# Building integrative environment for the improvement of scientific communication

Radovan Vrana, MSc
Department for information sciences,
Faculty of philosophy, Zagreb, Croatia
rvrana@ffzg.hr

## Introduction

Judging from the increased number of electronic sources of information available on the Internet (On August 24th 2001. 10836 electronic journals and bulletins have been registered available online.[1]) it has become evident that electronic publications have strong influence on the process of scientific research and communication as well as on teaching at universities and on other segments of our lives that include and promote collaboration in electronic environment.

Today, scientific community represents a special but significant group of users that is of special interest to computer science and information science professionals who try to do research in influence of information technology on everyday work routine in scientific community in order to improve the quality of communication. One of the key reasons for choosing scientists for the focal point of this paper is a fact that significant portion of new information technology is being designed, tested and used in scientific community. In addition, scientific community is generally considered to be a source of new knowledge and an environment that promotes creation and transfer of knowledge by use of information technology.[2]

This paper will present some of the key points of changing process of scientific communication with focus possibilities of communication and collaboration on the Internet which is now the main medium that enables better and easier communication with publishers, libraries and with colleagues plus with all other parties interested in transfer of knowledge.

## Scientific communication and maturing of electronic communication environment

The roots of the existing paper-based (journal) model of scientific communication can be traced to the 17th century and appearance of the first true scientific journal. Since then the same model has been improved mainly by introduction of new printing and publishing techniques and new ways of dissemination of published knowledge. Two and a half centuries later, publishing industry based on paper products such as books and journals has reached its peak. Advent of the Internet and its wide acceptance has only accelerated the existing communication processes by offering new means for information transfer. These new means helped scientists to communicate using e-mail, video conferencing, mailing lists, different databases available online etc. Another important event happened in 1987, when the first electronic journal "New horizons in adult education" appeared marking the beginning of an era of electronic publishing.[3] This event also marked the beginning of the process of modification of the existing scientific communication and publishing routine based on paper and unquestionably introduced the Internet as the new electronic working environment supporting better

[1] <URL:http://gort.ucsd.edu/newjour/ > (24.8.2001.)
[2] Valente, Adriana; Luzi, Daniela. Different contexts in elecotronic communication: some remarks on the communicability of scientific knowledge. // Journal of documentation 56, 3(2000), 305.
[3] Okerson, Anne. Are we there yet? Online resources ten years after. // Library Trends 48, 4(2000), 673.

and easier collaboration and creation of scientific papers in many new ways among colleagues from different institutions.[4]

The last 5-6 years were very intensive, as the Internet has finally become a widely accepted mass media. The most important asset of the Internet is its potential for the development of new custom-made user services and modes of communication and collaboration. Moreover, the Internet has been recognized as a phenomenon which determines the way of life of certain user groups serving them for the exchange of ideas, work results, discoveries and general aspiration for research. We may ask ourselves (judging by the number of the users and user groups on the Internet[5]) who might those users be and what do they expect from the Internet as an advanced communication environment? In order to start understanding their information and communication needs, and improve the existing communication and collaboration environment or design a new one (e.g. Internet 2), it is necessary to distinguish different forms of information sources which might be relevant for scientists using the Internet as their primary choice in search for quality information for research.

## Where it all starts - computer mediated communication

Computer mediated communication (CMC) is one of the freshly coined terms related to the scientific communication, and also related to some other forms of human communication. Basically there are two viewpoints in regard to the CMC:[6]

- Utopian vision: knowledge, competence and learning are considered as three complementary ways of cognitive transaction; the network creates new anthropological space, a new system of human proximity based on techniques, meanings, language, culture, conventions, representations and human emotions.
- Pragmatic vision: global network is recognized as the most recent form adopted to organize the knowledge system.

While the first viewpoint focuses on the knowledge in its broadest meaning, the second one draws attention to the knowledge that has already been codified and is organized according to certain criteria.

Indeed, when we take a closer look at the Internet, we may see that it offers a large quantity of content. However, it offers relatively small set of tools providing help in organizing that content. Therefore we can complain that the Internet actually represents a loosely organized chaos lacking proper organization and some advanced system of links between quality (in science – peer reviewed) primary information sources. The biggest problem is the inability to measure the exact size of the Internet[7] and to set the size of the sample necessary for the research in general communication patterns and the volume of the content available online. Sometimes, it is also difficult to find a solution to many interoperability issues of data exchange between heterogeneous information sources and the very nature of information sources: whether they are formal or informal, built by an individual or and institutions, are they regularly funded and with what amount etc. Taking into account all these problems and issues in research of the phenomenon called the Internet, we can say that numerous attempts to organize the content and provide persistent links to other information sources on this global network have reached certain level of success but these levels are still unacceptable when compared to quality standards of organization of the paper-based scientific publishing.

---

[4] Meadows, Jack. Can we really see where electronic journals are going? // Library management 18, 3(1997)m 152.

[5] According to the last available data from Nielsen/Netratings there are 495 million users on the Internet at this moment (http://www.nielsen-netratings.com/hot_of_the_net_i.htm - 29.8.2001.)

[6] Valente, Adriana; Luzi, Daniela. Different contexts in elecotronic communication: some remarks on the communicability of scientific knowledge. // Journal of documentation 56, 3(2000), 300.

[7] Inspite of many Internet statistics, it is impossible to measure the size of the Internet since many Web pages are simply out of reach of the most famous search engines and directories. This part of the Internet is also known as «deep Web».

The current schematic view of organization of content and hyperlinks towards different information sources can be presented on several levels:

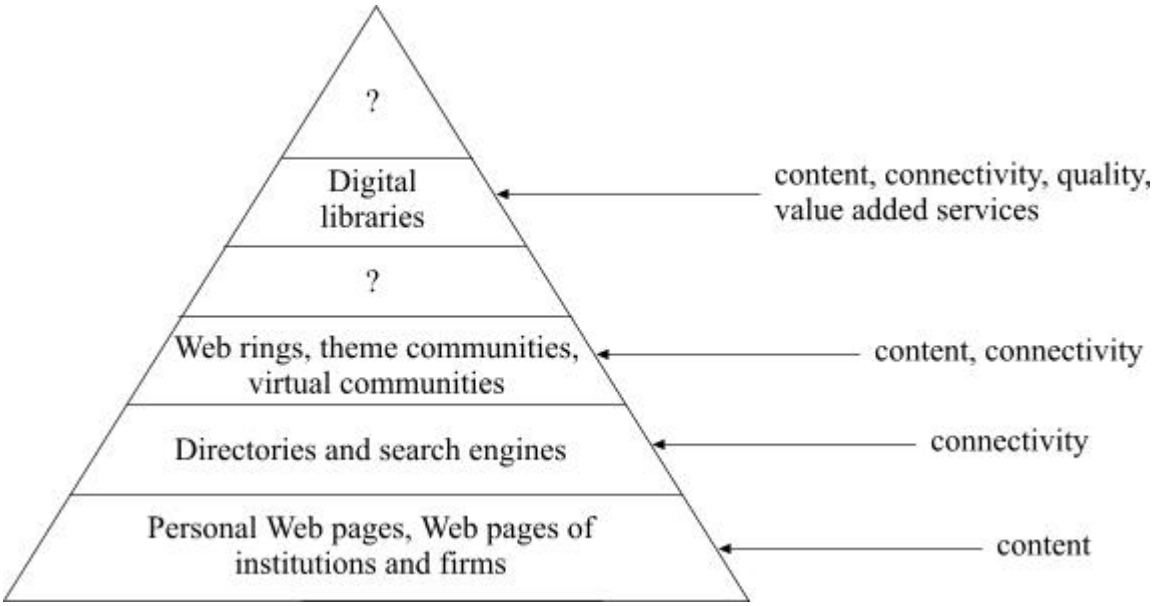


Figure 1. Organization of content and hyperlinks towards different information sources – categories are presented in their broadest meaning

Although these are not by any means all forms of content organization and content linking, such approach to the analysis of the current state of the maturity of the Internet as the possible primary source of information in science can still helps us to enumerate the most common forms of content organization and content linking on the Internet. This diagram can also clarify the position of digital libraries as one of the most prominent new information source/organization, which offers quality content, connectivity, and possibility of development of the custom-made user services, so much needed by scientists.

In addition to the numerous papers covering the topic about the Internet and digital libraries, it is still important to present some of the key advantages of the Internet as a mass medium for the delivery of information in comparison to traditional libraries and in cooperation with libraries.[8]

- Material is delivered quickly,
- Material is delivered directly to a computer station without patron having to collect material from around the library building,
- Material is accessible 24 hours a day, during the whole year.

Except enumerated advantages, the author Robert McGeachin points out several additional open questions, which can be easily mirrored to any other discussion about advantages of the Internet in comparison to the traditional paper-based information sources. Still, there is a strong need of mentioning some additional functional requirements that will help such an global electronic environment like the Internet (and all new organizations such as digital libraries which also use the Internet) to function as good as possible especially for specific user groups – scientists.

[8] McGeachin. Selection criteria for Web-based resources in a science and technology library collection. // Issues in science and technology librarianship. (1998) <http:www.library.ucsb.edu/istl/98-spring/article2.html> (15.11.2000.)

Some of these open questions cover:

- Additional software requirements
- Additional hardware requirements
- Regulation of user access rights
- Funding the cost of Internet materials
- User education and assistance
- Stability of Internet location
- Long term access
- Product licensing
- Cataloging of Internet materials
- Delays in access due to heavy Internet traffic.

Only when the most of these requirements are met we can consider having a quality electronic working environment prepared for scientific collaboration. Are all the requirements absolutely necessary for a successful collaboration among scientists in electronic environment? Most of them are prerequisites for long-term functioning of any collaborative environment and they represent the actual framework for building of integrative electronic environment for the improvement of scientific communication. Not meeting some of these requirements may in some cases be a restraining factor in development of collaboration between heterogeneous user groups and communication environments on the Internet.

**From bibliometrics to cybermetrics and beyond – exploring the possibilities of connecting the scientific research results in electronic environment**

One of the key methods for identifying core sets of articles, authors and journals in particular fields of study is called bibliometrics.[9] Bibliometrics explores number and forms/types of citations of authors and their works in paper-based (printed) sources of information - journals and books. By use of mathematical and statistical methods, it is possible to find out which are the most cited authors in a particular discipline or within certain academic field. This method exists for a long time, and it appears to be very suitable for doing comparative researches in order to gather data about quantitative presence of authors in a particular period of time in a particular publications.

Cybermetrics represents a step further by applying most of the elements of bibliometrics to the electronic environment. In case of the Internet and publishing of works in for instance HTML format, cybermetrics can show us, which authors are most linked to, and which authors make electronic citations in a particular area or discipline. Such results can be graphically presented as clusters of authors working together on a problem or issue even if they aren't from the same discipline. This valuable information can help us prove the thesis that scientific disciplines are losing their boundaries and that interdisciplinary collaboration becomes inevitable in science today. Existing electronic communication and collaboration environments like the Internet and organizations like digital libraries can only substantiate such direction in the development of scientific collaboration and publishing.

If we take into consideration current state in the development of scientific publishing and ever increasing number of printed paper-based journals which still have the leading role in development of scientific communication,[10] we can notice the emergence of new areas and topics related to the Internet as an environment suitable for the publishing of scientific papers, their linking with other papers and information sources and growth of collaboration between authors. In spite the growing number of papers published on topic of electronic publishing and in spite the improving conditions for

[9] Rowlands, Ian. Who can count the dust of Jacob? From bibliometrics to cybermetrics. // The Internet : its impact and evaluation. / ed. by David Nicholas and Ian Rowlands. London: ASLIB, 2000. 114
[10] Tenopir, Carol. Moving toward electronic journals. // Library journal 125, 12(2000), 36.

the publishing of papers on the Internet, it seems that authors still avoid publishing in the electronic environment.

In order to attract authors to publish their works in newly established electronic journals, it is necessary to develop methods with which it is possible to measure visibility of works of certain authors in question on the Internet and their linking with works of other authors, their other works, full text databases in which works are stored etc., as bibliometrics does it for the printed paper-based publications. The results of such researches, which will apply such methods, will directly help the development of organizations like digital libraries and other forms of organized electronic sources of information in the global electronic environment.

In her article about generations in the development of electronic journals Ellen Finnie Duranceau noticed that we are now using the third generation electronic journals (the first generation which appeared in mid 90s included mainly plain ASCII text format distributed by e-mail, and the second generation made its breakthrough on the Internet in form of HTML and PDF).[11] This is very important fact that should be taken into consideration when planning the future of the development of electronic journals in the next 5 years and seeing them as new pillars of scientific publishing, and, of course, when thinking about problems which make scientists discard electronic journals for publication of their works.

She also expects from the third generation to be flexible in linking the content of particular works i.e. it is of utmost importance to enable backward and forward linking (speaking about the time line) of the content of works published in electronic form and to avoid introduction of radical new interfaces when accessing electronic journals until the second generation interfaces are fully absorbed.

Such content connectivity across time and different platforms is also a component of the digital library structure which should enable users to access documents stored in other digital libraries by help of customized services.

[11] Duranceau, Ellen Finnie. Third generation e-journals and beyond: is there agent in your future? // Serials review 25, 2(1999), 77-82.

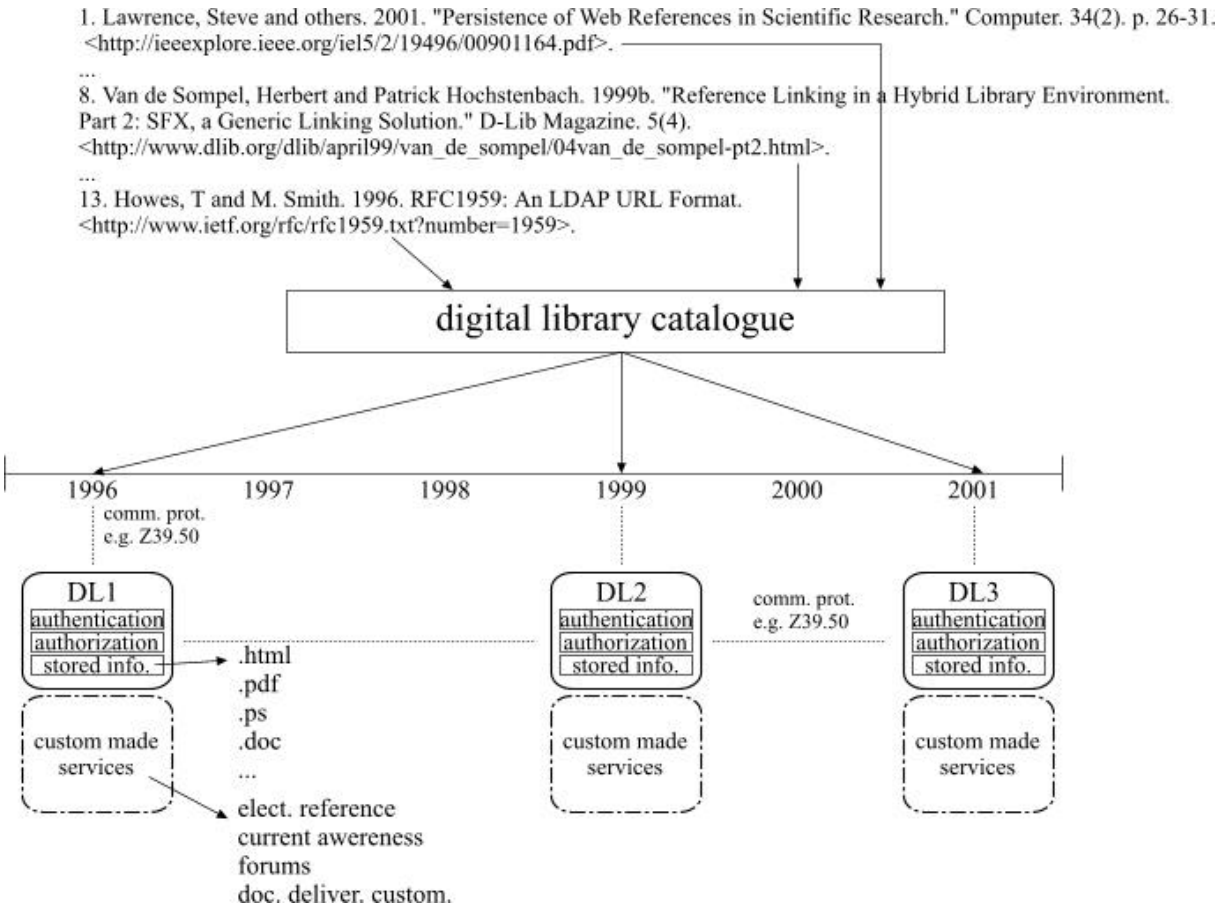


Figure 2. Backward and forward document accessibility in digital libraries

At the same time, Duranceau puts a question of quality control of the works published on the Internet. This is a key issue since without it very few scientists will agree to a long-term collaboration in the electronic environment. Can a peer review system, which represents the guarantee of quality in the world of paper-based publishing be mirrored into the electronic environment and there be widely accepted? Can digital libraries, the most prominent new electronic information organization and source, become the centers of excellence of scientific work based on electronic peer review system? Answers to these questions are affirmative, but in order to achieve these goals, there is still a lot to be done especially in negotiating terms of collaboration among all parties involved.

Let us take a closer look at some of the excellent examples of collaboration on the Internet that already exist and that can give us better insight in pros and cons of the new generation collaboration.

## Theme communities on the Internet

Theme communities represent the first step and an attempt to organize content for a particular group of people interested in one topic or area in order to make the content more available to users of that group. Since the most of the content on a particular topic is widely scattered across Internet, this initiative provides help for scientists (and other user groups) who need content and pointers to other information sources in one place. Many of such theme communities are informal in their nature and are built and maintained by enthusiasts (as a contrast, libraries maintain electronic references on the Internet). The idea of having related content in one place has most certainly come from libraries and therefore there are good chances that scientists will recognize new generation libraries i.e. digital libraries as an environment suitable for their work.

## Virtual communities

Another form of collaboration on the Internet is a virtual community. Virtual community is formed by integration of several organizations' competencies and resources thus forming virtual network of independent sources.[12] Virtual communities appeared 15-20 years ago and they were introduced when they enabled people to work at home by the use of information technology. There are several types of virtual communities summed up in three key points:[13]

1. First type are organizations that outsource some of their business activities, forming virtual alliances to achieve organizational objectives,
2. Second type is related to a conceptual organization that is abstract and exists only in minds of those who form a particular organization; it is antithesis of the physical organization,
3. Third type includes organizations built up by virtual links through use of IT; IT contributes to organizational structures and provides virtual mechanism for realizing organizational objectives.

The third type is (accompanied by definition of virtual communities/organizations) of greatest interest to us in this papers since it touches the very essence of the current problems on the Internet related to availability of quality information sources i.e. their linking.

In order to see how other sources of information are linked to the Croatian institutions present on the Internet, we have chosen one major search engine and three sample URLs. The first URL belongs to the National and university library in Zagreb, which is the central and most important library in Croatia. The second URL belongs to the library of "Rugjer Boškovic" institute. There we can find a great number of organized links towards Croatian libraries and some other information sources to satisfy information needs of scientists and librarians. The third URL belongs to the Croatian-British joint project of making a portal for exchange of scientific links.

Links to all three URLs were tested by use of Google (www.google.com) search engine by issuing "link:URL" command. As a result, Google returned all URLs in its databases linked to each of our sample URLs. In this manner we can see who has put a link to our information sources and we can also see the type of organization who is behind originating URL. The result would be a partial picture of organizations and (possibly) individuals interested in content provided by those three Croatian URLs.

Since the search engine presented several hundred URLs as a result, only first 100 were chosen as data for the URL count and categorization across domains.

| National and university library in Zagreb (11.7.2001.) | |
|---|---|
| link:www.nsk.hr – total of 326 results, only first 100 counted | |
| national domains | 72 |
| .edu domain | 15 |
| .org domain | 7 |
| .com domain | 6 |

In many cases national domains hide presence of different organizations, which can again, in some cases be very important and related to URL we tested. For instance, 72 national domains represent URLs of: magazines/newspapers (1/1 link), government bodies/institutions/ministries/agencies (13 links), faculties/universities (2/4 links), libraries (27 links), conferences (1 link), portals/directories/uncategorized pages with variable content (13 links), commercial firms (10 links)

[12] Shao, Y.P.; Liao, S.Y.; Wang, H.Q. A model of virtual organizations. // Journal of information science 24, 5(1998), 305.
[13] Ibid.

and personal Web pages (1 link). Some of them (universities and libraries) are more interested in linking to the sources of the National and university library in Zagreb than others.

| Library of the "Rugjer Boškovic" institute (11.7.2001.) | |
|---|---|
| link:nippur.irb.hr – total of 215 results, only first 100 counted | |
| national domains | 86 |
| .com domains | 7 |
| .edu domains | 3 |
| .net domains | 2 |
| .org domains | 2 |

In this case, 86 URLs with national domains represent: European organizations (6 links), faculties/universities (3 links), professional associations (1 link), directories/ uncategorized pages with variable content (8 links), journals (1 link), libraries (27 links), government institutions/agencies/institutes (39 links) and personal Web pages (1 link).

However, we must point out one specific and very important issue when counting the number of unique URLs linked to our test URL nippur.irb.hr. 62 of 86 national domain URLs represent the case of "self linking" i.e. URLs provided by the search engines represent links from the "Rugjer Boškovic" institute itself (between different organizational units and even inside library Web site itself) thus decreasing the number of unique URLs to 24.

| HEUREKA portal for exchange of scientific links on the Internet (11.7.2001.) | |
|---|---|
| link:heureka.foi.hr – total of 1 result | |
| national domain | 1 |

The last example represents the most concrete effort to collect links toward the scientific information sources. However, there was only one Web site recorded in Google's database linked to this project's URL, and this Web site is the URL of portal itself (heureka.foi.hr). Other important Internet information sources such as the National and university library in Zagreb are not aware of the existence of such Web site.

In spite of relatively small number of different institutions linked to the first two URLs, it is worth noting that both information sources have well organized content and list of active links towards other information sources (this is especially true for the library of the "Rugjer Boškovic" institute) which is one of the goals modern information sources should have. On the other hand, they use different database tools, different communication protocols for information exchange and different file and document formats, which can be an obstacle in the process of linking of their content in future thus creating the virtual communities of the third type.

## Issues regarding advancement of communication and development of digital libraries

As it was presented in the previous sections of this paper, scientists and other user groups feel strong need to communicate in advanced electronic environment in order to achieve better availability of their discoveries and to exchange ideas with other colleagues. So far, we have presented two types of content and link organization (theme communities, virtual communities), which are predecessors of many other types of organizations offering content like digital libraries, the most advanced form of libraries today.

There are several issues in regard to digital libraries and scientific communication and collaboration, which should be discussed.

First question is related to the concept of digital libraries in wider scientific community (except at universities which are cradles of digital library projects, mainly in the USA) and of course the notion of advantages in communication and delivery of original quality content for users outside the primary circle of interest for such services i.e. for common users. Extensive literature review has shown that the authors of digital library projects almost always point out that digital libraries are intended to become centers for exchange of scientific information and guarantee of quality of material they will offer through their services. One of the goals of the development of digital libraries is that become online research libraries i.e. electronic environment in which journals and books will be online and all citations will be converted into active links.[14] Another important goal is a transparent linking between digital libraries and their services and users, as well as enabling universal access to digital libraries and their services.[15] At the present moment digital libraries lack promotion in various scientific disciplines in order to encourage interdisciplinary research by use of their information sources stored in form of digital collections.

Second question is related to the interoperability issue. Earlier collaboration forms on the Internet has shown numerous problems in relation to the diversity of file formats and communication protocols. These issues still remain unresolved although there are some solutions especially among standards and protocols (Z39.50, DIENST etc.) that can help us avoid conversion and inability to connect to other sources. These problems are visible mostly in modern libraries that try to connect their online public access catalogues and holdings databases. This is especially true in Croatia where different types of libraries have different software applications and it will be difficult to put the information about libraries' collections on the Internet without extensive record conversion.

Third question is not actually a problem but it is rather a topic, which requires attention and additional research. It's about the question of access to electronic sources of information. Earlier in this paper it was mentioned that we are now using the second generation of e-journals, which use mostly HTML and PDF file formats. It was pointed out by the same author that interfaces toward such primary electronic resources shouldn't be changed until fully accepted and "consumed" by users. In that sense, we can say that digital library can benefit from the current state of development of the Internet technologies and browsers and use them to attract both scientific and non-scientific population. In addition, digital libraries have another major advantage when compared to publishers and individual Web sites offering electronic content: library material is well organized, there is also trained staff, which can help users in information retrieval. For those who don't need or want help, material in digital libraries is usually organized in such manner that users can browse and search content for themselves 24 hours a day, directly on their computer screens. They can use full text databases with various document delivery mechanisms, electronic references to locally and remotely stored documents, dictionaries, thesauri, forums for exchange of ideas etc.

Since the very beginning, digital libraries have been oriented in their projects and realization toward interdisciplinary collaboration and there is no special need for proving that they kept the same direction until now. What is important is that digital libraries have constant need for professionals from different disciplines since their development is influenced by changes in information technology. Digital library as an organization carries collaboration in its core and it would be appropriate that scientists and professional from different field take part in its development in future.

Each of these questions or issues represents a small but significant part in structure of existing and future frameworks for the advancement of scientific communication. We have seen the beginnings with personal Web pages, directories, search engines, portals and different type of online communities, and every online form has attributed to the development of new collaboration tools in the electronic environment.

---

[14] Crane, Gregory. Commercial digital library and the academic community. // D-Lib magazine 7, 1(2001) <URL:http://www.dlib.org>

[15] Sloan, Bernie. Service perspectives for the digital library remote reference services. // Library Trends 47, 1(1998)

## Conclusion

As the Internet becomes more complex and advanced communication environment, different groups of users seek different ways of organizing content and establishing links between related content. What was once known as a personal Web page with loosely organized hyperlinks today is known as portal or directory. Today, we also have user many communities on the Internet with foci on various topics. All these forms of collaboration on the Internet illustrate the necessity for finding at least one form of content organization and linking that will at the same time serve smaller user groups (scientists) as well as wider Internet population. Such a form could be a digital library.

Digital libraries are dynamic organisms which structure is constantly changing. As their structures grow, so are developing the conditions for the development of new electronic services. Most projects for digital library development were started in academic community, so most of the projects have been strongly influenced by scientists and their vision of new electronic source of information and therefore digital libraries can be adapted for the needs of scientists more easily than any other information source. They can help build their own collections with electronic documents of proved quality and develop new services required for their research and teaching at the universities. They can also help implement some of the content linking schemes such as OpenURL[16] in order to build network of permanent links between information sources.

And how can scientists help digital libraries to become centers of scientific communication? They can influence organizational and thematic segment of development of digital libraries by:

1. Using information stored in digital libraries
2. Citing information sources in digital libraries
3. Influence and encourage development of peer review in electronic environment
4. Publish their works in electronic form and store them in digital libraries
5. Participating in coordination of standards for storage and exchange of data in electronic environment
6. Helping librarians in development of their profession by giving them support in negotiations with publishers about copyright and other important issues
7. Carrying out projects in collaboration with digital libraries and by using them to communicate the results of their projects.

Scientists shouldn't avoid publishing in electronic environment and use of quality information sources at their hand. The framework for the advancement of scientific communication is still in development but there are already some solutions that will help more effective communication between scientists. We should user them more extensively.

## Literature:


1. Crane, Gregory. Commercial digital library and the academic community. // D-Lib magazine 7, 1(2001) <URL:http://www.dlib.org> (29.8.2001.)
2. Duranceau, Ellen Finnie. Third generation e-journals and beyond: is there an agent in your future? // Serials review 25, 2(1999), 77
3. McGeachin. Selection criteria for Web-based resources in a science and technology library collection. // Issues in science and technology librarianship. (1998) <URL:http://www.library.ucsb.edu/istl/98-spring/article2.html> (15.11.2000.)


[16] http://www.sfxit.com/openurl/openurl.html (31.8.2001.)

4. Meadows, Jack. Can we really see where electronic journals are going? // Library management 18, 3(1997), 151-154.
5. Okerson, Anne. Are we there yet? Online resources ten years after. // Library Trends 48, 4(2000), 671-693.
6. Rowlands, Ian. Who can count the dust of Jacob? From bibliometrics to cybermetrics. // The Internet : its impact and evaluation. / ed. by David Nicholas and Ian Rowlands. London: ASLIB, 2000. 114-130.
7. Valente, Adriana; Luzi, Daniela. Different contexts in electronic communication: some remarks on the communicability of scientific knowledge. // Journal of documentation 56, 3(2000), 299-311.
8. Shao, Y.P.; Liao, S.Y.; Wang, H.Q. A model of virtual organizations. // Journal of information science 24, 5(1998), 305-312.
9. Sloan, Bernie. Service perspectives for the digital library remote reference services. // Library Trends 47, 1(1998), 117-143.
10. Tenopir, Carol. Moving toward electronic journals. // Library journal 125, 12(2000), str. 36.