought to connect with WorldCat. In addition, special resources have been integrated in cooperation with OCLC, such as the "Scipio" database, an index to 300,000 auction catalogues.<sup>4</sup> It is also of great importance that a high proportion of articles from journals and collected

1. Rüdiger Hoyer, "The "Virtuelle Katalog Kunstgeschichte" (VKK): Current Status and Future Development", *AKMB-News* 9, no. 2 (2003): 15–16.

2. Jan Simane, "Der neue Art Discovery Group Catalogue: Anmerkungen zum Paradigmenwechsel im internationalen Kooperationsnetzwerk der Kunst- und Museumsbibliotheken", *AKMB-News* 21, no. 1 (2015): 3–9; Geert-Jan Koot, "The Art Discovery Group Catalogue: a New and Freely-available Tool for Art Historical Bibliographic Research", *Art Libraries Journal* 40, no. 3 (2015): 41–48.

3. <https://artlibraries.on.worldcat.org/discovery>.

4. Kay Downey, "Not my father's auction catalog database! SCIPIO, 1980 to 2003", *Art Libraries Journal* 29, no. 2 (2004): 17–22.

writings have found their way into the group catalogue. This is where WorldCat's Central Index plays a major role, as it enables the retrieval of references especially for articles that cannot be found in most of the local catalogues of the partner libraries, even though the journals themselves may well belong to the holdings of these libraries. For all these functionalities two insights were essential. First, the potential of the formation of digital networks and virtual sources – in other words data aggregation – for discipline-specific requirements (around the year 2000), and second, the group catalogue's quality as an alternative to the traditional subject bibliography. The constitutive and also formal analogies of library catalogues and bibliographies have been emphasized many times.<sup>5</sup> The similarities were particularly pronounced as long as bibliographies were printed and catalogues were organised as card indexes. Both resources were usually subject to strict and long-term fixed patterns of order, both in terms of the categorisation and hierarchisation of knowledge (systematics) and the required search strategy and orientation (alphabet, chronology). Under these conditions, the accidentality of information acquisition, which is indispensable for the research process, was mainly guaranteed at the shelf, where content 'browsing' took place in its early form. All this changed radically, if not immediately, with the establishment of electronic catalogues and bibliographic online databases. Equally, however, the rapid growth of primarily digital resources and their accessibility, with the permanent refinement of search tools, has increasingly challenged the traditional model of the intellectually controlled subject bibliography compiled at great human and logistical expense.<sup>6</sup> However, similar to catalogues, (digital) bibliographies have experienced an analogue development from printed indexes to electronic databases, especially with regard to retrieval comfort, although here the selection of literature and resources is of course much more limited, but at the same time more consistent and controlled than in the case of library catalogues. In the latter, the mapping of the literature landscape does still follow internationally agreed rules, but the formally pre-determined navigation system with pre-marked paths has been replaced by a far more intuitive, cross-area discovery that allows the user to leave the pre-assigned routes and also integrates resources that go far beyond the actual holdings and can even exceed them significantly in terms of quantity.

So, what does this mean for the Art Discovery Group Catalogue, which was, after all, considered an alternative to the bibliography? The decisive difference to the bibliography is not so much the proper search strategy, but the calibratability of the search space which in the case of the Art Discovery Group Catalogue can be extended to the entire WorldCat. However, also without this extension, the network building of art libraries catalogues has been from the very beginning something like a tactical process of generating a virtual big data source. Self-evidently, the data collection itself is strongly determined by the group identity. When the 'Virtual catalogue of art history', the precursor of the Art Discovery Group Catalogue, was founded at the end of the last century, the accumulation of data, whether in shared databases or virtually through federated search procedures, was the state of the art. Consequently, the profile of the data source, in this case the selection of partnering libraries, was the decisive criterion, for which the Steering Committee pursued a well-thought-out implementation strategy. Selection has been a core principle for many years and a reminiscence of bibliographies. This concept no longer plays a role today, as the idea of a firmly defined group profile has largely been abandoned and only professional involvement in the field of art history and the art museum in the broadest sense is a sufficient admission criterion. Likewise, a conscious opening to neighbouring disciplines such as classical archaeology is becoming apparent, not least due to the fact that quite a few member libraries already house corresponding collection segments. These changes have both a topical and a technological background. In the first case, and here we come back to the bibliography discourse, the question arises as to how a discipline-specific identity can be determined for the visual arts in the context of a global dimension and a multicultural history with fluid geographical and historical boundaries. Should a decidedly transcultural and transdisciplinary view, as it is increasingly establishing itself in art research, also be demanded of the most important working tools, the libraries? For an approach to such a goal, the size of the group and also the increasing diversity of member libraries has tangible advantages. A single library would never be able to meet this demand alone. Together with the many, partly unique, special collections and the

5. Jan Simane, "The 'Crisis' of Art Bibliography", *Art Libraries Journal* 36, no. 3 (2011): 5–9.

6. Carlo Bianchini, Fiammetta Sabba, Lucia Sardo, "Proposal for a New Methodological Approach to the Study of 20th Century Bibliography", *Bibliothecae.it*, 10, no.2 (2021): 138–151, <https://doi.org/10.6092/issn.2283-9364/14051>.

documentation in particular of the mostly-neglected article literature, this network holds enormous potential.

Accordingly, the main innovations are to be seen in the amount of data available as well as in the far more comprehensive access to this data. Another specific feature is the heterogeneous quality, sometimes paired with language diversity, of the aggregated data. Thus, quite different concepts are required for retrieval than from uniform bibliographies or catalogues of individual libraries. In other words, with the opening of access to new resources and the demand to make relevant discoveries, the problematic nature of the traditional catalogue concept has become evident. Certainly, there are integrated concordances and hybrid systems that contribute to a unification of, for example, authorities. The greatest deficit of such efforts, however, is in the area of content description, which is crucial if reliable navigation is required through the vastly growing and intellectually unmanageable number of documents that are still being catalogued to a significant extent. Two fundamental questions - certainly not raised for the first time - result from these findings: firstly, whether, in view of the quantity of accessible documents and resources, a far more precise and powerful instrument is required for their detection, interpretation, and evaluation, and secondly, whether, at the same time, traditional subject indexing with controlled vocabulary and an unwieldy set of rules can be the appropriate answer to the above-mentioned demands in view of their minor importance for the observed retrieval behaviour.<sup>7</sup>

Despite all the technical improvement and conceptual development of library catalogues and similarly organized bibliographies, both nevertheless follow a traditional ontological principle that can literally be called meta-data existence, i.e. an 'otherworldly' realm in which descriptive surrogates for entities of a different kind are collected and structured. The fact that more and more of these entities are also available and accessible as data calls into question the sense of separation, not at all suggesting that catalogues could become obsolete. But they can and must change.<sup>8</sup> Some steps have already been taken in this direction. This is to be seen in context with the technical evolution in librarianship and the increasing role of catalogues as complements in a linked and interoperable data universe, experiencing exponential growth. The Linked Open Data paradigm, for example, roots in a machine-based analysis of semantic relationships of data and a resulting automatic processability, domain-independent and across distinct data spaces. From the perspective of technology and library science, we are simultaneously observing a transfer of the metadata concept, which is changing from a structured work description based on cataloguing rules and controlled vocabularies to identity or entity management according to the Linked Open Data paradigm with the goal of universal interoperability.<sup>9</sup> Occasionally, MaaS (Metadata as a Service) is also spoken of in this context, referring on the one hand to the orientation of library metadata to sectors outside the library world and on the other hand to the operative-functional role of catalogue data as active nodes in a linked system of ontologically coded knowledge correlations.<sup>10</sup> For all these considerations, the connection of artdiscovery.net to WorldCat is of crucial importance. OCLC has been investing in corresponding research for several years in order to support libraries with know-how and robust infrastructure to achieve this transfer. Recently, OCLC started to develop a Shared Entity Management Infrastructure. Structured as a central repository, bibliographic metadata – represented as Linked Data – will be curated and their potential for interoperability enhanced. In recent years, the widely and controversially discussed phenomenon of artificial intelligence has been brought into play in connection with the analysis and interpretation of large amounts of data. Libraries play an important role here, as they are producers of not only very extensive, but also well-structured and highly qualitative data.<sup>11</sup> In addition to accessibility and discovery, the traditional metadata principles, now linking or data flow is added as a further characteristic. In the context of library catalogues and the Art Discovery Group Catalogue in particular, two insights can be derived. On the one hand, the (growing) amount of high-quality data and its diverse contexts on numerous fields of knowledge has a considerable heuristic potential, which machine-based semantic analyses can increase considerably. Secondly, it seems obvious that automated analysis and selection procedures for identifying relevant information meet the quality and comfort demands of the searching individual far more than the required 'deciphering' of terminologies from established bibliographic metadata.

7. For the complexity and also the shortcomings of traditional subject indexing see Jens-Erik Mai, "Deconstructing the Indexing Process", *Advances in Librarianship* 23, (2000): 269–298, in particular 272–277. Already in the pre-internet era in 1992 surveys on search behaviour related to databases revealed that thesauri (i.e. controlled vocabularies) are steadily neglected by users as the number of accessed databases increases. See Raya Fidel, "Who needs Controlled Vocabulary?" *Special Libraries* 83, no.1 (1992): 1–9, in particular 6.
8. Richard Gartner, *Metadata: Shaping Knowledge from Antiquity to the Semantic Web* (Cham: Springer, 2016) 93–96.
9. Susan Brown, "Same Difference: Identity and Diversity in Linked Open Cultural Data," *International Journal of Humanities and Arts Computing* 16, no. 1 (2022): 1–16.
10. Karen Smith-Yoshimura, *Transitioning to the Next Generation of Metadata*, (Dublin, OH: OCLC Research, 2020), 25–28.

11. Andrew M. Cox, Stephen Pinfield, Sophie Rutter, "The intelligent library: Thought leaders' views on the likely impact of artificial intelligence on academic libraries," *Library Hi Tech* 37, no. 3, (2019): 418–435.

Certainly, these are still, in part, scenarios for the future, but the transformation process has long since arrived in the present. And the art and museum libraries involved in the Art Discovery Group Catalogue are fortunate to have found a powerful partner in OCLC and WorldCat, which is at the forefront of this development. OCLC is active in both the definition of the required data models and the programming of the associated tools. Moreover, it would not be the first time that institutions from the cultural sector have been identified as particularly suitable and therefore welcome data providers in the context of the Linked Open Data paradigm. Those libraries that are members of the Art Discovery Group Catalogue are explicitly characterised by the fact that their catalogues contain information and data with numerous content links to neighbouring and non-library environments. They describe not only books and articles, but also objects and works, places and persons that are relevant in other, content-correlating contexts. As is well known, this is the initial finding underlying the Linked Open Data model. At the same time, it was recognised and emphasised years ago that, complementary to the development of data models and infrastructure, cataloguing practice in libraries must also change if the machine-assisted identification of semantic relations and the linking of data based on these relations are to achieve the desired development.<sup>12</sup> Or, in other words, as Karen Coyle put it back in 2013: “Our job today, as librarians and information scientists, is not to translate library data to linked data; our job is to create a new system for access and use of bibliographic data that is compatible and works within the web”.<sup>13</sup> Decisive steps in library policy were already taken with the introduction of RDA on the one hand and the establishment of the BIBFRAME ontology on the other and have significantly influenced metadata production in libraries worldwide.

This all sounds quite promising, but we are certainly far from having arrived in the ‘brave new world’ of Linked Data. Even the somewhat old-fashioned ‘manual’ interoperability of catalogue data, namely the simple accumulation, exchange and re-use of bibliographic descriptions, has been a focus of international library policy for a generation, although we are some way short of a satisfactory result. For example, redundancy in the worldwide cataloguing of publications is still high. This problem becomes all the more apparent when, as in the case of the Art Discovery Group Catalogue, catalogues of specialised libraries closely related in subject matter are merged into one overall data pool. From the point of view of library policy, it is regrettable that the time spent on redundant routines is not invested more purposefully in complementary tasks such as the cataloguing of article literature, the indexing of OA publications or in-depth subject indexing. Even if this may sound utopian at present, a corresponding coordination in the distribution of tasks in the network of art and museum libraries would be a worthwhile attempt. The result would be a significant enrichment of metadata production for the art sector. Furthermore, should the aforementioned adaptations and shifts concerning the principles of the Semantic Web and Linked Data come to fruition – here, the art libraries are not autonomous in all cases – and should a more intensive cooperation with OCLC also become possible, then the largest international network of art and museum libraries could develop into an element to be taken seriously in the ecosystem of the Digital Humanities. Admittedly, there is still a lot of wishful thinking involved here. There are also still a number of formal, logistical, institutional and probably also financial hurdles to overcome. However, this should not be a reason to relinquish all of these prospects as unrealistic. One fundamental insight must be recalled in this context: if we had not dreamed big and acted thoughtfully a good 20 years ago, the international art discovery network and the Group Catalogue would not even exist.

12. Nazia Wahid, Nosheen Fatima Warraich, Muzammil Tahira, “Mapping the Cataloguing Practices in Information Environment: A Review of Linked Data Challenges,” *Information and Learning Science* 119, no. 9/10, (2018): 586–596.

13. Karen Coyle, “Library linked data: an evolution,” *JLIS.it: Italian journal of Library Science* 4, no.1 (2013): 57.

---

Jan Simane  
 Library Director  
 Kunsthistorisches Institut in Florenz, Max-Planck-Institut,  
 Via Giuseppe Giusti 44,  
 50121 Firenze,  
 Italy  
 Email: [simane@khi.fi.it](mailto:simane@khi.fi.it)