

ProvenanceFinder

Preparing a search engine for the retrieval of provenance data

Abstract *ProvenanceFinder is designed to be an efficient tool of research with free access for everyone who wants to search internet databases for information on provenance. For the time being, twelve online catalogues in libraries, archives and museums have been designated as target databases. A virtual catalogue comprising these databases which have been analysed regarding contents and structure of provenance entries will be developed. As a research tool a multilingual thesaurus of provenance terms will be offered. In cooperation with the Deutsche Nationalbibliothek, the Bibliothèque municipale de Lyon and the Koordinierungsstelle für Kulturgutverluste Magdeburg, there are plans to promote and support the feasibility, the public awareness and the internationalisation of the project beyond the area of libraries. The project has not started yet.*

ProvenanceFinder ist ein effizientes Recherchetool mit freiem Zugang für alle, die in Internet-Datenbanken nach Provenienzinformationen suchen. Als Zieldatenbanken wurden vorerst zwölf Onlinekataloge in Bibliotheken, Archiven und Museen bestimmt. Auf der Basis dieser Datenbanken, die auf Inhalt und Struktur der Provenienzeinträge hin analysiert wurden, wird ein virtueller Katalog entwickelt. Als Rechercheinstrument wird ein mehrsprachiger Thesaurus der Provenienzbegriffe angeboten. In Zusammenarbeit mit der Deutschen Nationalbibliothek, der Bibliothèque municipale de Lyon und der Koordinierungsstelle für Kulturgutverluste Magdeburg ist geplant, die Praxistauglichkeit, öffentliche Wirksamkeit und Internationalisierung des Projekts über den Bibliotheksbereich hinaus zu fördern und zu unterstützen. Das Projekt ist noch nicht angelaufen.

Why provenance research now?

Provenance research deals with the origin and the previous ownership of copies, but also with copy-specific use: marginal notes, dedications, and signs of censorship. In Germany, this kind of research has gained significance over the past years in a twofold way.

Collections, as well as copies, contain a multitude of contextual information on their origin, tradition and their use, that has so far attracted little attention in the librarian practice. As a result, this information has quite often been hidden or even lost. Once this information is lost, the same thing will happen to important provenance data. As a rule, we can find this kind of data in accession books and archive records, as well as in the books themselves in the form of labels, stamps, handwritten notes or insertions of all kinds.

Bernhard Fabian's *Handbuch der historischen Buchbestände in Deutschland* has drawn attention to the significance of such contextual information and has consequently changed our approach to collections and copies. The *Handbuch* regards libraries, archives and museums as interfaces and hubs of collections, whose structures give us information about the various regional connections of such institutions.¹ Provenance research makes these connections visible and can thus be considered an important issue of stock exploitation.

The origin and profiles of use of individual copies and entire collections are an important tool for historical and literary studies, for text editions, literary sociology, art history and bindings research. In libraries these profiles serve as a selective criterion for exhibitions and book presentations and they are indispensable when it comes to the assessment of preservative measures. In a broader sense, antiquarians, auction houses, genealogists and bibliophiles also benefit from provenance information.

In Germany, provenance research has been attracting broader public and political interest thanks to the Initiative *Lost Art Internet Database* of the Koordinierungsstelle für Kulturgutverluste Magdeburg, Germany's central office for the documentation of lost cultural property (<https://www.lostart.de/de/start>). The Koordinierungsstelle has been collecting data on the losses and the changes in ownership of cultural property, that was lost due to prosecution during the National-Socialist rule or due to events during the Second World War. Provenance research has thus increasingly become a field of research where experts and politicians meet directly and benefit from each other's information. These days, the preparation and implementation of restitution projects in Germany make provenance research a topic in which not only

librarians with a preference for historical research are interested. Rather, it is becoming obvious that provenance research must refer to any form of cultural property in libraries and that it is a task with a potential for a sudden increase in political and moral significance.

At present, the most important task is to provide access to already existing provenance information and to link it up. It is the retrieval of and the connection to such information that ProvenanceFinder aims to provide. For a better understanding of its operating processes, the method and standards of previous provenance descriptions as well as the techniques and tools used in ProvenanceFinder research are explained below.

Provenance marks: description and access

David Pearson's reference handbook *Provenance research in book history*² is undoubtedly one of the milestones of bibliographical provenance research. According to Pearson, the term provenance is not confined to origin and previous ownership, i.e. the legal relationship of the owner and his copy, it also includes the use of copies. It is against this background that each single copy represents something that it has not yet been seen from the point of view of bibliographical stock exploitation: a physical object and artefact, that – in addition to its format and its size – can also be characterised by a number of individual features. With regard to provenance research, these features, called »evidences«, supply clues to the use and function of the individual copy. The German thesaurus of provenance terms³ enumerates more than 60 terms for such evidences, the American equivalent⁴ nearly 100.

We find various types of evidences on the binding, the edges, on the inside covers of the case and, of course, in the text block of the copies, either in connection with the printed text or on interleaved sheets, and quite often on inlaying letters, photographs and newspaper clippings. There are two groups of features that help us reconstruct the origin and profiles of use of individual copies: (1) direct marks of ownership, e.g. autographs, ex libris, coats of arms; and (2) traces of reading and using copies that supply clues to previous owners and can help identify him or her, e.g. annotations, marginal notes, blackened marks and erasures.

These two groups of features give hints to the type of the copies, i.e. the function of the copies for the previous owner, e.g. withdrawn copies and presentation copies. From the point of view of provenance research, these asso-

ciation copies are the most significant for the identification of the book's intrinsic value because it tells us something about the non-material relationship between previous owner and copy.

However, provenance information is, in general, not easily accessible. There are two reasons for this situation: Until few years ago, there were hardly any bibliographical standards for the description of such copy-specific features, in contrast to bibliographic description and subject indexing. Moreover, even if such information was retrieved and noted as free text in the note area of card or data files, there was no access or research tool to find it.

Consequently, the only hints for the documentation of provenance information could be found, if at all, in bibliographical codes and instructions for the drawing up of author or alphabetical catalogues. Until recently, there were no specific guidelines for the description of copy-specific features. The current editions of the *Anglo-American Cataloguing Rules* (AACR-2) and the *Regeln für die alphabetische Katalogisierung* (RAK-WB) are not appropriately suited for the demands of copy-specific description either. Copy-specific data are normally documented only casually and are not suited for retrieval. The term »provenance« was introduced in the AACR-2 in 1978, and in the RAK-WB as late as 1993 and this was done rather casually at that.⁵

Neither the old card-based librarian catalogues nor the new EDP-based ones are suited to the representation and communication of copy-specific features of books and other library material (fig. 1).⁶

The handling of the notes (whether general or specific), i.e. entries at the bottom of index cards (the note area), is but one example of this problem: On the catalogue cards reproduced in the *British Library general catalogue of printed books to 1975* (1982), copy-specific data are at the end of the actual bibliographic title description, highlighted by a different typography. Sometimes they give information on previous owners as well as on notes and marginalia of the »writing reader«. If, during the conversion of the catalogue cards, such entries are transferred to the electronic data record without further qualification, the research of provenance information amounts to the search for a needle in a haystack. Entries that have been drawn up, in German, »nach Vorlage« or that have been standardised only in part (»With MS. corrections by the author.«) can only be evaluated through search requests that combine various searchable fields in a sophisticated way and use tools such as multilingual thesauri with time-consuming routines of synonym- and abbreviation retrieval.

Fig. 1: »With MS. corrections by the author.«

HUME (DAVID) <i>the Historian.</i> — A treatise of human nature: being an attempt to introduce the experimental methode of reasoning into moral subjects, etc. [By David Hume.] London: John Noon, 1739, 40. 1133.c.4-6. 3 vol. 8". Δ <i>MS. notes by George Chalmers.</i> — [Another copy of vol. 3.] A treatise of human nature, etc. [By David Hume.] 1740. C.175.c.8. Δ <i>MS. notes by the author.</i>	Title: A treatise of human nature: being an attempt to introduce the experimental method of reasoning into moral subjects, etc. [By David Hume.] Author: David Hume, 1711-1776. Publication Details: London : John Noon, 1739, 40. Notes: MS. notes by George Chalmers. Vol. 3 was printed for Thomas Longman. Physical Description: 3 volumes ; (8") Holdings Note: General Reference Collection C.175.c.8.(1.) [Another copy of vol. 3.] A treatise of human nature, etc. [By David Hume.] 1740. ***MS. notes by the author. Obsolete shelfmark: C.175.c.8. Shelfmark(s): General Reference Collection 1133.c.4-6. General Reference Collection C.175.c.8.(1.)
vol. 155, p. 264	<http://explore.bl.uk/BLVU1:LSCOP-ALL:BL01001765099>
— An abstract of a book lately published; entitled, A treatise of human nature, &c. Wherein the chief argument of that book is farther illustrated and explained. London: C. Borbet [sic], 1740. C.175.c.8.(2.) pp. 32. 8". □□ An anonymous abstract, by David Hume, of his "Treatise of human nature". With MS. corrections by the author.	Title: An abstract of a book lately published; entitled, A treatise of human nature, &c. Wherein the chief argument of that book is farther illustrated and explained. Author: David Hume, 1711-1776. Publication Details: London : C. Borbet [sic], 1740. Notes: An anonymous abstract, by David Hume, of his "Treatise of human nature". With MS. corrections by the author. Physical Description: 32 pages ; (8") Shelfmark(s): General Reference Collection C.175.c.8.(2.)
vol. 155, p. 264	<http://explore.bl.uk/BLVU1:LSCOP-ALL:BL01001765111>
The British Library general catalogue of printed books to 1975, Saur: London / München / New York / Paris, 1982.	British Library, online main catalogue, <https://www.bl.uk/catalogues-and-collections/catalogues>.

When converting data from the British Library general catalogue of printed books to 1975 into the British Library computerised catalogue system, provenance data are stored in various types of note fields (»Notes«, »Holdings Note«). Figure and detail, redesigned and abridged: Jürgen Weber

The activities of the Association of College and Research Libraries of the American Library Association and others have brought about some progress in the field of copy-specific description since the middle of the 1980s. Part of this progress has been the development of various thesauri⁷, the establishment of specific cataloguing guidelines⁸ and the enlargement of the data format MARC by fields that allow description both at a minimum and a higher level.

Inspired by these initiatives, a model of provenance description has been developed by the Herzogin Anna Amalia Bibliothek under the data format PICA in 1997. This model was published as a »recommendation for the description of provenance« by the Arbeitsgemeinschaft Alte Drucke beim Gemeinsamen Bibliotheksverbund (GBV) in 2003.⁹

Despite some progress,¹⁰ the possibility most frequently used by scholars to learn about previous owners of books is to inquire personally at the respective library. For the time being, it is impossible to carry out a simultaneous re-

search in databases. This situation is even more unfortunate as book and art collections from literary remains have very often been scattered and the reconstruction of this scattered property depends on the very possibility of carrying out a search.

Services offered by ProvenanceFinder

The future ProvenanceFinder is going to offer an accessible tool for the research of this scattered information which is so important for scholarship and for parts of the public. There is a plan to establish a catalogue that is accessible via the Internet and that can be searched specifically for provenances. The catalogue consists of a combination of selected online-catalogues, that enable external users to get free access to provenance information via the respective OPAC. In addition to local library systems, supra-regional catalogues (Manuscripta mediaevalia, Inkunabelkatalog deutscher Bibliotheken (INKA), Lost Art Internet Database) that combine data from further local catalogues, are also taken into consideration.

Consequently, one single search will initiate indirect queries of provenance information from several hundreds catalogues of libraries, archives and museums.

Connected with the search engine a multilingual thesaurus of provenance terms will form the centre, that will specialise in the query of provenance features in designated target databases. The structure and the terms of the thesaurus have been inspired by the thesaurus *Provenance evidence*.¹¹

There will probably be four search fields offered: previous owner, provenance mark, author/creator, free text.

Searches with the fields »previous owner« and »author/creator« will be supported alternatively by databases with authority files for persons and corporate bodies, by adopting certain forms of names from a table with a click on the search field.

Searches with the field »provenance mark« will be supported by the thesaurus by adopting a certain descriptor and synonyms from a table by clicking on the search field. Tests carried out during the project will help determine if queries can be broadened automatically and if so, by how many synonyms and abbreviations can it be broadened without forcing the user to select them specifically.

The combined search with author- and provenance categories makes sense whenever the author is no longer directly named in the provenance entries, but is rather indirectly mentioned (»With MS. corrections by the author.«).

There is a plan to integrate ProvenanceFinder into a portal on provenances. During the preparations of this project a list of links with online resources regarding the subject of provenances has been established (cf. endnote 3). In this way, we hope to facilitate the communication between the future users and the cooperative partners and also to make it possible to carry some initial tests.

Cooperative partners

During the preparations of ProvenanceFinder, the Herzogin Anna Amalia Bibliothek has been collaborating with three partners: The Bibliothèque municipale de Lyon, the Deutsche Nationalbibliothek, and the Koordinierungsstelle für Kulturgutverluste Magdeburg/Lost Art Internet Database.

The Bibliothèque municipale de Lyon has been establishing a provenance database since 2001. In 2002 the libraries in Weimar and in Lyon harmonised the French-German terminology of the thesaurus. For a lot of provenance terms the option of doing research in three languages (English-French-German) was added as early as December 2002.¹² There is a plan to intensify the work on terminology by carrying out further checks during multilingual research of provenance information. The Bibliothèque municipale de Lyon has taken on the task of translating the prospective thesaurus into French.

The Deutsche Nationalbibliothek: There is a plan to support the search for names of previous owners (persons and corporate bodies) with the help of information supplied by the Personennamendatei (PND) and the Gemeinsame Körperschaftsdatei (GKD). It is true that we will find only the names of famous previous owners in these databases, and it is also true that not every target database registers names in accordance with PND or GKD. However, this supply of information can be used as a kind of »dictionary« or research tool in accordance with our needs, adapted to the search engine via ProvenanceFinder. It is important for the Deutsche Nationalbibliothek to check the application of the norm data in a new subject matter and technical context. There is a plan to check specifically to what extent PND and GKD are suitable for the search in heterogeneous data, where names are registered in a standardised as well as in a non-standardised form. Furthermore, we will investigate which possibilities exist to improve the suitability of PND and GKD.

The Koordinierungsstelle für Kulturgutverluste Magdeburg/Lost Art Internet Database now cooperates with more than 400 museums, libraries, archives and other public and private institutions all over the world, as well as with many private individuals. The conception and the use of the portal are expected to benefit from the acquired experience in managing information on inquiries and detection of data on previous owners and on objects. It will also benefit from the possibility to help interested institutions and individuals collecting provenance data. It is planned to cooperate on the following tasks:

(1) Research of provenances registered in a non-standardised way: development and tests of research strategies and their terminology, suitable for finding provenance information, that has been registered heterogeneously in databases and electronic publications.

(2) Development of a minimum standard for the documentation of provenances: There are plans to supply interested colleagues in museums, archives and libraries with tools for the retrieval and the description of provenances for Lost Art Internet Database. This measure also aims to standardise technical terms for the documentation of provenances, e.g. owner's annotations, reading dates, sentence structure of dedications.

It was during the preparation of this project that two colloquiums on provenance research were organised and carried out: In 2003 the Herzogin Anna Amalia Bibliothek organised a colloquium »Provenienzforschung für die Praxis. Dokumentation und Recherche von Provenienzen in Bibliotheken« (Provenance research for the practice. Documentation and research of provenances in libraries) in cooperation with the Initiative Fortbildung für Spezialbibliotheken und verwandte Einrichtungen e. V. and the Koordinierungsstelle für Kulturgutverluste Magdeburg in Weimar. In September 2004, there was a follow-up colloquium »Von der Provenienzforschung zur Restitution geraubten Kulturguts: politischer Wille und praktische Umsetzung« (From provenance research to the restitution of looted art – political intention and its implementation) in Berlin.¹³

Both conferences have shown that provenance research is being pushed forward by the current political debate on the restitution of looted art. Provenance research will be successful and widely accepted only if it is carried out in a cooperative way. ProvenanceFinder will be a significant and essential tool to support the search of provenances, but it will also contribute to the international standardisation of provenance documentation.

Erstveröffentlichung: Der Aufsatz ist unter dem Titel *ProvenanceFinder—preparing a search engine for the retrieval of provenance data* zuerst erschienen in: David J. Shaw (ed.), *Books and their owners: Provenance information and the European cultural heritage. Papers presented on 12 November 2004 at the CERL conference hosted by the National Library of Scotland, Edinburgh, London: Consortium of European Research Libraries, 2005 (= CERL PAPERS, V), S. 71–79.* – Die in der Erstveröffentlichung angegebenen Internetadressen (13. April 2004) wurden am 26. Oktober 2023 aufgerufen und aktualisiert. Die frühere Bezeichnung »Die Deutsche Bibliothek« wurde durch die 2006 eingeführte Bezeichnung »Deutsche Nationalbibliothek« ersetzt. Die Abbildung des Katalogeintrags wurde neu gestaltet.

Anmerkungen

- 1 Bernhard Fabian (ed.), *Handbuch der historischen Buchbestände in Deutschland*, vol. 3, Hildesheim/Zürich/New York: Olms-Weidmann, 1992, preface, p. 10.
- 2 David Pearson, *Provenance research in book history. A handbook*, London: The British Library, 1994.
- 3 Multilingual version under construction: Stiftung Weimarer Klassik und Kunstsammlungen, http://www.weimar-klassik.de/de/haab/prov_tpr_o.html, 2023 online resource no longer available, cf. ProvenienzWiki – Plattform für Provenienzforschung und Provenienzerschließung. T-PRO | Thesaurus der Provenienzbegriffe, processing status: 13 December 2022, https://provenienz.gbv.de/T-PRO_Thesaurus_der_Provenienzbegriffe.
- 4 Provenance evidence. Thesaurus for use in rare book and special collections cataloguing. Prepared by the Standards Committee of the Rare Books and Manuscripts Section (ALA/ACRL), Chicago: Association of College and Research Libraries, 1988.
- 5 Michael Gorman/Paul W. Winkler (ed.), *Anglo-American cataloguing rules*, prepared by American Library Association/British Library/Canadian Committee on Cataloguing/Library Association/Library of Congress, 2. ed., London: Library Association, 1978 (1. ed. 1967), AACR-2. In the index, the keyword »Provenance (early printed monographs)« refers to 2.18F with references to the description of the copy in hand in the »note area«:

»Copy being described | Make notes on special features of the copy in hand. These include rubrication, illumination, and other hand colouring, manuscript additions, binding (if noteworthy), provenance, and imperfections.« – Commentary 2023: The statement that the concept of provenance was »introduced in the AACR-2 in 1978« should be qualified, since the provisions from 2.18F of the AACR-2 can already be found in comparable formulations in 185D of the 1. edition of the AACR of the London edition (Anglo-American cataloguing rules, prepared by American Library Association/Library of Congress/Library Association/Canadian Library Association, British Text, London: Library Association, 1967) and in 184E of the Chicago edition (Anglo-American cataloging rules [...], North American text, Chicago: American Library Association, 1967). In it, the description of provenance marks is broken down according to different types of media (incunabula, manuscripts collections, pictures), and for the description of pictorial material it then says in 272C5 with reference to Frits Lugt, *Les Marques de Collection de Dessins & d'Estampes* (1921, Supplement 1956, as online resource 2010 of Fondation Custodia, Collections Frits Lugt, Paris, <http://www.marquesdecollections.fr/>): »Provenance. The names of previous owners may be given if known and if useful for identification. [...] Previous collectors' marks on a print are noted in inverse chronological order, with Lugt designations.« Frits Lugt's catalogue has a tableau of provenance features (e.g. initiales, numéro, paraphe), which are indexed according to visual search criteria. – Kommission des Deutschen Bibliotheksinstituts für Erschließung und Katalogmanagement (ed.), *Regeln für die alphabetische Katalogisierung in wissenschaftlichen Bibliotheken*, erarbeitet von der Kommission des Deutschen Bibliotheksinstituts für Alphabetische Katalogisierung (bis 1990) und der Expertengruppe RAK des Deutschen Bibliotheksinstituts (seit 1991), red. Hans Popst, 2., überarbeitete Ausg., Berlin: Deutsches Bibliotheksinstitut, 1993- (1. Ausg. 1983), RAK-WB. In the out-of-the-way place 2.6.2.6.4.4. on the »Angaben zum vorliegenden Exemplar«, § 163a states: » [...] 2. Weitere Angaben zum vorliegenden Exemplar können gemacht werden, wenn es für zweckmäßig gehalten wird. Diese Angaben betreffen z.B. Unvollständigkeit (fehlende Blätter), Einbandart, Erhaltungszustand, Kolorierung, handschriftliche Vermerke einschließlich Widmungen, Provenienzen, alte Signaturen, Bücherpreise, numerierte Exemplare, in Zentimetern gemessenes Format.«

- 6 Jürgen Weber, »The copy in hand«. Ziele und Voraussetzungen exemplar-spezifischer Erschließung, in: Bibliotheksdienst, 36, 2002, 5, p. 614–624.
- 7 Thesauri for use in rare book and special collections cataloguing, prepared by the Standards Committee of the Rare Books and Manuscripts Section (ACRL/ALA): Printing and publishing evidence (1986), Binding terms (1988), Provenance evidence (1988), Paper terms (1990), Type evidence (1990).
- 8 The Library Association Rare Books Group, Guidelines for the cataloguing of rare books, London: The Group, 1997.
- 9 Arbeitsgemeinschaft Alte Drucke beim Gemeinsamen Bibliotheksverbund (GBV), <http://aad.gbv.de/empfehl.htm>, 2023 online resource no longer available, cf. https://provenienz.gbv.de/DBV-Empfehlungen_zur_Provenienzverzeichnis.
- 10 Monique Hulvey, Provenance description and access, in: CERL Newsletter, 7, July 2003, p. 3–4, https://www.cerl.org/_media/publications/newsletterjuly03.pdf. – Jürgen Weber, Provenance information: steps towards cooperative recording and indexing, in: IFLA Newsletter. Section on Rare Books and Manuscripts, Winter 2003, p. 10–13, <https://archive.ifla.org/VII/s18/pubs/winter03.pdf>. – Jürgen Weber, Kooperative Provenienzer-schließung, in: Zeitschrift für Bibliothekswesen und Bibliographie, 51, 2004, 4, p. 239–245. – The description of provenance can be funded along with other cataloguing projects by Deutsche Forschungsgemeinschaft (DFG): Aus der Deutschen Forschungsgemeinschaft, Förderungsgrundsätze im Programm »Erschließung von Spezialbeständen«, in: Zeitschrift für Bibliothekswesen und Bibliographie, 48, 2001, 5, p. 281–286, p. 283.
- 11 Provenance evidence. Thesaurus for use in rare book and special collections cataloguing. Prepared by the Standards Committee of the Rare Books and Manuscripts Section (ALA/ACRL), Chicago 1988. – Deborah J. Leslie, Folger Shakespeare Library Washington, has offered the support of the Bibliographic Standards Committee of the Rare Books and Manuscripts Section for the works on the thesaurus.
- 12 Bibliothèque municipale de Lyon, Numelyo, <https://numelyo.bm-lyon.fr/plansite.php>.
- 13 Initiative Fortbildung für wissenschaftliche Spezialbibliotheken und verwandte Einrichtungen e. V., texts and reports 2003 and 2004: <https://www.initiativefortbildung.de/html/schlaglichter/schlaglichter.html>.

