How should libraries respond to new forms of publication? Some reflections on the expectations of users and providers and what could be realized at the present state of technical development

Schmiede, Rudi; Stoll, Julia; Sonkajärvi, Hanna; Körnig, Stephan

Preprint / Preprint
Konferenzbeitrag / conference paper

How Should Libraries Respond to New Forms of Publication?

Some Reflections on the Expectations of Users and Providers

and

What Could Be Realized at the Present State of Technical Development.

Julia Stoll [Correspondence Author]

Working Group telos (technology of the electronic library organization and the semantic web)

University of Technology Darmstadt, Merckstr. 25, 64283 Darmstadt, Germany

jstoll@ifs.tu-darmstadt.de

Hanna Sonkajärvi

European University Institute, Villa Schifanoia, Via Boccaccio 121, 50133 Florence, Italy

hanna.sonkajarvi@iue.it

Stephan Körnig

Working Group telos

University of Technology Darmstadt, Merckstr. 25, 64283 Darmstadt, Germany

koernig@ifs.tu-darmstadt.de

Rudi Schmiede

Working Group telos

University of Technology Darmstadt, Residenzschloß, 64283 Darmstadt, Germany

schmiede@ifs.tu-darmstadt.de

Abstract

The paper focuses on the possibilities and problems created for the humanities by the usage of Digital Libraries. The exchange of information, forms of discussion and the process of publication in the scientific communities have changed over the recent years. On the other hand the user group of digital libraries is heterogeneous and difficult to define. We discuss the organization structure and the working process of two communities, which have quickly adapted their communication process to the use of web-technologies: the Open Source community and the scientific research community. We have to distinguish the communication technique, the applications of the World Wide Web (WWW), including search engines, and the barely started developments to realize new Web Services for DL-oriented applications, to support the library services and finally to obtain a new concept, which we call long-term access to digital structured resources by contract. A DL user views a Web Service as a “black box” and the service has an interface. The user has only to know the preconditions so that the postconditions will support his/her requirements. We conclude by stating that we should accept that publishing and distributing structural digital resources/information is a new way of (scientific) publication based on collaborative structures. Libraries should respond to this development by providing dynamic metadata in a decentralized network of libraries.

1 Introduction

Today scientific research requires an exchange of results between disciplines all over the world, at any given time and place enabling dialogue between different researchers, working groups, institutions and universities. One way of exchanging scientific results is by publishing. We observe that the exchange of information, forms of discussion and after all the process of publication in the scientific communities have changed over the recent years. For the characterization of the usability of the libraries we consider the following aspects: What kinds of libraries exist and what are the groups of users/visitors addressed by them? How are library services influenced by technical developments? Who uses what kinds of library services? What kind of changes can be observed in the needs of the scientific community using a library?

The paper focuses on the possibilities and problems created for the humanities by the usage of Digital Libraries (DLs). We argue that important changes have taken place in the way scientific community works with media. The libraries have so far not responded to this shift. Therefore, we start by considering different forms of libraries, Sec. 2. We discuss the term “public use” concerning the needs of visitors/users of libraries. Sec. 3 focuses on the parallel operations and differences between the classical publication process of a publisher and the direct definition of a digital publication by the researchers themselves. We discuss the process of scientific community building through the publication and distribution. We present the development of Open Source software establishing a new working process in Sec. 4. This working process correlates with the way information is handled in the scientific community today. We transfer the inherent working concepts to the process of publishing scientific results. As a result we formulate requirements for a working process defining the structure of a digital resource. This concept introduces preconditions and postconditions for structuring the digital resources, so that an encapsulation is possible. This encapsulation is based on the use of

metadata. The considerations presented in Sec. 4 result in a new concept Sec. 5, named (long-term) access of digital structured resources by contract. Such a concept of contracts allows the implementation of decentralized networks for the scientific communication and the installation of trusted archives. In Sec. 6 we conclude by arguing that the future of the institutional library depends on its ability to respond to challenges posed by the development of digital media and to provide dynamic metadata, which are web-processed.

2 Forms of Libraries and the Problem of "Public Use"

We have to distinguish between several forms of so called "libraries". We start by taking into consideration libraries located in buildings and run by librarians. We call these organizations "traditional libraries".. Further, we have to distinguish between academic and public libraries. Academic libraries are often located in universities and financed by the universities themselves which in their turn are supported by countries and states. The purpose of the academic libraries is to serve educational and research demands. On the other hand public libraries are organized, located in and financed by cities, churches, companies and other non-governmental organizations. Such libraries address different kinds of audiences. Visitors of a city library might be the local inhabitants interested in information concerning their area of residence, or local newspapers and journals, fiction, music and other entertainments. The needs of a company are steered by their business interests. The aim of such a library should be to help the employee to do a successful job. A library affiliated to a church addresses the parishioners. Additionally, national libraries are responsible for the preservation of the "cultural heritage" of the nations. Such libraries are reference libraries. The national legislation directs the collection of the presentation copies.

If library services are supported by an (computer-based) information system or applied to

the WWW, this kind of a library is often named a “virtual” or a “digital” library (DL). In traditional libraries the information search is based on metadata, whereas in DLs there is no coherent scheme of metadata available up until now even though some implementations, such as Dublin Core, exist. Thus, the access to information always remains limited by the availability of the additional descriptive and administrative (meta)data provided. DL users cannot access information that is not administrated using metadata. The search on the WWW requires that metadata of structured documents and other structured digital resources are handled and presented in an accessible form for the user. Assuming that the metadata is available, everyone with access to the Internet and a browser can use a search engine to find structured information on the WWW. Today, it is still an open question if the user will find what (s)he is looking for when using freely accessible search engines like Google, Yahoo etc.

We conclude that the term “public use” is not well-determined. In the case of “traditional libraries” the target group and the location of the institution library are mutually conditional and allow a specialisation between different libraries, where as residence does not play a role in the use of a DL. As a consequence, the librarians know their group of visitors and their needs, where as the user group of the DLs is heterogeneous and hard to define. Everyone is a potential DL user, but not everybody goes to a university library. The development of the usability and the services of a DL, thus requires knowledge on the composition of different target groups using the WWW as information resource. In the following we question the organization structure and the working process of two communities identified as such groups, the Open Source community and the scientific researcher community. The Open Source community plays an essential role in the development and promotion of web technologies. The researcher community represents a typical example of a group who has quickly adapted their communication process to the use of such technologies.

3 The Ways of Communication and Publishing in the Scientific Community

3.1 The Researchers' Needs and their Role defined by and in a Scientific Community

Scientific discussions constitute a social process. Doing research is based on discussing problems, defining tasks, finding solutions and exchanging results. Discussions are based on informal communication between researchers and cooperative teams. Communication is supported by the use of adequate communication techniques: writing letters, telephoning, and exchanging e-mails, if there is no possibility of face-to-face communication. Cooperation is needed for answering complex questions. Internet-techniques are used worldwide for exchanging information and publishing research results on the WWW. Web-based joint-activity tools, like e-mail and search engines enable every researcher to look for recently published digital structured resources. Although the forms of distributing research results have changed over the recent years, scientific results rarely address the mass media. They are developed for others who belong to the scientific community. The question is how a community is built using different ways of publication. A researcher is accepted in a scientific community, where other researchers read and refer to his/her scientific results. In such a manner a validation or falsification takes place. Furthermore, a member of the (scientific) community is defined by his/her activities. Therefore, his/her activities have to be made public in a referable form. The communication establishes the community of experts by using references to already published results so that researchers share a common context in a given field of interest. In the following we consider the methods and mechanisms of the publication process located in a scientific community.

3.2 The Publication Process within a Scientific Community

We distinguish at least four phases in the exchange process of scientific information:

- a creation phase (authors and researchers)
- a phase of fabrication, advertising and distribution (publishers/publishing companies and researchers located in universities)
- a phase of evaluation (publishers/editors/proof-reader or researchers/readers), and
- a preservation phase ("traditional libraries" and digital libraries).

The following groups participate in the publication process:

- developers of communications and web-based technologies,
- authors, researchers, reviewers, editors, and publishers,
- publishing companies, universities,
- other forms of organization, e.g. scientific communities, using web-based technologies, and
- libraries and librarians.

We observe a divergence of services in the sphere of digital structured resources and scientific libraries. Many facilities of publication and distribution of information are available for the classical print media**,** but they have their counter parts in the field of digital media, like electronic newspapers, e-magazines and e-books.

Publications in the form of digital structured resources have introduced a new dynamic in the information exchange of the (scientific) community. Such publications are available in different forms and places. Documents not only permit a certain degree of validation through the scientific community, but they also provide a basis of conversation by contributing to the acquisition of knowledge. The merging process of the services of the institution-based libraries and the services of digital libraries delivering web-based documents results in changes in the use and handling of scientific publications, which is in its initial phase. Some differences can be observed in the process of publication. We identify two streams in the publication process for

distributing scientific results. One of these methods can be characterized by the term “review-before-commit” and the other method by “review-on-commit”.. The sequence of the phases of presenting, distributing and reviewing is different according to the publication method used. In both publication processes the researcher/reader is a reviewer, who validates the resource by studying the results. The method of review-before-commit describes the common way of publishing by publishing companies. This publication process includes the phases of writing, reviewing (by a publisher), printing (or presenting) and commercial distribution. There are at least three levels in a scientific publication process:

- a basic level, where discussions and personal communication take place in form of face-to-face communication,
- an intermediate level by experts, for example conference-based publications (program committee, selected reviewers of the discipline)
- publication level, article in a reviewed magazine (editors are experts in their particular field) or in a book, where the publishing company ensures that the reader gets a reviewed text and takes care of the marketing and the dissemination of the product.

The community of reviewers is located in the research community and in the publishing companies. The document is “closed” when it is printed/presented. If changes are required the process has to be restarted. The author has to produce a (complete) new document and send it for reviewing (to a publishing company). The company decides when the document will re-published again. Thus, the decision to publish is not taken by the author.

The method “review-on-commit” is well established in the Open Source community as we explained above. The digital structured resource becomes a dynamic “object” and its publication a dynamic process. The following table contrasts the phases of either using the method “review-before-commit” or “review-on-commit”

Phases in "review-before-commit"	Phases in "review-on-commit"
1. Writing/producing	1. Writing/producing
2. Reviewing	2. Presenting (on the WWW)
3. Printing/presenting	3. Distributing by discussion
4. Distributing by selling	4. Reviewing and modifying

Using the method of “review-on-commit” the modifications are made by the author and/or by the reviewer in the existing document. The document is immediately presented again on the WWW. In principle, evaluation is performed throughout the distribution process. The process of publishing is “open” for almost every researcher who is interested in the subject of the document in form of a digital structured resource. This is why the slogan “’Publish early, publish often’ is a perfectly reasonable way to treat documentation releases as well”, according to Deb Richardson, leader of the Open Source Writers Group. Of course, the model of “review-on-commit” can bring problems. By assuming that a digital structured resource is “dynamic”, there might be no rollback, so that (scientific) results could be fragile.

4 Innovation Provided by the Open Source Community

4.1 The Open Source Phenomenon

In recent years the development and implementation of software in an Open Source manner has gained importance. Open Source systems have been introduced in industry, at universities and in governmental organizations at several levels. The success of this software is based on the philosophy of „publish early, publish often”.. The approach is based on the maximal delegation of tasks and requires mutual trust between the developers.[i] The development of Open Source software is subversive. But, nobody starts the implementation of new features from scratch. The

Linux system is based on Linus Torvalds' transmission of Tanenbaum's Minix and Unix as a well established (computer) operation system and a huge amount of GUI/GNU packages initiated by Richard Stallman. The GNU organisation started in 1991. The first Linux operation system was successfully "distributed" in 1994. Special distribution organizations with the responsibility of building interoperable applications have been established. Their qualification is that they know what kind of hardware and software work together. Distributors - "RedHat", "Mandrake", "Debian", "SuSE" etc. - do not sell complete system architectures and software applications. They distribute so-called packages (of binary code ready for installation) and sell configuration tools for a more or less simple installation of an operation system. This kind of software development is realized via integration of applications into a complete system and the implementation of working processes that span new forms of collaborative work and do not require a common working location.

4.2 The Working Process beyond the Open Source Phenomenon

As we have seen (sub)systems are technically distributed, but constitute a common working environment that supports specific tasks. The working process behind the Open Source approach of solving problems is the following. When someone recognizes a problem and defines a task (s)he designs an interface that expresses certain information and/or requirements, checks or conditions at an abstract level independent of the later implementation. Therefore, so-called assertions are used in the form of preconditions, postconditions and invariants to increase the system in a distributed environment. Note that the implementation of every system component is independent of the person who signalled the given problem to the community. It is only necessary to reason about the compatibility of interfaces both statically and dynamically to provide an acceptable level of reliability. A digital resource interface is defined as XML schema

or a DTD, where the representation and the content are divided.[ii] In principle, the resource interface is the only portion visible to other resources, like communication mechanisms and representation facilities.[iii] The interface of a digital resource should be able to express what services are available from the resource and also specify the behaviours of such services by using the concepts of contracts. Every resource should provide a service as a part of a contract with other resources. When the preconditions of the digital resource have been satisfied according to the contract, it becomes the resource's responsibility to ensure that the postcondition is fulfilled.[iv] This contract specifies exactly what is required and guaranteed before and after each invocation in form of preconditions and postconditions.

4.3 The Transfer of the Concept of Contracts to Digital Resources

We believe that a digital resource should be able to provide an interface that contains as much structured information as is required to communicate to other resources, while encapsulating any content or services that is not public. Such an interface includes: (1) the resource name, its public attributes as different types of e.g. metadata and public methods for handling such metadata and their corresponding types; (2) a specification of behaviour of the resource typically in terms of preconditions and postconditions; assertions represent a fulfilled contract.

Assertions can be considered as a part of the digital resource interface. On the other hand the preconditions and the postconditions are viewed as being independent of the digital resource implementation. Thus in a distributed environment preconditions and postconditions are often presented by an interface language.[v] We have to distinguish refinements of encapsulation in a distributed environment. Often different people, groups or (scientific) communities are involved in the development. One of these people can be viewed as an "architect", who designs or creates the interface, presents the interface service provided for a digital resource, including

preconditions and postconditions and invariants. Others might be understood as implementers, who take the designs or a specification, implement the code for each of the services. Others can be seen as users, who use certain services of the implemented digital resource. This results in three different views of the same interface of the digital resource: the public view, a restricted view, and an internal view. The public view is exactly what the user of a digital resource can access: its name (or identity [technical named: Uniform Resource Identifier, URI][vi], public attributes, and public services (for accessing the content of a digital resource by using the DLs). The restricted view includes the public view of a digital resource together with the services which have been provided for the use of the interface - the URI and the methods for representing its content. Thus the content required for contract definition is available but may be hidden form the public interface. The internal view is typically one which is visible to the implementer only and which remains hidden from the restricted and the public views.

The realization of the interface is defined by its specified preconditions, postconditions and invariants. In practice we apply a so-called black-box concept for the identification of such preconditions and postconditions. The “black box”-approach is necessary for the exchange of information between different interfaces. We have to apply such concepts to the definition of digital structured documents and resources. This means we have to implement a syntactical description for access to the digital structured resource. For this we do not have to know anything about its content. The access of a digital structured document or resource is controlled by using a URI, see above. The content of the digital structured resource is characterized by its (descriptive) metadata. The metadata are defined in the form of preconditions. In a distributed environment much effort is devoted to passing message. But we can trap errors at the interface using assertions before the message including a digital resource has been sent to the target resource. It is possible to use this kind of a specification to automatically produce objects of digital resources based on

specific interfaces containing metadata. In the next section we present a brief overview of the role of metadata in general. For a more profound explanation of metadata see Dublin Core (DC) and MARC.[vii].

4.4 Metadata and their Ambiguous Role

In general, metadata is data about data. In our context of the DLs a digital resource contains content as such data. Common categories of metadata include descriptive metadata (e.g. bibliographic information), structural metadata (information about formats and structures), and administrative metadata (which includes rights, permissions and other information used to manage access).[viii] One item of metadata is the identifier, which identifies an item to the outside world (URI see above[ix]). Further metadata contain information about the resource they describe. The metadata must conform to a specification of syntax. It enriches the origin data and can define new structures. Such definition of a structure is the basis to "understand" the ambiguous role of metadata. See the following example: Catalogue records and abstracts are usually considered to be metadata because they are used to describe the content of a digital resource, but in an online catalogue they are data. The problem is that data about data are only determined by their use in a certain application. In our case the application is a DL and such a (scientific) DL is defined by the requirements of the research communities. Researchers would be particularly interested in metadata in digital form for the support of a reference service so that digital resources would be referable.

4.5 Web Services

Web Services are defined by the W3C Web Service Architecture Working Group, "a software application identified by a URI, whose interfaces and bindings are capable of being defined,

described, and discovered as XML artifacts. A Web Service supports direct interactions with other software agents using XML-based messages exchanged via Internet-based protocols" (Web Service definition from W3C Web Service Architecture Requirements, W3C Working Draft).[x] A large part of web-based services are in process of being realized via integration within applications, working processes or institutions and enterprises that span new forms of collaborative work all over the world, at any time, in any place.[xi] From the technical point of view application software has been developed to allow systems running in different environments to interoperate via XML.[xii] Furthermore, other web standards are still in a process of development[xiii] and will influence future web-technology and the connected services.

5 Long-term Access to Digital Structured Resources by Contract

A distinction has to be made between the communication technique and the applications of the WWW, including search engines. The web standards for the integration of different applications based on the concept of Web Services are in process of being developed. DL users are often researchers established in (scientific) communities, who are able to apply the functionalities of information systems and their applications on the WWW. Generic portals are a helpful and useful development as they support the information exchange based on digital structured documents. The term "generic portal" means the supply of the standardized Web Services so that different applications have been integrated in (scientific) DLs. A part of such services is the integration of DL-oriented web-based applications. These new Web Services might also support the following concept, which we call long-term access to digital structured resources by contract. A DL user views a Web Service is as a "black box" and such a service has an interface. The user has only to know the preconditions so that the postconditions will support her/his needs. Applying a Web Service means to agree to a contract. For the contract to be agreed, we assume that we can

identify authors, researchers and editors as “architects” or creators of digital structured resources.

An open question is who takes care of the refinement of metadata so that the contract can be established in form of Web Services for DLs. Presumed that the concept of agreeing to a contract represented in preconditions and in postconditions is accepted, then long-term access could be defined by a contract as well. For this there exists a complete concept of so called “trusted archives”.. This concept implies that it is the author’s responsibility to create the preconditions and postconditions for the digital structured resources. The trusted archives represent a technically decentralized as well as a decentraly organized network of institutions for the long term preservation of digital resources.[xiv]

6 Conclusion

We conclude by stating that we should accept that publishing and distributing structural digital resources/information is a new way of (scientific) publication based on collaborative structures. This new way of publication requires new forms of assertion. These assertions also include new forms of organizing the administration of digital resources. Such organizations have to be more or less decentralized and located in a distributed environment.

In this paper, we have avoided the usage of the terms “preservation” and “archiving” because there are different approaches to the usage of digital technologies in the community of librarians and the research community. This means that the primary aim of the researchers is to exchange results. The results of these processes do not necessarily need to take the form of a physically existing, long time storable media. But, at the same time they have to be referable. Therefore, the problems for the scientific community do not lie in the preservation of a fragile digital resource, but in the access to the metadata. In contrast to this, librarians tend to think in terms of physically existing (final) products. For this final product a single set of metadata is

produced.

What is needed, are dynamic metadata schema – which allow a continuous updating along the publication process based on the method of “review-on-commit”. Traditional libraries could play an essential role as members of a decentralized network of libraries providing the dynamic metadata. In the future, the librarians work could consist of the support of the scientific working process by delivering metadata in form of a Web Service-based standard. The standardization of Web Service-based standards has barely been started.

[i] Raymond, E. (1999). *The Cathedral and the Bazaar – Musings on Linux and Open Source by Accidental Revolution*, O'Reilly, Beijing, Cologne.

[ii] Canonical XML Version 1.0 [15 March 2001] http://www.w3.org/TR/XML-c14n; Extensible Markup Language (XML) 1.0 (Second Edition) [06 October 2000] http://www.w3.org/TR/REC-xml.

[iii] Extensible Markup Language (XML) 1.0 (Second Edition) [06 October 2000] http://www.w3.org/TR/REC-xml; Resource Description Framework (RDF) Model and Syntax [22 February 1999] http://www.w3.org/TR/REC-rdf-syntax.

[iv] Mayer, B. (1997). *Object-Oriented Software Construction*. 2. Ed. Prentice Hall Professional Technical Reference, ORT XXX.

[v] Resource Description Framework (RDF) Model and Syntax [22 February 1999] http://www.w3.org/TR/REC-rdf-syntax; Web Service Description Language (WSDL) Version 1.2, Part 1: Core Language [11 June 2003] http://www.w3.org/TR/2003/WD-wsdl12-20030611; Part 2: Message Pattern [11 June 2003] http://www.w3.org/TR/2003/WD-wsdl12-patterns-20030611; Part 3: Bindings [11 June 2003] http://www.w3.org/TR/2003/WD-wsdl12-bindings-20030611.

[vi] Resource Description Framework (RDF) Model and Syntax [22 February 1999] http://www.w3.org/TR/REC-rdf-syntax

[vii] Dublin Core: http://www.dublincore.org; MAchine Readable Catalogue: http://www.loc.gov/marc

[viii] Arms, W. Y. (2002). *Digital Libraries*. MIT Press, McGraw Hill, 2. Ed.

[ix] Resource Description Framework (RDF) Model and Syntax [22 February 1999] http://www.w3.org/TR/REC-rdf-syntax.

[x] Resource Description Framework (RDF) Model and Syntax [22 February 1999] http://www.w3.org/TR/REC-rdf-syntax; [WD1] Web Service Description Language (WSDL) Version 1.2, Part 1: Core Language [11 June 2003] http://www.w3.org/TR/2003/WD-wsdl12-20030611; Part 3: Bindings [11 June 2003] http://www.w3.org/TR/2003/WD-wsdl12-bindings-20030611.

[xi] Web Service Description Language (WSDL) Version 1.2 [11 June 2003]; Web Content Accessibility Guidelines 2.0 [24 June 2003] http://www.w3.org/TR/WD-WCAG-20-20030624.

[xii] Extensible Markup Language (XML) 1.0 (Second Edition) [06 October 2000] http://www.w3.org/TR/REC-xml.

[xiii] Web Service Description Language (WSDL) Version 1.2 [11 June 2003].

[xiv] Beagrie, N.; Doerr, M.; Hedstrom, M.; Jones, M.; Kenny, A.; Lupovici; C.; Russell, K.; Webb, C.; Woodyard, D.; Dale, R.; Bellinger, M. (2002). *Trusted Digital Repositories. Attributes and Responsibilities*. A RLG-OCLC Report, RLG Mountain View CA. http://www.rlg.org/longterm/repositories.pdf..; Lavoie, B.; Dale R.; Alexander, M.; Bradley, K.; May, M.; Guenther, R.; Hruley, B.; Lupovici, C.; Rieger, O.; Sergant, D.; van der Werf, T.; Webb, C.; Wendler, R.; Woodyard, D. (2002). *Perservation Metadata and the OAIS Information Model. A Metadat Framework to Support the Perservation of Digital Objects*. A Report by The OCLC/RLG Working Group on Perservation Metadata, RLG Mountain View CA. http://www.oclc.org/resaerch/pmwg.

References

Classical References

- Arms, W. Y. (2002). *Digital Libraries*. MIT Press, McGraw Hill, 2. Ed.
- Mayer, B. (1997). *Object-Oriented Software Construction*. Prentice Hall Professional Technical Reference, Upper Saddle River, New Jersey, 2. Ed.
- Beagrie, N.; Doerr, M.; Hedstrom, M.; Jones, M.; Kenny, A.; Lupovici; C.; Russell, K.; Webb, C.; Woodyard, D.; Dale, R.; Bellinger, M. (2002). *Trusted Digital Repositories. Attributes and Responsibilities*. A RLG-OCLC Report, RLG Mountain View CA. http://www.rlg.org/longterm/repositories.pdf.
- Lavoie, B.; Dale R.; Alexander, M.; Bradley, K.; May, M.; Guenther, R.; Hruley, B.; Lupovici, C.; Rieger, O.; Sergant, D.; van der Werf, T.; Webb, C.; Wendler, R.; Woodyard, D. (2002). *Perservation Metadata and the OAIS Information Model. A Metadat Framework to Support the Perservation of Digital Objects*. A Report by The OCLC/RLG Working Group on Perservation Metadata, RLG Mountain View CA. http://www.oclc.org/resaerch/pmwg.
- Raymond, E. (1999). *The Cathedral and the Bazaar – Musings on Linux and Open Source by Accidental Revolution*, O'Reilly, Beijing, Cologne.

World Wide Web Consortium (W3C) Recommendations (http://www.w3.org)

- Canonical XML Version 1.0 [15 March 2001] http://www.w3.org/TR/XML-c14n.
- Extensible Markup Language (XML) 1.0 (Second Edition) [06 October 2000] http://www.w3.org/TR/REC-xml.
- Resource Description Framework (RDF) Model and Syntax [22 February 1999]

http://www.w3.org/TR/REC-rdf-syntax.

W3C Published Working Drafts (http://www.w3.org)

- Web Service Description Language (WSDL) Version 1.2.
    - Part 1: Core Language [11 June 2003] http://www.w3.org/TR/2003/WD-wsdl12-20030611.
    - Part 2: Message Pattern [11 June 2003] http://www.w3.org/TR/2003/WD-wsdl12-patterns-20030611.
    - Part 3: Bindings [11 June 2003] http://www.w3.org/TR/2003/WD-wsdl12-bindings-20030611.
- Web Content Accessibility Guidelines 2.0 [24 June 2003] http://www.w3.org/TR/WD-WCAG-20-20030624.

Others

- Dublin Core: http://www.dublincore.org.
- MAchine Readable Catalogue: http://www.loc.gov/marc.