



**Bibliographic and universal access:
Historic sources and fictional works with FRBR**

Frank Förster
MA, MLIS
Christian-Albrechts-Universität Kiel,
Kiel, Germany

Meeting: 93. Cataloguing

WORLD LIBRARY AND INFORMATION CONGRESS: 76TH IFLA GENERAL CONFERENCE AND ASSEMBLY
10-15 August 2010, Gothenburg, Sweden
<http://www.ifla.org/en/ifla76>

Abstract:

The continued development of the bibliographic data model FRBR_{oo} is on the agenda of the Strategic Plan of the IFLA Cataloguing Section. As part of CIDOC Conceptual Reference Models (CIDOC CRM; ISO 21127:2006, latest version: v.5.0.2.) differently structured data from the domain of the cultural heritage can be integrated and exchanged; bibliographic information will be integrated by FRBR_{oo}. Due to visible changes in publication practices former library catalogs with their bibliographic fixation on the book as the main indexing element are no longer up to date. A comprehensive index of the cultural heritage together with museums and archives and their “objects” and such an ontologically based representation of the completely digital publication diversity in the near future in collaboratively cultivated catalogs of academic libraries only succeeds when on the one hand the single text as a “bibliographic object” including historic sources, fictional works, as well as scholarly essays is indexed formally and subjectively and on the other hand the event is considered as the starting point of indexing work. Two conclusions result from this: 1) the semantic levels of knowledge representation according to Erwin Panofsky should be expanded to a fourth level designated as “acrossness” and 2) the task of the academic librarian should consist of an intellectual management of new publications and also of an intellectually thorough comprehension of one’s own subject area.

Half a year ago I purchased a book in Helsinki on the cover of which was the title *Mörkrets Hjärta* (Heart of Darkness), the author Joseph Conrad, the publisher Lindelöw Göteborg, the translator Einar Heckscher. It is a Swedish edition with two stories of the British author Joseph Conrad, letters and other texts or text fragments ; also there are four photos, a rather long afterword, and three short descriptions. The book was issued in 2009. Surprisingly on this book I found the following: it refers purely in its form—as a book—back

to the tradition of the printed work, in its contents composition however it refers to hypertextual thought and forward strategies as they could correspond to the reading patterns of the younger generation; it is a type of semantic-web-environment in miniature. I did not say that it was ahead of its time for this reading pattern, it is not just ahead of its time, but it shows much more, how inherent connected and linked thinking, web-based interaction, is to publishing.

The aforementioned book contains the thematically connected stories *Mörkrets Hjärta* (Heart of Darkness) and *Pionjärerna* (An Outpost of Progress). They are thematically connected because in both of them the colonial pillage of the former Belgian kingdom in the Congo at the end of the 19th century is pointed out and represented fictionally; then the book contains the foreword of Joseph Conrad to his short novel “The Nigger of Narcissus.” This foreword serves as a programmatical written basis for the understanding of Conrad’s complete works. A short excerpt from Conrad’s biographical sketch *A Personal Record* is cited as well. Then the book contains an afterword of the Swedish Conrad researcher and biographer Gunnar Fredriksson which is in Swedish in the original copy; in addition there are three reproductions of original photos from the years 1890 and 1921, next the reproduction of a map of Africa with a quote from Conrad next to it, and finally also five of Joseph Conrad’s letters translated into Swedish which in the original copy were in English and French respectively. I am going to ignore the additional paratexts like page numbering, printing information, and the attached short descriptions.

I am showing you this book in this graphic way as an interlinked connection in the bibliographic data model FRBR_{oo} which is part of the formalized and universal ontology of the CIDOC Conceptual Reference Model (CRM). The continued development of FRBR_{er} and FRBR_{oo} is part of the Strategic Plan 2009/2011 of the IFLA Cataloging Section (goal number 2). With the CRM differently structured information from the area of a cultural heritage can be integrated to each other and exchanged with each other; thus from a user’s point of view the attempt will be made to depict the continual flow of reality in its infinite variety within space and time and its reduction by a single person as “his own” perceived, subjective reality with its specificity of relevant, discrete, and persistent objects, persons, and events as a final virtual and user-specific suitable cosmos. As in the case of this Conrad book mostly a further subjective excerpt from the already reduced objective reality—a view—for its own purpose is sufficient. Because generally speaking history appears to consist of and be subjectively observed as recognizable objects, acting persons, and connected events often as a meaningful, organized, and purposeful unfolding process over a time span, so the depiction of events stands as the main element, the middle point, of the CIDOC Conceptual Reference Model and also FRBR_{oo}. They are set up in relation to persons, time spans, places, and physical (that is to say) conceptual objects from which history is formed. Thus the participating persons, places, time periods, and objects or concepts interact with each other and construct in this interaction events which are documentable and through the participation of the named entities often correspond in varied ways. The documentability of events mirrors itself in narratively constructed texts like fictional works, chronicles, letters, documents, government texts, and also photos, maps, and scholarly essays.

In the cataloging guidelines “Resource Description and Access” (RDA) the data elements are stored in data fields (as is stated in the guidelines) which means that physically

distinguishable objects with their determining characteristics are retrievable like index cards where individual characteristics receive information input by linking to standard fields. In an ontology like CRM the depiction in data fields or so-called cites is possible. The cites for the desired excerpt out of this cosmos can be self-defined, can be individually adapted and changed. They are not predetermined. The CRM describes an unlimited network of statements in the form of a semantical network without recognizing a document unity. The CRM neither stipulates development features nor does it represent an average documentation; it offers room for expansions as well as for simplified usage.

It was the former practice in bibliographic indexing to transform more and more to new data, from a book catalog to a card catalog, from AACR2 to RDA with a varied mixture of formal and subject indexing features. This always resulted in a static, in a tabular designation of the holdings by overtaxing the user with multiple hits in the case of broad searches which reflected the publisher, variations of formats, forms, issues, revised, enlarged, and expanded editions, authors and editors, translations, publishing practices. And the bibliographic universe has always been seen from the book. In order to meet this the Functional Requirements for Bibliographic Records has created an initial stage. An ontologically founded depiction of this publishing variety only succeeds – if at first the hidden events are thought of, and secondly, cataloging is not from the book but from the text.

The introduction in the FRBR_{oo} from the idea of “work” up to the single copy “item” is owed to various groups of events. On the one hand it is the direct origin environment of texts, that is to say the materializing or the actual creation from the work to the expressions level, then to the manifestations level:

- F27 Work conception
- F28 Expression creation
- P73 Translation
- F30 Publication event
- F31 Performance
- F29 Recording event

At the item-levels there are again other events which model the object in actual reality which in the library catalog however play a role formerly in the background because they are not important to the library user as long as it is not a question of single pieces, rarities, or similar things. They bundle administrative, technical information. At the same time another change already takes place in part, namely from FRBR_{oo}-inventory to CIDOC CRM entities, from texts to artifacts, from the intellectual to the physical object.

- F32 Carrier Production Event
- F33 Reproduction Event
- E8 Acquisition
- E11 Modification
- E81 Transformation
- E6 Destruction
- E10 Transfer of Custody

E87 Curation Activity

I am coming back to the contents of the book. It contains two fictional works, a scholarly essay, photographs, and historical sources.

The definitive object from a library viewpoint was up to now the single book, therefore the monograph or collected work was simply to be a physically understood product of a publisher. At the same time the independence of an intellectual contribution in publications is completely ignored. I am thinking much more that it should be the single text that formally and subjectively ought to be defined and depicted in the catalog. The boundaries between single essays and various other publication forms have long been blurred; not lastly they are doing it in the virtual space of the World Wide Web and the constantly growing aftereffects from it on physical publication products. It is to depict the single story, the afterword, the letter, something which the author, that is the originator, would consider as buried (it would be then that it became voluntarily or involuntarily left incomplete).

In a perhaps not all too distant future every written and published text in printed form will be available in digital form. The digitalization of the holdings of libraries, museums, and archives is already taking on a not-to-be-overlooked large undertaking. Text recognition (OCR) is always being improved and through cooperative efforts it can be assumed that our printed cultural heritage in digitalized format will lie before us completely within the next few years.

By that I do not only mean texts from printed books, but also written historical sources like correspondence, diaries, newspapers, magazines, manuscripts, and much more.

In the digital age every text will be available digitally as a character string on at least one server in this world . I am convinced of it.

It is therefore time for a new form of the catalog. The card catalog schemata has finally served out its time. Libraries as guardians of the cultural heritage next to museums and archives should see it as their duty to document our intellectual legacy in a defined form among themselves, so that realities, artifacts, and historical events can be brought together.

Conceptionally seen every single completed text and every picture and every map should be seen as an object, often even as several intellectual objects inseparable in a so-called incorporation. In the concrete case named here the book contains two stories by Joseph Conrad, five letters, three photographs, Conrad's foreword to "The Nigger of Narcissus," his poetological bequest, then a reproduction of a map and an afterword by Gunnar Fredriksson. Much has been written and thought about the concept "work" in the FRBR-sense, but that is not intended with this listing. There are text-objects and picture-objects which have been reproduced here, thus manifestations of single texts along with pictures. I am pleading consequently for a new way of thinking about the concept "object" in the bibliographic universe.

It would be the next step that the contents of texts be recognized intellectually, cooperatively, and increasingly automatically. The path has already been taken by one project or another.¹ The intellectual development will not be able to become obsolete then (at least I hope so), but steps as an addition to automatic development.

Thus an automatic text recognition can determine similar texts, agreements in determined percentages, not least of all to be able to point out plagiarisms even much better than title comparisons of cataloged books. Already integrated text and picture objects in new publications and translations can be automatically recognized and classified intellectually or automatically into the ontological network of FRBR_{oo} and thus CIDOC CRM.

I see the task of the academic librarian in the 21st century on the one hand in the intellectual administration of new imprints, thus a cooperative connection achievement together with scholars, publishers, and memory institutions in the form of a fundamental knowledge contribution and on the other hand in the intellectual penetration into one's own subject area in the form of an intellectual administration of the cooperatively lost worded text corpus into a granularized, fragmented text level. At the German Library Congress in Leipzig in March 2010 in one of the lectures² it was correctly warned against placing works of rules too much in the middle point of librarian development work. Because in the area of intellectual subject indexing the content-subject quality of the indexing is by no means unimportant. The management of heterogeneous content indexing data from different sources, like both speakers said, would move into the focus of academic-librarian work. They stress that a shift in emphasis from the title development to work on the normal data be observed. I agree with that and even warn against it. Naturally a standardized and centrally organized data bank of persons, geographical data, and other objective articles beyond one's own national nose is desirable in order to validate the exchange of catalog data internationally, completely in the sense of the Virtual International Authority files. But the content-subject dispute with published texts can by no means be forgotten. Invest in it, do it cooperatively, use existing resources, that means earlier access accomplishments from the card-and-online-catalog times! But assign this task to the publishers and academics. A task assignment in the national libraries should not only be of a physical nature, but should also contain an intellectual new form of "Cataloging in Publication."

Already completed indexing work of around the last hundred years cannot be allowed to be lost. Only one example: a subject card catalog exists in the German National Library in Leipzig in which German belletristic literature has been made accessible to such a measure unknown in Germany.³ The catalog has gained its accessibility on the one hand from the literary understanding of the employees at the time. Such knowledge cannot be lost! In addition, other card catalog schemata, search terms in the online catalog, and the classification arrangement contain the intellectual efforts of many decades and many employees. That should be carried over meaningfully into an appropriate and collaborative subject index. Search terms are found which touch relevantly on the semantic levels of knowledge representation, according to Erwin Panofsky the area of "ofness," but also such levels from the field of "aboutness."⁴ This pair of concepts recognizes still a third meaning level, that of the iconological knowledge level which turns up not in the catalog but only in

academic interpretation for which direction the researcher needs both of the other meaning levels.

I am pleading for an expansion of these semantic levels of knowledge representation to a fourth concept, that of “acrossness.” The Panofsky concept field relates solely to intratextual conditions. Even in doing so real facts are obviously assumed from the objectively learnable world as contents, the concept field always still remains caught up in the fictional hold of the text. What is lacking is an intertextual added value: historical origin information, comparability of fictional as opposed to objectively learnable reality, and above all in the literary sense—understood intertextuality.⁵ The “acrossness” consequently ties questions of space-time placement and concepts of real-fictional text creation in the conception of a work (F27 Work Conception) and the origin of an expression (F28 Expression Creation) into a fragmented text level: when and where has the text (and excerpts from it) been written? Which fictionally reinterpreted times and places bring to mind (obviously) epochs and places of the non-fictional world? What are idea and contemporary moments (subject matters, motives, models) which have stimulated the text genesis? How does it stand with the participation of various persons and groups? Thus the concept of “acrossness” is not only planted in the fictional tangle of the text like the three other Panofsky concepts, but mis-spaces and mis-times it in the perceived reality and at the opposite of other text and picture objects. A library subject index never gives information about it. Giving geographic areas, a person, a concept, a time span, and a genre as the only possibilities of a subject index remains relatively abstract and says nothing about their respective participation, their correlation to each other, and their content-directed contribution for the particular text or their transferability into an objective reality. The context, the actual reality-based historical origin, is largely ignored and not indexed as a subject.

An equivalent intertextual or also hypertextual relevancy of texts to each other is lacking, like I am finding in the representation of texts in a book like *Mörkrets Hjärta*.

- 1) Both stories are the only stories of Conrad whose chief location is in inner Africa
- 2) The five letters belong together because they are the only evidence from the narrow time frame of the Congo trip of Joseph Conrad from May to September 1890 and for that reason refer indirectly to both published stories. That he also wrote a “Congo diary” and “Up-River-Book,” a type of log book for the Congo River, remains disregarded. It would have happened exactly like the German edition of Herz der Finsternis (Heart of Darkness) of 1922.
- 3) The map represents Africa with its white fleck in the middle, the “unexplored region.” With the map, the quote, and the anecdote, a fragment from one of both of his autobiographical works, the prophecy of the ten-year-old Conrad to travel there sometime is connected. Conrad’s foreword from “The Nigger of Narcissus” clarifies his poetological and artistic understanding of what literature from his point of view can and ought to achieve; it is to be seen as a portrayal of reality. Literary and non-literary texts

are personifications of a relationship between the author and the named, known, and captured events.

And predominantly over the captured events, over the external context conditions, but also through the iconological meaning there is an intertextual correlation (“acrossness”) between otherwise materially and motively (“ofness,” “aboutness”) heterogeneous indexed texts, pictures, and maps. In addition there is the already mentioned incorporation, if—overlooked by a subject index—several artistic-intellectual objects are combined inseparably in a text, a picture, or a map.

No library catalog can give me an answer which texts or text fragments—whether fictional or non-fictional, even historical sources—were written at the same place and/or at the same time. The search is also lacking as to which texts or artifacts or historical contingencies served as sources for other texts. Likewise, the publication history and the search for translations of certain texts (and I do not mean books here) and their time and space spread are not possible. I can also not examine and discover which texts are cited in others directly or indirectly as allusion. A literary and historical need and the contents of Critical Editions are always able to determine what inspired the origin of a text, how literary influences and motivations (P15 was influenced by/influenced, P17 was motivated by/motivated), how places, persons, and events gain influence over the text creation. The search for thematically similar texts proves likewise to be just as difficult, those that look after the various times of the material or motif which would allow the analysis of a diachronic development of the literary finish of a definite theme.

Too frequently the catalog user is overwhelmed with the convenience of the online catalog searches which are strongly oriented to internal library processes rather than being open to outside cooperative control. It would be a goal of an ontologically modeled catalog data bank to show how a researcher arrived at his results but also to be able to pose questions on further links and in the process it cannot be forgotten that museums and archives also collect texts and other products of intellectual and artistic processes. There, indexed in similar form, by an exchange format like CIDOC CRM or FRBR_{oo} backgrounds going beyond institutions and more global connections and origin histories can be researched, completed, and criticized. Former bibliographic extensive research would gain a new quality and the time-saving would be enormous.

I am pleading for a totally new cataloging, certainly owed to the scope of the publications, for a cataloging to be achieved cooperatively using existing data standards. Every national library should achieve and introduce in coordination with each other a collaborative and complete index on the four FRBR or RDA levels. Cataloging of intellectually or artistically created products should consider the various circumstances of the origin itself and the iconographical and iconological interpretation which can be achieved on these texts best by researching scholars, and truly not according to a prescribed scheme, but in relation to the research question per se.

The publishers should also be taken to task to prepare digitalized versions of all their publications with the agreement of authors and copyright holders and they should be required to complete standard data and formal information where this is asked for and helpful.

All of this should not obscure the fact however that in the case of an intellectual index it is a question of either a prescribed and above all narrowed classification or component word scheme or a reduced portrayal to a scholarly query about the intellectual contents of a written source which in no way, I stress, no way can substitute for the continuing intellectual work of a scholar, but serves solely as a lead to further sources and to their plausible connection. The links within the networks created here are only as good as they are programmed by the inputters and filled in by the subject catalogers. On the other hand it is true that the availability within a research experience of indexed knowledge realms can never create such a model. The network presented in the catalog offers access and a more or less base framework as an illustration of events and conditions, also partly as “surprisingly” appearing links in the academic sense which a rationally thinking researcher’s brain would never notice, but what can also lead to overlooked relations between very relevant sources.⁶ The requirements of academic knowledge are born in the intensive exchange among researchers, students, and other academics. The cost in the mind of an academic, above all not measurable in intellectual research, goes beyond objective value and is not to be misplaced into an arranged hierarchy according to space, time, and subject entries.

The relation of academic queries to data banks whose formal ontologies are based on CIDOC CRM or FRBR_{oo} is still not sufficiently theoretically cultivated. The role of research data banks can only be much more exactly specified, unsuitable structures better identified, and access methods optimized if a series of research questions have been solved with the aid of formal ontologies. A data bank modeled with CIDOC CRM, in my opinion, can make academic evidence comprehensible and stimulate criticism and further work.⁷

What remains to be done, what has made possible the grasp of certain connections, is namely a grouping of events with the help of retrieval of relations among places (E53 Place), times (E52 Time span), persons (E39 Actor), conceptual objects (E28 Conceptual Object), and things (E70 Thing)—at the same time always inclusive of nearby searches which are to be understood spatially, timely, and conceptually. Historians and literature scholars must break through from the words of the text to the worlds of the text, from formulations to reconstructions of the actual meaning. The assumed, iconological construction of a text in this way as expression of the intellectual and material moments in the text genesis and also to form their causal, conditional, or final connections to other texts and real events through search queries in data banks is not achievable now, but is to some extent possible with an event-centered data bank model like CIDOC CRM. In doing this a way should be found for the effective rendering of relations recognized by scholars to lead to new subject connections.

I have devoted part of my life to research on the literary reception of Joseph Conrad in the German-speaking world. Besides the publication history this consisted of pulling together articles and discussions of his life and work from German newspapers and magazines, to index them intellectually and to make them available in a user-oriented data bank. In light of the possibilities which FRBR_{oo} and the CRM offer and which I have described here, I see myself faced with entirely new challenges: not only will the search time for newspaper and magazine articles be shortened as well as texts being available at

the press of a button and many trips to various libraries and archives made superfluous, there will be more time for actual research questions and it will be an enormous advantage for academic activity. And the scholar will gladly increase the indexing of the results of his research queries if he receives a bibliographically comprehensive modeled diversity.

Critical editions register the naming of event, places, persons, etc. which stand in relation to certain texts (thus works). Such a library catalog misses this. An ontological-structured bibliographic data bank should be able to achieve this.

For example, the conception of a work, thus the text genesis—a genuine need in every literary research—would be more achievable. Thus the two letters from July 22, 1896, to his publisher, T. Fisher Unwin, and his editor, Edward Garnett are of significance for the text cosmos of the Conrad book in which Conrad depicts the developmental history of *An Outpost of Progress* which he was just writing. In the same manner there are relevant letters from the developmental phase of *Heart of Darkness* from December 1898 until February 1899, among others to the publisher, William Blackwood, or the author and friend, Ford Madox Ford. They are all lacking in this collection, just as the aforementioned *Kongo Diary* and the *Up-River-Book*.

The tasks of a user, which an RDA-based catalog ought to fulfill, are defined as: find, identify, clarify, and understand. Over the years of bibliographic indexing at least the first two tasks have remained essential and irreplaceable. The other two tasks are in my opinion superficial and false; they arise from the attempt to want to bring into harmony the automatic calculation of associations, which a computer fulfills, with human thinking. That is completely unnecessary however. In the personal compilation and the reading of fictional works, of historical documents, and of scholarly essays, the user creates in his head his own cosmos with its own connections and associations, its own backgrounds and thought stimuli which a search query only casts if the data bank is correspondingly fed what a meta-data bank cooperatively with museums and archives can be in a position to create on the basis of CIDOC CRM. The clarification and understanding have a catalog molded on the side of the user. The former possible search queries do not arise unconditionally from a need for knowledge acquisition but generate purely coincidental, empirical, partly surprising statements which seldomly meet the needs of an academic search query, which is initially asked of the system but whose value and usefulness is only surrendered retrospectively and mostly after a costly, personal examination for the user and the library of (from a standard scheme) rejected documents.

But every user is different, please make that clear! Individuality should also not stop in front of the library catalog. What can the users find? What do they want to find in the catalog? What should they find? The formal description covers details like book size and weight. Subject indexing covers contents. But no one needs the entire scope of these details all the time. Generate user-specific sites for the cataloged bibliographic details and documented contents!

In conclusion I would like to show you still a catalog environment based on CIDOC CRM, which as a user would not overwhelm me, but where I would feel well taken care of and sufficiently informed about my

search queries in my area of research and one could better comprehend my academic research work.

The offer that you will make with RDA gives access in its beginnings to the daily user in order to index a text cosmos that he wants to “choose” because it satisfies his needs, a user who wants to read works with certain plots or by certain authors or a certain epoch. But what do you do with scholarly users? Current online library catalogs are not an instrument out of which scientific knowledge acquisition can be generated; it cannot be denied, the constant learning of the world, that is to say more scholars and more research, cannot be halted.

Therefore my request to you: don’t catalog the books! List the texts, catalog the contents! The intellectually indexing academic librarian should not be superfluous. Dedicate your time as a subject specialist to in depth contents indexing of works, but in conjunction with researchers in their fields. Work on a universal view of the contents of the bibliographic world. Join together with museums and archives and also with business concerns. Work cooperatively together! Do not be bothered by the activities of Google, cooperate with them, make offers, use them, yes, take advantage of them!

And above all: page through the books, ask the impossible, and think of the unthinkable! It’s worth it.

Kiel, April/May 2010

¹ For example in the Projekt Wiss-KI, see URL: [http://wiss-ki.eu/\[4/8/2010\]](http://wiss-ki.eu/[4/8/2010])

² Braune-Egloff, Dörte; Q, Esther: Die RSWK-Revision vor dem Hintergrund Aktueller Entwicklungen. Lecture on 3/17/2010 at the 4th Leipziger Kongress für Information und Bibliothek. [URL: http://www.opus-bayern.de/bib-info/volltexte/2010/874/pdf/RSWK-Revision-final_korr.pdf]

³ Compare—Förster, Frank: Die Erschließung belletristischer Literatur in Sachkatalogen wissenschaftlicher Bibliotheken im deutschsprachigem Raum [engl. Indexing of fictional literature in subject catalogues of academic libraries in German-speaking countries]. Cologne University of Applied Sciences, Master’s thesis 2008, revised ed. 2009.

⁴ According to: Panofsky, Erwin: Introductory. In: Studies in Iconology: Humanistic Themes in the Art of the Renaissance. New York, 1939 [Ger. Studien zur Ikonologie, Köln, 1980].

⁵ Intertextuality is differentiated here in the sense of “Transtextuality” of Gérard Genette (compare: Genette, Gérard: Palimpsestes. Paris, 1982).

BIOGRAPHICAL INFORMATION

Frank Förster, M.A., M.L.I.S.; studied German language and literature, English language and literature, journalism, and library and information science in Leipzig and Cologne (Germany); works as Coordinator of the Humanities Data Bank at the graduate school “Development of Human Societies in Rural Areas” at the Christian-Alberts-University at Kiel (Germany); was an intern at the German National Library in Leipzig and on the academic staff at the University of Leipzig and at the Technical College of Cologne; latest publication: “Vernetztes Wissen—Ereignisse in der bibliographischen Dokumentation,” Zeitschrift für Bibliothekswesen und Bibliographie 57 (2010) 1, 15-25 (mit Hans-Georg Becker, Dortmund)

Frank Förster, MA, MLIS
(Christian-Alberts-University Kiel, Germany)

Bibliographic and universal access:

Written historical sources and works of fictional literature in FRBR_{oo}

