



World Library and Information Congress: 69th IFLA General Conference and Council

1-9 August 2003, Berlin

Code Number: 085-E
Meeting: 151. Social Science Libraries
Simultaneous Interpretation: Yes

Standardization, heterogeneity and the quality of content analysis: a key conflict of digital libraries and its solution

Jürgen Krause

University of Koblenz-Landau and Social Science Information Center (IZ Bonn),
Lennéstr. 30, D-53113 Bonn, Germany

Krause@bonn.iz-soz.de; <http://www.uni-koblenz.de/~krause/>

Abstract

Solutions for the problems of building up information services go beyond the hitherto conventional way of thinking about information centers and libraries. The disputed guidelines of “Standardization from the view of the remaining heterogeneity” and the paradigm of “Net publishing” clearly characterize the change. It isn’t just technological but content-conceptualized. Both guidelines taken together are an answer to the new technological and user demands in the changing setting of the Web. And a clear sign that the traditional methods of standardization are no longer sufficient to keep the desired interoperability and data consistency that should be maintained.

Keywords: digital library, content analysis, metadata, heterogeneity, Web publication; text-fact integration, cross-concordance

1 Introduction

“Access Point Library”, what should this motto mean to the libraries today? What does it really mean, especially with its historical background of the past 30 years. Will the library be reduced to an access point contents others created or should it keep the old central role as the information provider for science and the supply from publishers, scientists on the Web and technical information centers become secondary to the library.

Presently in Germany, the interplay can be well studied among technological development, new user needs, shifts in the organization and service structure, and the resulting changes in the modeling of information process. In the scope of social science, this is a part of informational society, as described by Castell (2001) as a “network society” broadly taken from the perspective of economics and sociology.

2 Poly-central information provision

In looking back today, an important change took place for libraries about 30 years ago. They released an important part of the content development of scientific information to information service providers and technical information centers, that store and record literature databases in specific technical areas, along with independent literature from journal articles. In general, the actual journal stays, along with the borrowing function, at the library. Today, one can ask oneself, why did this function distribution come about? Why should an institute have primarily journal articles and another only books? Why didn't libraries make the move, then, towards information technology, didn't get involved in the development of literature database, but gave it to a new found institution; and what does that say about the situation today? Will libraries do justice to the division started by the information technology development? Will they continue to be important players in the scientific information provision or will they give up more functions, until they mutate to collections of physical documents which other information organizations (as long as there is a need) provide in electronic form?

But first: How does the libraries and more generally, the literature provision in the scientific sector look like and what is their main problem?

The world of information providers is (as opposed to the political) no longer centralized or bipolar, but poly-central. Technologically speaking, access is available worldwide from different information sources, relatively easy and readily available (at any time of day and from any spatial distance). This multiplies (in contrast to conventional media) the amount of active content distribution. In parallel to other areas of e-commerce it 'lowers the barriers of market entry' and works against existing monopolies (Cigan 2002;15). Information providers can directly reach their goal audience world-wide. At the same time, 'the internet shifts the market power from producer to consumer' (Cigan, 2002;19).

The scientific information provision shows the technological changes as door openers for a growing decentralization. The alternative by Cigan (2002;15;19) of knowledge markets as thinkable developments to a new monopoly formation does not take place here. Libraries with their OPACs and the I&D database are only now part of a versatile heterogeneous service. Currently, user information services are facing a high-grade decentralization and heterogeneous documentation space.

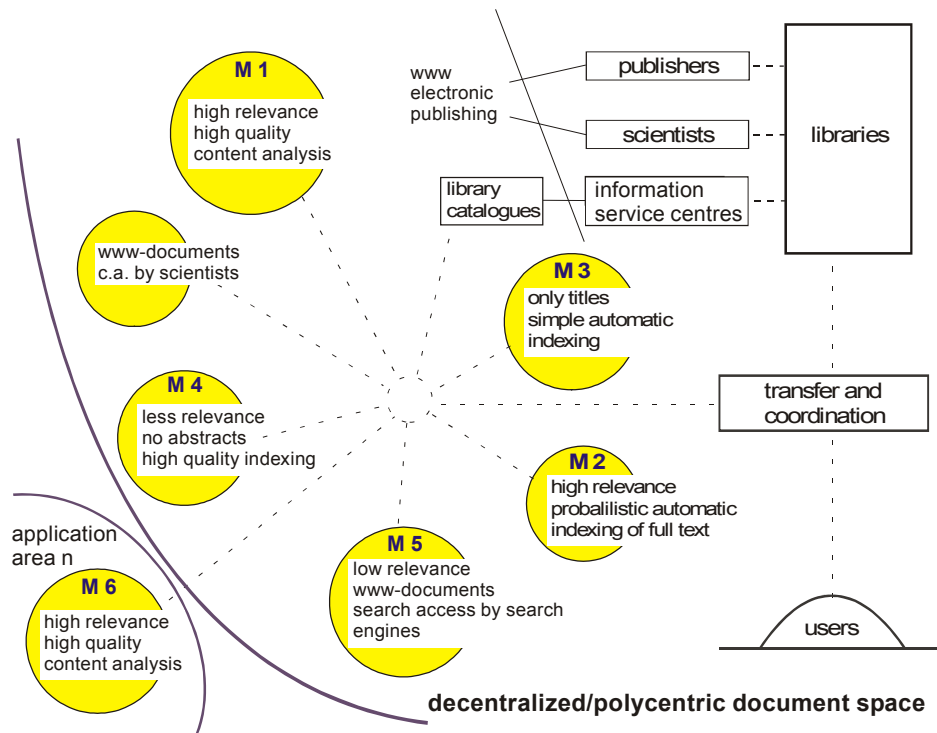


Figure 1: Decentralized/polycentric document space

Besides the traditional information providers (the publishers with their printed media; the libraries, which record their books according to intelligently assigned classifications; and the technical information centers that provide their information through hosts) are strengthened by the scientists themselves. They, who develop in all these areas, over the WWW, independent services which supply differently, have different points of relevance and presentation processing. Generally, those who collect information in special areas, could crop up anywhere in the world. A result of this is the different lack in consistency:

- Relevant, quality controlled data is among irrelevant and possibly data that can be proven false. No editorial system ensures a division between rubbish and potentially desired information. Whoever, e.g., as a social scientist has the research area of couples and sexual behavior, knows, what that entails when searching the Web?
- A descriptor X can take on the most diverse meaning in such a system. Even in limited technical information areas can a term X, which is ascertained as a highly relevant, with much intellectual expense and in high quality, not be matched with the term X being delivered by an automatic indexing from a peripheral field.

The user, despite such problems, will want to access the different data collections. No matter which process he chooses or in which system they are provided. He also has in the world of decentralized, inhomogeneous data, the justified demand on the information science to ensure that he receives when possible only and all relevant documents that correspond to his information need.

How is management, and which changes in the library's and I&D organization's traditional, well loved procedures and way of thinking, are attracted by the new circumstances?

2.1 Digital libraries as hybrid libraries

For 30 years, determined by the technological development libraries and technical information centers were forced to organize centrally and in so doing, align conceptually the

content indexing in a centralized approach. A mainframe computer was set up to run the data. The clientele are served by terminals or offline by inquiry at a reference desk.

This corresponds to the theoretical basis of the context indexing. By a normed intellectually controlled procedure, that was developed and carried out by the reference office, a uniformed indexing of the documents. In this way of thinking, the data consistency received the highest priority. It, and the uninterrupted indexing of relevant documents without wasting time with accurate delivery, becomes more time consuming and difficult in today's environment.

Centralization attempts in terms of complete data collection into a database by an organization are barely evident now. Even in the library environment this concept has been replaced with the speculation of networks. This model presentation best explains the concept of digital libraries. Digital libraries should make it possible for scientists to have an optimal access from their computer to the worldwide available electronic and multimedia full-text, literature references, facts databases, and WWW information. Which enables availability to teaching material, special listing for experts, for example. Digital libraries are in a manner of speaking, hybrid libraries with a mixed collections of electronic and printed data (and if applicable other types). The latter is available through electronic document ordering and delivery services. On the technical side this requires among other things access to distributed databases on the Net; on the conceptual side, the integration of different information contents and structures. Examples for the obtained technical integration of heterogenetic data collections are the library union KOBV (<http://www.kobv.de/se/cont.html>) or the digital library union NRW.

What until now the library unions and similar projects do not usually do is take adequately into account the different content indexing processing of partial collections.¹ In fact, this yields, in comparison to the weaker search machines of the general WWW (see Krause 2003: section 1) an improvement in choice of relevant data material, the conceptual difference here between different content indexing processes become, however, also unbalanced.

Traditionally, in context of the digital libraries an attempt is made to secure conceptual integration through standardization². Scientists, librarians, publishers and the providers of technical databases have to agree for example, to the Dublin Core-Metadata and a uniformed classification such as the DDC. In this manner, a homogeneous data space is created, that allows for consistence high-quality data-recall. Unfortunately there are clear signs that the traditional processings of standardization have reached their limits. Already in traditional library areas there was often more claims than reality. On the one hand, standardization appears to be indispensable and has, in some sectors, clearly improved the quality of information search. On the other hand, it is only partially applicable in the framework of the global provider structures of information, with rising costs. Especially in some sectors of the content indexing does it become clear that for digital libraries (despite much needed efforts) it cannot go by the applicability uniformed standards of content descriptions. Therefore, a different view has to be found for the unending demands for consistency and interoperability. This will occupy us in the third section under the motto of "Standardization from the view of the remaining heterogeneity".

¹ An exception to this rule is the ViBSoz, the digital library for the social sciences, that first employed heterogeneous components as discussed in part 3, see http://vibsoz.bonn.iz-soz.de/ViBSoz_Start.html. An overview of the relevant research program from BMBF and the DFG including the hitherto attained level of service performance is given by Schöning-Walter 2003

² For an overview see Krause/Niggemann/Schwänzl 2003

2.2 Publishing on the Net

The Web goes beyond the consideration of modeling a clear decentralized information space of library section archives i.e., beyond a Z39.50 interface. System development and the data format from information collections refer to the paradigm of “publishing on the Net”, which gives the clearest expression of the semantic web-approach along with initiatives such as DDI or OAI (open archive initiative)³.

The vision behind these efforts is clearly seen in, e.g., the project NESSTAR and Faster⁴ from the area of social scientific data archives, which goal focus is presented by the following illustration. It also contains the connects of textual elements (e.g. publications) with factual data⁵.

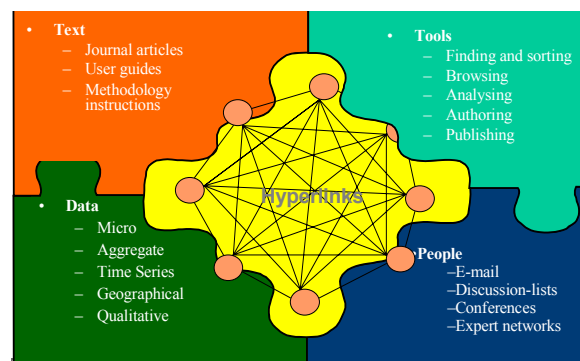


Figure 2: NESSTAR by Ryssevik 2002:14

The paradigm “publishing on the Web” makes one thing clear: It was never so difficult as today to model new information systems and put into practice, which is the foundation of every Web-activity, that comes about from this premise. Every new offer “*. is designed to fit into a wider data input and output environment*” (Musgrave 2003;05). Earlier system developments “only” need to worry that their system, within itself, accepting efficient and fast inquiry and acting upon the user’s needs. Today this is not enough. No one works isolated for himself and his user group anymore. Everyone is part of a global demand and fulfills, in this technical scientific information context a small, unique task. This goes for libraries, as well. The user of a special database would not limit himself to this one offer, but would want to, in an integrated way, access many similar collections. Some of these clusters are already known at the start of development of a new provision. More importantly is, however, that the security in the upcoming years after completion of one’s offer new information collections are added to the Web, where the user would like to have integrated access.

And since one knows this, the difficulty lies not in the concrete system programming but in the modeling of the system as a fitting sub-unit. Ideally the Web-community sees itself as a

³ For an introduction to semantic web see Matthews 2002; Data Documentation Initiative (DDI) (<http://www.icpsr.umich.edu/DDI/>); Open Archives Initiative (OAI) (<http://www.openarchives.org/>)

⁴ Networked Social Science Tools and Resources (NESSTAR) (<http://www.nesstar.com/>); Flexible Access to Statistics, Tables and Electronic Resources (faster) (<http://www.faster-data.org/>)

⁵ In Germany this approach is equivalent to the planning of the KVI-Project (see Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik 2001). The point is the combination of fact data from the censor bureau, which are above all process generators, and the produced data from science (e.g., survey such as ALLBUS). The existing database should be enriched by the missing metadata, to a great extent standardized and as much as possible, be provided in an integrated manner. New metadata systems should be tailored for microdata at different locations by social scientists for science and statisticians for censor bureaus.

community of system providers, whose contributions are adjusted such that each sub-element fits with the other without any prior agreement between the participants. Regardless of whether or not it has been modeled and programmed correctly in the sense of the Web-paradigm. Each system provider should be able to read and process any data collection without a problem. Each provider of system services should ideally be able to integrate and further develop any system module without having to redo the developmental work done by others because some existing module does not fit.

The protocol level today hardly causes any problem under this paradigm (e.g., http, JDBC) furthermore, neither does the syntax level (HTML and at least in the modulation almost established XML). Today's professional development systems work on this standardization and "fitting" bases. Only then can a search engine be constructed, which can search in any server and index their data without prior agreement. That this standardization falls short and none can achieve enough result quality for the technical scientific aims, it is seen today as certain. Further standardizations in the structuring and the contents are necessary. Parallels to the consideration of content indexing and remaining heterogeneity, in the paradigm of digital library were discovered by Musgrave for the example of social scientific data archives:

"On top of the syntax provided by XML and the structure provided by the DDI there is a need to develop more standard semantics via thesauri and controlled vocabularies in order to make for better interoperability of data." (Musgrave 2003;1).

The structuring is the international cooperation of data archive thank DDI already very widely expanded (see Ryssevik 2002 as introduction and the DDI homepage: <http://www.icpsr.umich.edu/DDI/ORG/index.html>)⁶. Controlled vocabularies and thesauri in many subsectors cannot be summarized to a so-called metathesauri. This is not at all necessary, since the heterogeneity components retrieves well (see section 3).

The limits of today's development are in the exchange and "fittingness" of the functionality. Pursuits such as the agent system or the semantic web-initiative show the way as a rough outline for future systems (see Matthews 2002).

Conclusion: The discussion of guidelines of "publishing on the net" goes beyond the decentralization discussion of digital libraries. The IT-change of the past decade is most clearly characterized by the expansion of the WWW. All libraries are subjugated by it. It is not only technologically but content conceptualized. It allows cooperation only in combination with all of whom participated in the information service so far, who bring with them their technical know-how and open new solution possibilities. The times are over, where simple technical oriented only solutions are for every access point type, for and in the libraries; as well as, the hope that information technological know-how for technical science servicing can be reduced to only a subsequently acquired program knowledge by technical scientists.

2.3 Standardization with decentralized organization

Also in the paradigm of "publishing in the web" are efforts to bring back homogeneity and consistency in today's decentralized information world; when creating suitable information systems, that can deal efficiently with divided data collections and the keeping of standards.

- First step is for the technique-oriented viewpoint of a problem solution. One ensures that the different document spaces can be physically retrieved simultaneously and that it happens efficiently. The majority of publications to a literature search about "divided

⁶ An overview about the library landscape, the technical scientific services of the information centers and the digital libraries is given Krause/Niggemann/Schwänzl 2003.

databases, integrated information systems” are meant here. These technique-oriented solutions of the decentralized document spaces problem are an indispensable prerequisite. It still does not solve the main problem of content and conceptual differences between the individual documentation collections.

- The approach of implementing metadata goes a step further. Metadata are agreed specific characteristics of a document collection in an arranged form applied to one’s own data, no matter how different they are from other characteristics. An example for this is the Dublin Core (<http://dublincore.org/>) which plays an important role in the scope of global-info. Metadata supports at least a minimum of technical and conceptual exchanges. E.g., see Jeffery 1998 in the acceptance of metadata in important solution strategies for the European project searchable databases union.

Efforts to standardize and initiatives for acceptance and expansion of metadata are unquestionably important and a prerequisite for a broadening search process in a daily decentralizing and poly-centralizing information world. In principle they try (at a low level) to do the same as the centralized approach of the 70’s, which tried to leave out the cooperation agreement without having the hierarchical authority. Especially in the area of content indexing, they try to restore the data homogeneity and consistency through voluntary agreement of all of those involved in the information processing. The above considerations on dealing with heterogeneity using the classical demand of massive standardization efforts, is closer and not wrong per se: When everyone uses the same thesaurus or the same classifications, we won’t need heterogeneous components. If the individual provider deviates from the basic premise of any standardization procedure, it must be “somehow” possible to make (force) him play by the classical rules.

As long as one understands that this can only be partially achieved, everything speaks in favor of this kind of initiative. No matter how successful the implementation of metadata can be in a field, the remaining heterogeneity, e.g., in terms of different types of content indexing (automatic; varying thesauri; different classifications; differences in coverage of the categories) will become too large to neglect. The wider access to the network would affect the centralizing doctrine of information indexing (which is also clearly felt in the metadata activation) if not be lost. Here, one can rank positively the efforts made by the German library scene, to overcome the RAK-WB and the MAB-formats and to establish the mandatory transformation to AACR2 and MARC21 for all German libraries⁷.

All of the world different groups can crop up, which gather information for specialized areas. The user will want to have access to them, independent of which approach they use or in which system they provide. The above mentioned cooperation model would demand, that the information agent responsible should get in contact with this provider and try to convince him to keep certain document norms and content indexing (e.g. the Dublin Core). That may work in individual cases, but never as a general strategy. There will always be an abundance of providers who will not subdue to the stipulated guidelines. Previously, central information service centers would not accept a document which did not keep certain rules of indexing. In this way, the user (ideally) always confronted a homogeneous data collection. On this, the whole I&D-methodology, including the administrative structure of the library and technical information center was arranged. Whether it was right or wrong, this initial situation no longer exists in a worldwide connection system nor in the weaker form of metadata

⁷ The first resolution in this direction published at the end of 2001 from the “standardization committee”, who is the body responsible for German libraries. There are many questions still open, which should be answered by one of the German library coordinated projects by the end of 2003.

consensus. For this reason, the data consistency postulate as an important cornerstone of today's I&D-behavior has been proved an illusion.

Today's I&D-landscape has to react to this change. Thus the question becomes, which model conception can be developed for the remaining heterogeneity on different levels.

3 Remaining heterogeneity in the area of content indexing

If one wants to find literature information (and later, fact information and multimedia data) from distributed and content differently indexed data collections, which are found together, not having anything to do with each other, heterogeneous organized structures and accessible contexts (institute libraries, special collections from university libraries, scientific specialized libraries, reference databases, digital full-text) with an inquiry for integrated searches, the problem of content retrieval from divided document collections must be solved. A keyword X chosen by a user can take on a very different meaning in different documentation collections. Even in limited technical information areas can a keyword X, which has been ascertained from as a highly relevant, with much expense and in high quality document collection, will not be matched with the term X being delivered by an automatic indexing from a peripheral field. For this reason a pure technological linking of different documentation collections and the formal integration at a user interface is not enough. It leads to falsely presenting documents as relevant and to an abundance of irrelevant hits.

In the context of expert scientific information is the problematic of the heterogeneity and multiple content indexing generally very critical. Because the heterogeneity of the data type is especially high, e.g., fact data, literature and research projects data should be accessed simultaneously and besides the traditional technical information user groups appear target groups send request in colloquial language instead of scientific language. In spite of these heterogeneous starting point, the user should e.g. not be forced, to acquaint oneself with the indexing system of the library in order to learn when extending their search intention on gray literature as secondary system of content indexing and to convert in the right search strategy, and again another, when he wants to search further scientific databases.

For this reason the existing heterogeneity of different content indexing systems has to be related to one another through suitable measures. The first step is the integration of scientific databases and library collections. It has to be supplemented by internet resources and facts data (e.g. time series from surveys such as in NESSTAR) generally by all data types that we can find today at digital libraries and the different technical portals and at electronic market places.

3.1 Bilateral transfer module

In the next short model sketch presents a general frame in which certain classes of documents with different content indexing can be analyzed and are algorithmically related. Central are intelligent transfer components between the different forms of content indexing, which carry semantic-pragmatic differential computation and which allow itself to be modeled as independent agents. They interpret, in addition, conceptually the technical integration between the individual data collection with different content indexing systems. The terminology of the field specific and general thesauri, classifications, eventually also the thematic terminology and inquiry structures of concept data systems etc. are related to each other. The system must e.g. know, what it means, when term X from a field specific classification comes, or is used in a thesaurus for intellectual indexing from a journal article, which the WWW-source only automatically indexes. Term X should only be found by chance

in the terms of the running text and only then when a conceptual relationship between the two are analyzed.

For this reason, transfer modules should be developed between two data collections of different types, that the transference form is not only technical but also conceptual (more details on bilateral transfer are given in Krause 2003).

Generally there are three approaches that were tested and implemented, for their effectiveness in individual cases. None of the approaches were solely responsible for the transfer burden. They were restricted by one another and work together.

- Cross-concordance in Classification and Thesauri

The different concept systems are analyzed in a user context and an attempt made to relate intellectually their conceptualization. This idea should not be confused with the metathesauri. There is no attempt made to standardize existing concept worlds. Cross-concordance encompasses only partial union of existing terminological systems, of which the preparatory work is used. They cover with it the static remaining part of the transfer problematic. Such directories offer possibilities by data-recall of terms of a concept system, which can be enhanced by another, in individual cases, in the sense of a synonym and similarity relation but also as a deductive rule relation.

- Quantitative-statistical Approaches

The transfer problem can be generally modeled as a fuzzy problem between two content description languages. For the vagueness addressed in information retrieval between terms within the user inquiry and the data collections, different operations have been suggested (probability procedures, fuzzy approaches and neuron networks (Mandl, 2001), that can be used on the transfer problematic. Procedures of these types need training data. The individual document can be indexed into individual documents in two concept schemata or whereby two different and differently indexed documents can be put in some relation to each other. For the multilingual IR the same text, e.g., can be in two languages.

The initial situation for such procedures is by the fusion of library collections with specialized databases of information centers, very favorable. This is due to the fact that the information centers in addition to articles almost always take over the independent literature so double keywords are furnished. In the digital library of the social sciences, all documents are e.g. of the specialized fields collection from the University of Cologne are also indexed in SOLIS.

In ELVIRA8, an IZ-project, which deals with the fusion of text and facts (time-series) for a united information system a co-occurrence analysis was made between time-series and automatic indexed text. Until now, all models that have been tested were based on one approach, that has been suggested by Sheridan/Ballerini 1996 for a multilingual retrieval. In contrast to the cross-concordance, the transformation is based not on general intellectually determined semantic relationships but the words are transformed in a weighted term vector that mirrors the term use in data collection.

Which approach proves to be the most promising, can only be rendered from the bases of real data.

- Quality-deduction Procedures

⁸ see <http://www.gesis.org/Forschung/Informationstechnologie/ELVIRA.htm>

Deductive components are found in intelligent information retrieval (Belkin 1996, Ingwersen 1996). Also in intelligent data-recall systems such as OSIRIS (Zillmann 1997, Ronthaler/Zillmann 1998) and in the area of expert systems.

What is important, is that the postulated transfer modules bilaterally work on the level of the database. They combine terms from different content description. This is conceptual, and the practical results are somewhat different from the fuzzy problematic handling between user query and the document collection of the database, which is integrated into the search algorithm. In this manner, the transfer modules can be e.g., between a document collection which is indexed using a general keyword list such as SWD⁹ and a second of these indexes which is based on a special field-specific thesaurus. Through qualitative procedures such as the cross-concordance and deduction rules, cross-referencing and creating search algorithms to connect the user terminology to a probabilistic procedure. The possibility that the transfer fits in each case, depends on the connection conditions from different data types and the problem of being able to encounter different concept systems, not only undifferentiated data-recall algorithms, is an important difference from the IR solutions so far.

3.2 Standardization from the view of the remaining heterogeneity

Heterogeneity components open new viewpoints on the existing remaining demands for keeping consistency and interoperability. It can be restated with the following premise: Standardization should be viewed from the standpoint of the remaining heterogeneity. Since technical provisions arise today from different contexts with different content indexing traditions (libraries, specialized information centers, Web communities) their rule and standards meet, which are valid in their respective worlds.

The quintessence “standardization from the view of the remaining heterogeneity” is further clarified in Krause/Niggemann/Schwänzl 2003 with today’s achieved standing of handling heterogeneity connected to the specialized information provisions. It refers, above all, to the clear deviation of the user from low quality standards of conventional general search engines on the Web and by doing so the type of content indexing, its intrinsic existence and unchangeable partial discrepancies between the different existing data collections and the solution path that should be taken:

“Despite voluntary agreement of everyone participating in information processing, is, nevertheless, a thorough homogeneity of data impossible to create. The remaining and unavoidable heterogeneity must be met, for this reason, with different strategies, new problem solutions and further development are necessary in both areas:

- Metadata
- The methods of handling the remaining heterogeneity”
(Krause/Niggemann/Schwänzl 2003;27)

Krause/Niggemann/Schwänzl (2003) have found that both demands are closely connected. Through further development in the area of metadata should, on the one hand, lost consistency be partially reproduced; on the other hand, procedures to deal with heterogeneous documents can be cross-referenced with different levels of data relevance and content indexing.

⁹ SWD is from the German scientific university libraries cooperatively constructed keyword nor data on the basis of the rule set RSWK „Rules for the word key catalogue“, see Krause/Niggemann/Schwänzl 2003.

4 Summary and Perspectives

The problem in constructing a technical information provision (whether it is an access point for libraries or as a “marketplace” or scientific portal for other information providers) goes beyond the current common thinking of information levels and libraries. The disputed guidelines “Standardization from the view of the remaining heterogeneity” and the paradigm “Publishing on the Net” characterize best the change. It is not only technologically but content conceptual. It can be surmounted only with cooperative, in a joint effort of all who have participated until now on the information provision, who each bring their specialized expertise and open new solution procedures. The times are over of simple only technically oriented solutions. Also the hope is, likewise, of reducing information technological know-how for technical provisions to just programming knowledge.

Recent user surveys clearly show that clients of information services have the following aims for technical information¹⁰:

- The primary entry point should be by a technical portal
- Neighboring areas with cross-over areas such as mathematics-physics and social sciences-education-psychology-business should have a build in integration cluster for the query.
- The quality of content indexing must clearly be higher than the present general search engines (no “trash”)
- Not only metadata and abstracts are wanted from the library but also the direct retrieval of full-text.
- Not only library-OPAC’S and literature databases should be integrated into a technical portal but also research project data, institutional directories, WWW-sources, and fact databases.
- All sub-components can be offered in a highly integrated manner. The user doesn’t want, as at the human help desk, have to differentiate between different data types and have to restate repeatedly his question in different ways, but only give once and directly his request: “ I would like information on Term X”.

The fulfillment of these types of wishes also means under the paradigm “Standardization from the view of the remaining heterogeneity” and by the acceptance of the guideline of “Publishing on the Net” that many other questions are left open. For example, the problem of interplay of universal library provisions and that of the field-specific preparation of literature archives from technical information centers need to be clarified, when one wants to create an overlapping knowledge portal like VASCODA in Germany (<http://www.vascoda.de/>). Both guidelines produce an acceptable starting point. The consequences of the changes that are mirrored in the above user demands, are highly complex structures that also lead in detail to questions because there are no complete solution models anymore, that the librarians and the information center “power” could fall back on, like before.

Literature

Belkin, Nicolas J. (1996): Intelligent Information Retrieval: Whose Intelligence? In: Krause, Jürgen; Herfurth, Matthias; Marx, Jutta (Hrsg.): Herausforderungen an die Informationsgesellschaft. Konstanz 1996, S. 25 - 31.

¹⁰ For the following results of the study from the library committee DFG see the report DFG in ZfBB h.1,2003, p.41-44; for the BMBF-studies see Schoening-Walter 2003; to GESIS-Studies :Binder et al. 2001; For the Infoconnex-studies see IMAC 2002.

- Binder, Gisbert; Klein, Markus; Porst, Rolf; Stahl, Matthias (2001): GESIS-Potentialanalyse: IZ, ZA, ZUMA im Urteil von Soziologieprofessorinnen und -professoren, GESIS-Arbeitsbericht, Nr. 2. Bonn, Köln, Mannheim.
- Castells, Manuel (2001): Der Aufstieg der Netzwerkgesellschaft. Teil 1. Opladen, S. 31-82.
- Cigan, Heidi: Der Beitrag des Internet für den Fortschritt und das Wachstum in Deutschland. Hamburg: Hamburg Institute of International Economics, 2002. (HWWA-Report 217)
- IMAC (2002): Projekt Volltextdienst. Zur Entwicklung eines Marketingkonzepts für den Aufbau eines Volltextdienstes im IV-BSP. IMAC Information & Management Consulting. Konstanz. September 2002.
- Ingwersen, Peter (1996): The Cognitive Framework for Information Retrieval: A Paradigmatic Perspective. In: Krause, Jürgen; Herfurth, Matthias; Marx, Jutta (Hrsg.): Herausforderungen an die Informationswirtschaft. Konstanz 1996, S. 25 - 31.
- Jeffery, Keith G. (1998): ERGO: European Research Gateways Online and CERIF: Computerised Exchange of Research Information Format, ERCIM News No.35.
- Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik (Hrsg.) (2001): Wege zu einer besseren informationellen Infrastruktur. Nomos, Baden-Baden.
- Krause, Jürgen (2003): Suchen und "Publizieren" fachwissenschaftlicher Informationen im WWW. In: Medieneinsatz in der Wissenschaft: Tagung: Audiovisuelle Medien online; Informationsveranstaltung der IWF Wissen und Medien gGmbH, Göttingen, 03.12. - 04.12.2002. Wien: Lang. (IWF: Menschen, Wissen, Medien), erscheint Spätsommer 2003.
- Krause, Jürgen; Niggemann, Elisabeth; Schwänzl, Roland (2003): Normierung und Standardisierung in sich verändernden Kontexten: Beispiel: Virtuelle Fachbibliotheken. In: ZfBB: Zeitschrift für Bibliothekswesen und Bibliographie, 50, Nr. 1, S. 19 - 28.
- Mandl, Thomas (2001): Tolerantes Information Retrieval. Neuronale Netze zur Erhöhung der Adaptivität und Flexibilität bei der Informationssuche. Dissertation. Konstanz: UVK, Univ.-Verl. (Schriften zur Informationswissenschaft; Bd. 39).
- Matthews, Brian M. (2002): Integration via Meaning: Using the Semantic Web to deliver Web Services. In: Wolfgang Adamczak, Nase, Annemarie (eds.): Gaining Insight from Research Information. Proceedings of the 6th International Conference on Current Research Information Systems, Kassel University Press. Kassel, S. 159 - 168.
- Musgrave, Simon (2003): NESSTAR Software Suite. <http://www.nesstar.org> (January 2003)
- Ronthaler, Marc; Zillmann, Hartmut (1998): Literaturrecherche mit OSIRIS; ein Test der OSIRIS-Retrievalkomponente. In: Bibliotheksdienst, 32 (7): 1203 - 1212.
- Rysevik, Jostein (2002): Bridging the Gap Between Data Archive and Official Statistics Metadata Traditions. Power Point Presentation. <http://www.nesstar.org> (January 2003).
- Schöning-Walter; Christa (2003): Die DIGITALE BIBLIOTHEK als Leitidee: Entwicklungslinien in der Fachinformationspolitik in Deutschland. In: ZfBB: Zeitschrift für Bibliothekswesen und Bibliographie, 50, Nr. 1, S. 4 - 12.
- Sheridan, Páraic; Ballerini, Jean Paul (1996): Experiments in Multilingual Information Retrieval using the SPIDER System. In: Frei et al. 1996. S. 58 - 65.
- Zillmann, Hartmut (1997): OSIRIS. Ein gemeinsames Projekt der Universitätsbibliothek Osnabrück und des Instituts für Semantische Informationsverarbeitung der Universität Osnabrück (ISIV). Osnabrück 1997.
http://www.ub.uni-osnabrueck.de/osiris/osiris_papers.html