

INTERLENDING & DOCUMENT SUPPLY
INTERNATIONAL CONFERENCE
1-5 October 2001 - Ljubljana, Slovenia

Evaluation of a Internet Document Delivery Service

S. Mangiaracina^{*}, P. Salamone^{*}, M. Buzzi[§], F. Gennai[§], L. Abba[§]

^{*} Research Area of Bologna Central Library. Research National Council. Bologna ITALY

[§] Institute per le Applicazioni Telematiche. Research National Council. Pisa ITALY

contact persons: mangiaracina@area.bo.cnr.it, marina.buzzi@iat.cnr.it

keywords: Internet document delivery, statistics gathering, performance measurement

ABSTRACT

An Internet-based Document Delivery Service (DDS) has been developed within the framework of the CNR (the Italian Research National Council) Project BiblioMIME, in order to take advantage of new Internet technologies and promote cooperation among CNR and Italian university libraries. Adopting such technologies changes the traditional organisation of DDS and may drastically reduce costs and delivery times.

An information system managing DDS requests and monitoring the temporal evolution of the service has been implemented, running on the local-area network of a test-site library. It aims to track number and types of documents requested and received, user distribution, delivery times and types (surface mail, fax, Internet), to automate repetitive manual procedures and to deal with the various accounting methods used by other libraries. Transmission of documents is carried out by means of an e-mail/Web gateway system supporting document exchange via Internet, which assists receiving libraries in retrieving requested documents.

This paper describes the architecture and main design features of the e-mail/Web gateway server (the BiblioMime server). This approach permits librarians to continue using e-mail service to send large documents, while resolving problems that users may encounter when downloading large size files with e-mail agents. The library operator sends the document as an attachment to the destination address; on fly the e-mail server extracts and saves the attachments in a web-server disk file and substitutes them with a new message part that includes an URL pointing to the saved document. The receiver can download these large objects by means of a user-friendly browser.

We further discuss the data gathered during the triennium 1998-2000; this consists of about 5,000 DDS transactions per annum with 300 other Italian scientific and bio-medical libraries and commercial document suppliers. Use of the instruments described above allowed us to evaluate the performance of service "before" and "after" the use of Internet Document Delivery and to extract some critical data regarding DDS. Those include:

a) libraries with which we have greater numbers of exchanges and their turnaround times;

- b) extraordinary reduction in costs and delivery times;
- c) the most frequently requested serial titles (allowing cost-effective decisions on new subscriptions);
- d) impact on DDS of library participation in consortia which allow user access to greater numbers of online serials.

1. Introduction

The BiblioMIME project aims at *“the development of advanced services of Document Delivery among the CNR libraries, based on Internet transmission”*, in order *“to reduce management costs and to achieve short turnaround times in satisfying Document Delivery requests from final users and libraries”*.

The BiblioMIME project is being developed under the coordination of the CIRT commission of the Italian Research National Council (CNR). The CIRT is the Study Commission to develop the computer network facilities at CNR, and has the following goals :

- To guarantee network connectivity among all the CNR Institutes and to guarantee their interconnection to the Italian network of the Scientific Research and Education Ministry (GARR).
- To coordinate network facilities management; to define the operating services and to design new ones to be offered to the CNR user community; to define and implement network security policy.
- To promote and participate in those research projects utilizing the computer network, which have a follow-up in the ordinary activities of CNR.

The BiblioMIME project falls within the last category .

2. Document Delivery service: quality indicators and problems in traditional DD

Document Delivery Service plays an important role in Italian research and university libraries. The main Internet resources for Document Delivery are the Archivio Collettivo Nazionale dei Periodici (ACNP, the collective serials Italian catalogue), and the Sistema Bibliotecario Nazionale Catalogue (SBN). ACNP records 565.937 serials holdings, which are indicated by 2.344 Italian research and university libraries, and accounts for 250.000 online searches/month. The first paper catalogue was realized by CNR in 1990; the online version was then developed by the University of Bologna. ACNP is now evolving towards an integrated tool to access various services, such as electronic serials holdings, abstracting services, and SSD (Serial Services Directory) [1].

The organisational models of DD services may be very different within each library, since they depend on the number of library patrons, the kind of patrons, the Institution directives, and so on. However, typical parameters of DD service are:

- Total number of borrowing or lending requests carried out

- Proportion of total number of requests to successful ones
- Unitary costs
- Personnel
- Use of national or international resources of DD services among libraries (such as ACNP, SBN...)
- Use of commercial document suppliers
- Policies for requesting or delivering documents
- Policies for payment
- Use of computer supported automation tools or management software
- Electronic transmission of requests/documents

The measurement of performances of DD services, the adoption of guidelines, such as the IFLA [2], or the design of a librarian "Carta dei servizi" [3] is under discussion among librarians in the last few years, as documented in the AIB-CUR archive, the Italian librarian discussion list. However, studies promoted abroad by ARL or other important organisations go much further and present very interesting experiences in the analysis of "best practices". The Interlibrary Loan and Document Delivery (ILL/DD) Performance Measures Study, a two-year effort to measure 1995/96 performance of ILL/DD departments in North American research and college libraries [4] and [5], examined four performance measures: cost, fill rate, turnaround time, and user satisfaction. The Study also examined differences between libraries and identified characteristics of low-cost, high-performing ILL/DD operations in order to suggest strategies for other research and college libraries for improving local performance. Materials were of returnable (ILL service) or not returnable (DD service) type. The basic indicators taken into consideration by the Study are:

1. *Direct cost*, that is the cost a library pays for a DD request (borrowing cost) or a DD delivery (lending cost) of a document;
2. *Fill rate*, that is the per cent rate of satisfied requests or delivery of documents;
3. *Turnaround time*, that is the medium of working days needed by a library to satisfy a DD request;
4. *User satisfaction*, based on rapidity, quality/integrity of documents provided, and interaction with library personnel.

Other interesting indicators are *weekly fill rate*, measuring the average number of requests a DD service is able to fulfil in a single week, and *percentage of requests sent to commercial suppliers*, which supposedly rates the grade of inter-library national or international cooperation (low percentage means high grade of cooperation).

However, several management problems arise in DD service organisation, some of a structural nature, such as the chronic shortage of personnel afflicting many Italian research libraries. Problems concern :

- Difficulty in assigning adequate personnel resources, especially when DD volume of requests increases;

- Service costs, due to:
 - Personnel costs, which can be reduced by the implementation of computer-supported procedures for DD;
 - Delivery costs, i.e., the expense of copying and sending non-returnable material. We have estimated that a library with an average number of 200 articles sent per month, that uses half per cent the fax and half per cent surface mail, spends about L.1.000.000 /month.
- Turnaround time. When a library requests material of other libraries turnaround time can rarely be calculated, while turnaround time in delivering material to other libraries depends on the transmission system used.
- Copy quality and integrity also depend on the transmission system chosen.
- Different refunding rates and means : in Italy there is a wide variety of these, most of which are time-consuming...
- The management of all "paper" related to DD transactions. Paper must be filed (also for Copyright reasons) and easily retrievable; it could thus be elaborated in order to satisfy these requirements. For instance, a library with an average DD of 4.000 articles requested/delivered per year must manage about 40.000 pages to be delivered to users or other libraries, 4.000 request letters sent/received, 4.000 answer letters sent/received, 4.000 refund communications and operations to be dealt with.... an incredible amount of paperwork could thus be avoided.

3. Goals of the BiblioMIME Project

The BiblioMIME project was created in 1999 with the first prototype system for electronic document delivery, whose main characteristic was its high usability for libraries. In fact, the system releases libraries from the very "technical" tasks of understanding and dealing with transmission protocols such as SMTP or FTP. The prototype was developed by the CNR Institute for Telematic Applications. After the prototype system underwent a short experimentation period by two libraries (the central library of the CNR Research Area in Bologna and the library of the MASPEC Institute in Parma), the advantages were immediately clear and the BiblioMIME project was proposed to the CIRT, with the following aims :

1. ***To develop and experiment the use of a document transmission system based on the Internet.*** This could reduce delivery costs such as xerox copying and postal or fax expenses. Receiving libraries benefit the most from such a system, since they would receive high quality documents in very short time. However, if the use of Internet DD spreads to many libraries, this would trigger a *virtuous chain reaction*, improving the overall DD national system and benefiting the final library users of the Italian research community.
2. ***To support the working flow of the activities in the entire DD process, from request to delivery of a document.*** This would reduce working time and eliminate the need for paper archiving. It would also automatically provide data for measuring DD performance according to a uniform system of indicators.
3. ***To suggest an experimental model of DD service,*** promoting inter-library mutual

exchange and cooperation, on the basis of special agreements. For instance, a electronic voucher-like system of payment (such as IFLA vouchers), or guaranteed turnaround times could be experienced by all libraries sharing the model.

4. The Document Transmission System Architecture

The simplest and most common way to send documents via the Internet is by e-mail, although sending very large files could create problems in transmission as well as reception. In fact, the use of FTP or HTTP protocols, while recommended for very large files requires the user to have considerable technical knowledge of the tools.

The proposed solution permits librarians to utilize e-mail service to send very large files, while at the same time resolving any problems that the receiver may encounter when downloading a large file with an e-mail client.

In the following, we introduce the basic and we will explain the problems that may arise when sending a very large attachment by e-mail.

4.1. The electronic mail system

In electronic mail we can identify two important components of the delivery process:

- The transport protocol SMTP (Simple Mail Transfer Protocol) [Pos 82] [Cro 82] used to send messages. Depending on the network configuration the sent message can pass through one or more mail servers (Message Transfer Agents) to reach the destination mailbox;
- Mailbox access protocols such as POP (POP Office Protocol) [Mye 96] and IMAP (Internet Message Access Protocol) [Cri 96] which enable users to access their own remote mailboxes, using a user-friendly e-mail client.

However, the transfer of large files via e-mail can present some limitations in both transport and mailbox access protocols:

- Mail servers can impose message size limitations. Sometimes problems can emerge when delivering large messages over congested networks. A possible solution is to fragment messages at the sender MTA, recomposing them at the destination MTA. Both MTAs must support the MIME (Multipurpose Internet Mail Extensions) message fragmentation/de-fragmentation mechanisms.
- Mailboxes access protocol time-outs. POP offers off-line and pseudo on-line ('leave messages on server' option) operational modes. Full messages (header and body) are downloaded in a sequential order. In transferring large files, time-out parameters set into the POP client can interrupt the connection making it impossible to download the message and those following. In this case the user will not be able to download new messages until the server's administrator removes the blocking message.

The IMAP protocol offers on-line, off-line and disconnected operational modes. In contrast to POP, IMAP selects remote messages by downloading only mail headers. Users can activate a selective download of the full body message or body parts. This implies that a user using IMAP is able to verify the message structure in order to understand whether it

is composed of large parts which he or she can then decide not to download – for example, when using a slow dial-up connection.

IMAP is a step ahead of POP in flexibility and features, but nevertheless it does not entirely solve the problem.

- The availability of remote and local resources (such as disk space) can influence the behavior of the Internet Delivery Service.

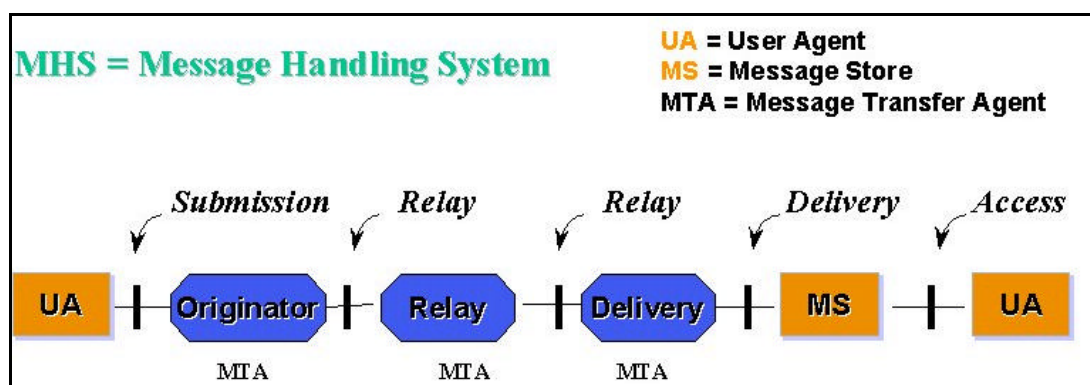


Fig. 1 - Message Handling System (logical scheme)

Our idea is to combine the e-mail service (for document transmission) with the web service (for document downloading). Downloading documents using a web browser is more flexible and reliable than when using one e-mail client. The user can move between different platforms (where a web client is present) without requiring any client reconfiguration (such as that required by e-mail clients).

The fundamental element on which our service is based is MIME (Multipurpose Internet Mail Extensions) the specification for including multimedia files in electronic messages. MIME was conceived as an extension of the format of e-mail messages defined in RFC822 to permit the inclusion of multimedia data in the body of the message, while retaining its compatibility with the standard format. A multipart MIME message is a structured message composed of several parts which can contain different types of data. The RFCs 2045 [Fre96a], 2046 [Fre 96b], 2047 [Moo 96], 2048 [Fre 96c], 2049 [Fre 96d] contain the basic definitions and descriptions of MIME mechanisms and of message format.

4.2. The e-mail/web gateway at work

The use of an e-mail server totally conforming to MIME allowed us to implement the functionality of a gateway between e-mail and web (e-mail/web gateway) by writing a simple script. The system allows the following operations:

1. The librarian receives a request (by e-mail), identifies and extracts the document from the library's archives, converts it into electronic format (by scanner), prepares the message for the requester by inserting the document as an attachment, and finally sends it.

2. On-fly the e-mail server extracts the files contained in the MIME parts of the message and, if they belong to one of the types registered for the service (.gif, .jpg, .pdf, .tiff) it saves them in a part of the disk that is accessible via web. Each extracted part is replaced with a new part of a text/html type containing the URL of the saved document. The parts of the message that do not correspond to one of the registered types remain unaltered within the message. At the end of the body of the message a new part of text/html type is added automatically, containing a brief description of the service, instructions for downloading the software for the visualization of data, and the copyright rules.
3. The resulting multipart message continues on its way on the network until arriving at its final destination.
4. The receiver opens the message, clicks on the URL and downloads the original files from the web-server.

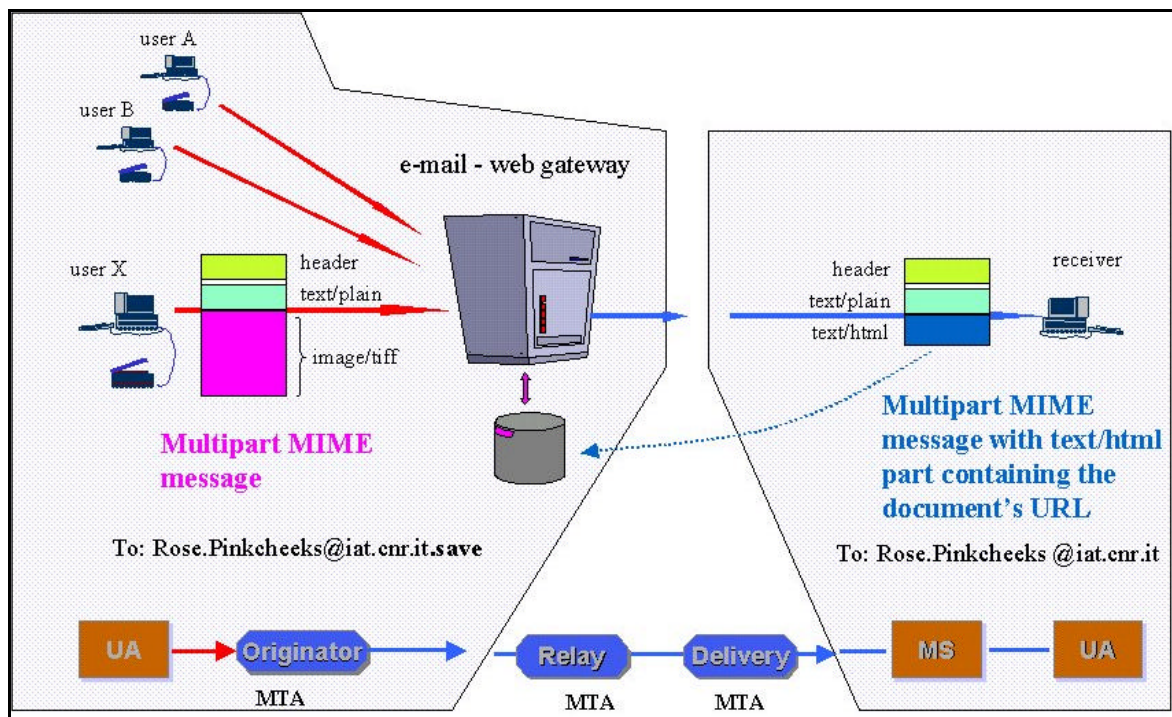


Figure 2 - E-mail/web gateway function

We offer multiple services, on the same e-mail system used to implement the gateway. We host e-mail services for different CNR domains, we offer a MIME attachment conversion service, e-mail to fax gateway service. For this reason we implemented a selective service. In order to have messages processed by our system, two conditions must be met:

- the IP address of the client must be registered in a service authorization list;
- the e-mail destination address must end with the fake top level domain (TLD): **.save** This fake TLD domain (added by the user) will be automatically removed by the e-mail server and it is necessary in order to request that the message pass through our service.

The first condition permits only authorized users to use the e-mail/gateway functionality.

The second condition is important in order to permit the user to switch between regular e-mail operations or those which can extract and save attachments.

At present, documents stored in the web area are not protected. However the filename is based on a unique identifier difficult to guess (i.e. 01K1NEXTLNZK984J6Y.pdf) and the listing of the directory is not allowed. In the near future, we may consider adding a mechanism to enhance document access protection.

Documents are stored in the web directories for a period determined by a system parameter (14 days) and after this time they are automatically removed. All system activities are logged with, and accessible from, protected web pages.

This solution is quite general, and can easily be "personalized" in order to satisfy the specific requirements of an organization.

5. System use and evaluation

Software testing of the Document Transmission System's first prototype was begun in 1999, by the two libraries involved in the project. The data collected by the Document Delivery Service (DDS) of the CNR Central Library in Bologna show the enormous advantages this project offers in terms of costs and reduction in turnaround times.

An information system managing DDS requests and monitoring the temporal evolution of the service has also been implemented in the first phase of the project, running on the local-area network of the test-site libraries. The DDS Management System permits tracking of the number and types of documents requested/received, users' and libraries' internal or geographical distribution, and delivery times and types (surface mail, fax, Internet).

Figure 3 shows the trend of DD service at the Bologna CNR Central Library during the three-year period 1998-2000. The great increase in the number of DD mutual requests confirms a more general trend also noticed by other library systems such as the Sistema Bibliotecario Biomedico Lombardo (50 biomedical libraries) and the Sistema bibliotecario d'Ateneo of Florence University [6], [7]. The increase mainly relates to articles delivered to other libraries, which has grown from 1.264 requests received in 1998 to 2.702 requests counted in 2000. Instead, the number of requests sent to other libraries remains fairly stable (at around 1.500 requests/year), with a slight decreasing trend. This is due to the impact on DDS of library participation in consortia which allow users direct access to greater numbers of online serials (up to tenfold). However, users also gain greater consortium access to bibliographic online databases, so the need to retrieve information through the library remains high. We suspect that this phenomenon would be more marked if we compared the requested titles of serials for the last two years (2000 and 2001, when consortium online access began) with the previous two years (1998 and 1999). The fact that titles requested by our users have "changed" is also proven by the increase in requests to commercial document suppliers, as shown in the following table. Note that our library only started using document suppliers (in our case, the British Library) in 1999.

DD Requests to other libraries	Number of requests	Fill rate	Unit cost in Lira (no personnel costs included)	Document Suppliers requests	Costs Incidence of Doc.Suppl. Requests
1997	2.138	82%			
1998	1.491	84%			
1999	1.693	93%	L. 1.177	3%	65%
2000	1.448	92%	L. 890	4%	79%
2001 (Jan-June)	924	95%	L. 2.212	10%	81%

Figures 4, 5 and 6 show the kinds of exchanges with other Italian libraries which took place during the year 2000. The library has contacts with about 400 libraries: 300 from university, 50 from CNR, 35 various others. It is clear that there is a privileged exchange channel between CNR libraries and the libraries of the local university of Bologna (Figure 4). Figure 6 shows the exchange of volumes with regard to requests from university libraries for copies of documents from our library, with an average frequency greater than or equal to 1 request per week. However, with a few exceptions (Universities of Bari, Padova, Pavia, Roma, Pisa, Ferrara, Napoli), it is clear that reciprocity is rare: we believe the reason is the lack of visibility of these libraries' holdings in the main collective resources for inter-library cooperation (ACNP or SBN).

The impact of Internet Document Delivery (as a provider library) is demonstrated by the following table and the graphic in Figure 7: both show the dramatic cost reduction that took place during the last two years of our DD.

DD requests from other libraries	Number of requests	Fill rate	Unit cost in Lira (no personnel costs included)
1997	281	85%	-
1998	1264	86%	-
1999	1965	87%	L. 1.792
2000	2702	89%	L. 914
2001 (Jan-June)	1217	91%	L. 500

The graphic in Figure 7 shows the number of papers sent per month (Y axis on the left) during the period January 1999-May 2001, depending on the transmission system used by our library: surface mail, fax, or e-mail/web gateway. The total costs in Lire per month (Y axis on the right, red line) are the sum of partial costs per month, where unit costs are calculated thus : L. 3.200 for single document transmission via surface mail, L. 6.000 via fax, L. 0 via Internet. We have estimated average document length to be 10 pages. The Internet Document Delivery effect is such that *expenses have fallen by a factor of 10* (from an

average/month of L. 800.000 in early 1999 to an average/month of L. 100.000 in 2001). A few documents are still sent via surface mail due to their length (usually, when a document is longer than 35 pages, xerox copies+surface mail is still the preferred method). Although we use fast scanners (4 pages/minute speed), when a great number of pages are to be copied, the Xerox machine may be a time-saver for library operators.

The use of the BiblioMIME system for scanning and sending files without the need for specialised software packages has also proved to be cost-effective, since personnel involved in the DD service need no special training, beyond learning to use a scanner and a MIME compliant e-mail program (i.e., Eudora, Outlook, Netscape etc...).

Figure 8 shows, from our library's point of view, the turnaround times achieved in document delivery (lending) with respect to the turnaround times performed by other libraries (borrowing). The graphic shows the percentage of documents delivered on the same working day, by the next one, by the third one, by the first working week, etc.... Turnaround times are calculated by taking the difference between the date of request satisfaction and the date of request arrival. The graphic shows that our library is able to deliver to other libraries 75% of the copies by the first two working days; it is able to satisfy almost all requests (95%) by the first working week. However, looking at the other libraries' turnaround times, they are able to satisfy almost half of the requests (56%) by the first working week, while the remaining part (up to 88%) is delivered by the first working month. Such differences are mostly due to the infrequent use of electronic delivery by other Italian libraries, consequently, their delivery times are greatly affected by the efficiency of the surface mail provider.

6. Conclusions and Future developments

The Document Transmission System developed within the framework of the BiblioMIME project offers libraries the following advantages:

- Efficiency, since it reduces turnaround times;
- Cost effectiveness, since it lowers transmission costs and the process of "scanning + sending electronic copies" saves time for librarians;
- Robustness, since the downloading of documents based on the e-mail/web gateway system has proven to be very reliable;
- High quality of the delivered documents;
- Ease of use for libraries.

We plan to concentrate our future efforts on the design and development of an integrated management system for Document Delivery. An initial prototype system has been implemented and has been in the testing phase for the last few months. The requirements for this integrated system are:

- a)Automation of the complete work-flow of DD activities;
- b)Visibility of the state of DD transactions to requesting and delivering libraries; all requests are filled through online web forms;
- c)Integration with the Internet Document Transmission system;

- d) DD performance measurement;
- e) Support to refund procedures; a reciprocity model should be adopted by all libraries using the system;
- f) Support to work distribution and ease of library personnel training.

ACKNOWLEDGEMENTS

The authors wish to thank all the CNR colleagues who helped with their support and fruitful discussion in this work: Rosella Magno (MASPEC Library), Maria Grazia Balestri, Silvia Roda and Riccardo Ricci (Bologna Central Library), Derek Jones (Bologna Central Library Commission), Claudio Baesso and Patrizia (IAT Institute), Maurizio Lancia and Giuseppe Martini (CIRT Commission).

REFERENCES

- [1] L'evoluzione di ACNP: da catalogo collettivo a strumento di accesso ai servizi. Alessandra Citti e Vincenzo Verniti, Roma, CNR, 6 settembre 2001
<http://www.bibliomime.cnr.it/ws2001/acnpevol.pdf>
- [2] International Federation of Library Associations and Institutions, "International Lending and Document Delivery: Principles and Guidelines for Procedure", DRAFT DOCUMENT,
<http://www.ifla.org/VI/2/p3/illdd.htm>
- [3] Carta dei servizi e servizi in rete per il Document Delivery. Gabriele Mazzitelli, Roma, CNR, 6 settembre 2001
<http://www.bibliomime.cnr.it/ws2001/mazzitelli.pdf>
- [4] : Mary E. Jackson, "Measuring the Performance of Interlibrary Loan and Docum. Delivery Services", ARL's Bimonthly Newsletter of Research Library Issues and Actions, Issue 195, December 1997 <http://www.arl.org/newsltr/195/illdds.html>
- [5]: Mary E. Jackson, "Mesure du Rendement des Services de Pret entre Bibliotheques" Conclusions d'une étude financée par la Andrew W. Mellon Foundation, Association of Research Libraries, Washington, DC 1998
<http://www.arl.org/access/illdd/illna.shtml>
- [6] Il Document Delivery nel Sistema Bibliotecario Biomedico Lombardo Enrica Veronesi, Roma, CNR, 6 settembre 2001
<http://www.bibliomime.cnr.it/ws2001/sbbl.pdf>
- [7] Il Document Delivery elettronico nel Sistema Bibliotecario dell'Università di Firenze

Paolo Salvi, Roma, CNR, 6 settembre 2001
<http://www.bibliomime.cnr.it/ws2001/salvi.pdf>

[Pos 82] J. B. Postel. RFC821: Simple Mail Transfer Protocol. <ftp://ftp.isi.edu/in-notes/rfc821.txt>, August 1982.

[Cro 82] D. H. Crocker. RFC 822: Standard for the format of ARPA Internet text messages - <ftp://ftp.isi.edu/in-notes/rfc822.txt>, August 1982.

[Mye 96] J. Myers & M. Rose. RFC1939: Post Office Protocol - Version 3. <ftp://ftp.isi.edu/in-notes/rfc1939.txt>, May 1996.

[Cri 96] M. Crispin. RFC 2060: Internet Message Access Protocol - Version 4rev1.
<ftp://ftp.isi.edu/in-notes/rfc2060.txt>, December 1996.

[Fre 96a] Freed, Ned and al. RFC 2045. MIME Part One. <ftp://ftp.isi.edu/in-notes/rfc2045.txt>
November 1996.

[Fre 96b] Freed, Ned and al. RFC 2046. MIME Part Two. <ftp://ftp.isi.edu/in-notes/rfc2046.txt>
November 1996.

[Moo 96] Moore, K. RFC 2047. MIME Part Three. <ftp://ftp.isi.edu/in-notes/rfc2047.txt>
November 1996.

[Fre 96c] Freed, Ned and al. RFC 2048. MIME Part Four. <ftp://ftp.isi.edu/in-notes/rfc2048.txt>
November 1996.

[Fre 96d] Freed, Ned and al. RFC 2049. MIME Part Five. <ftp://ftp.isi.edu/in-notes/rfc2049.txt>
November 1996.

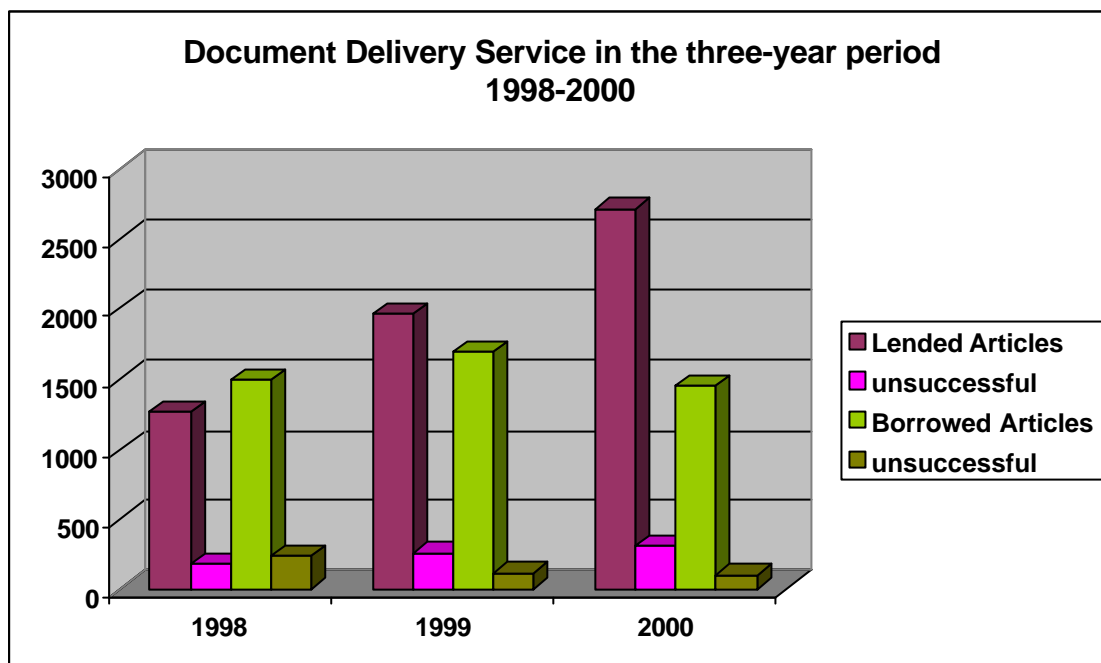


Figure 3

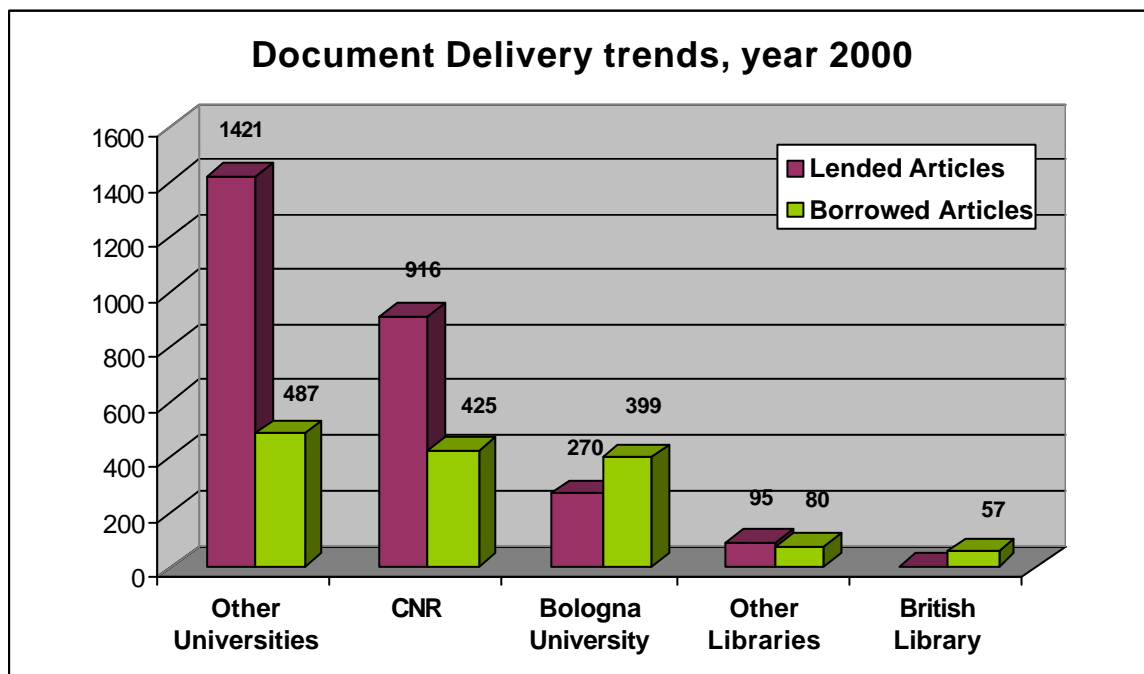


Figure 4

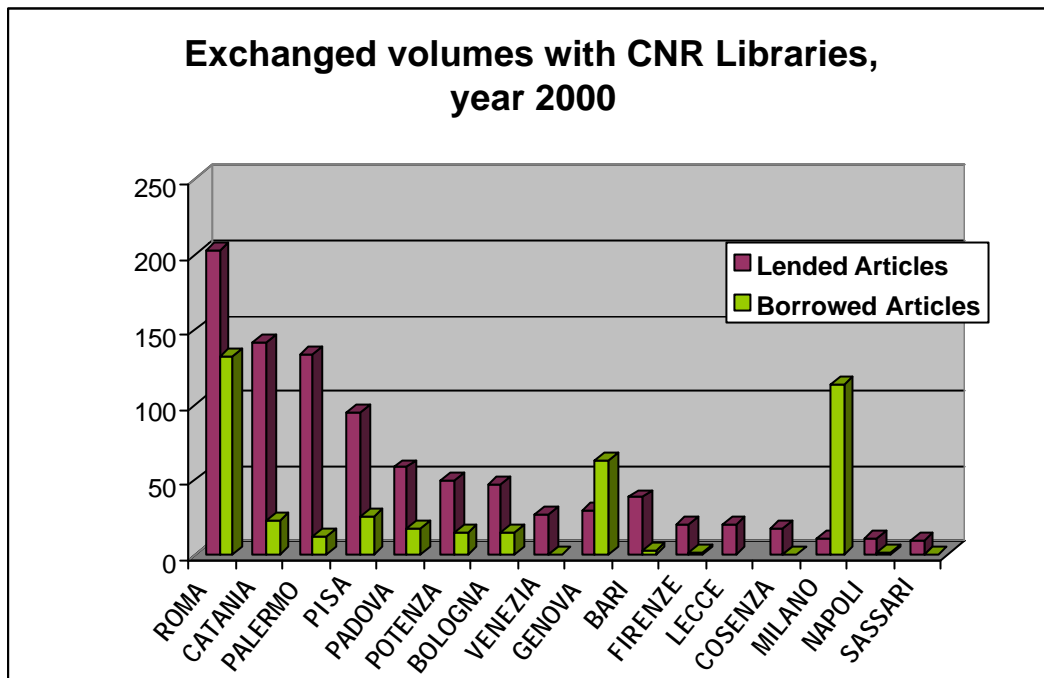


Figure 5

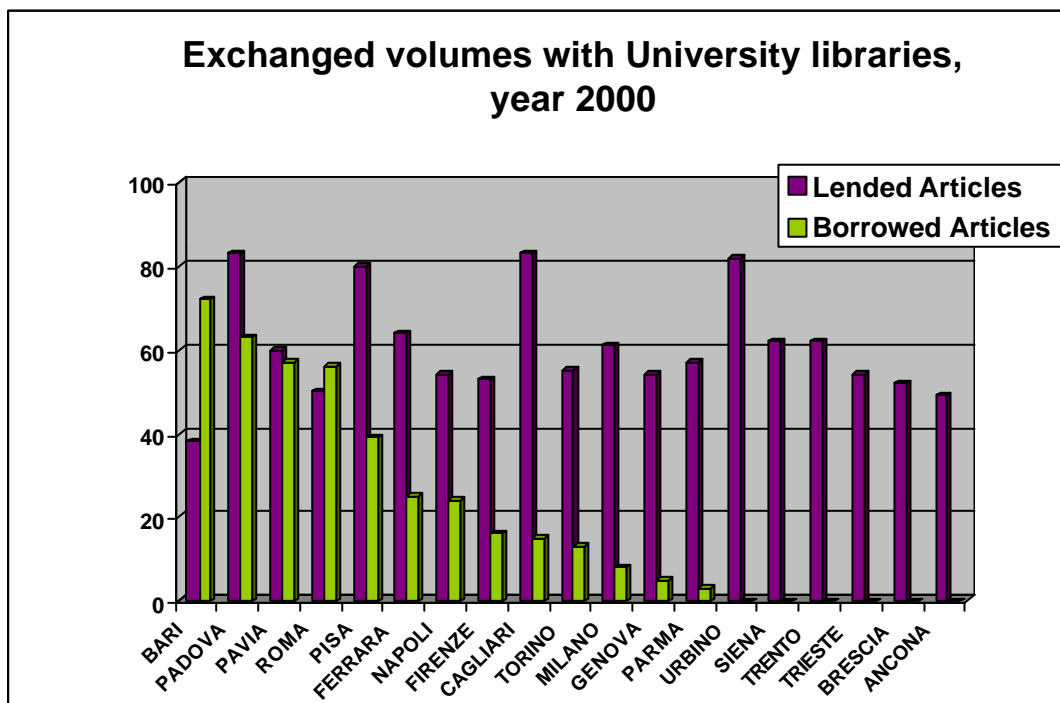


Figure 6

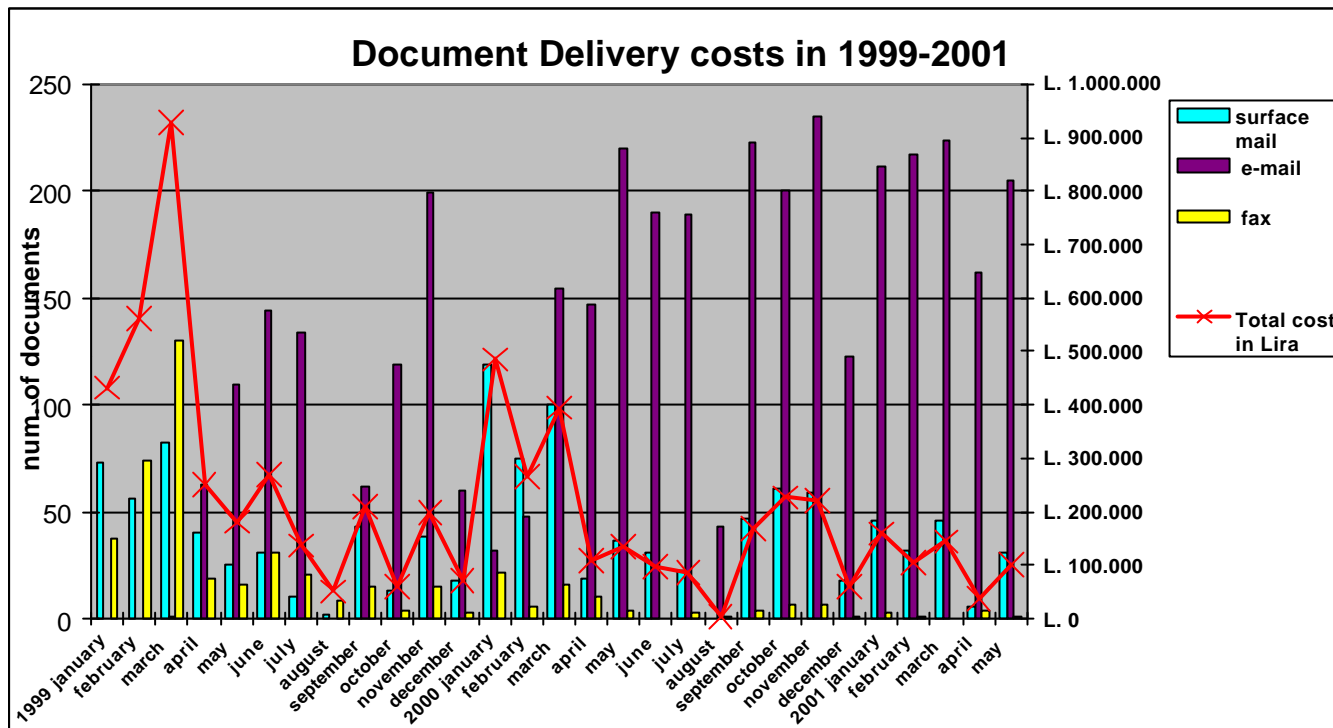


Figure 7

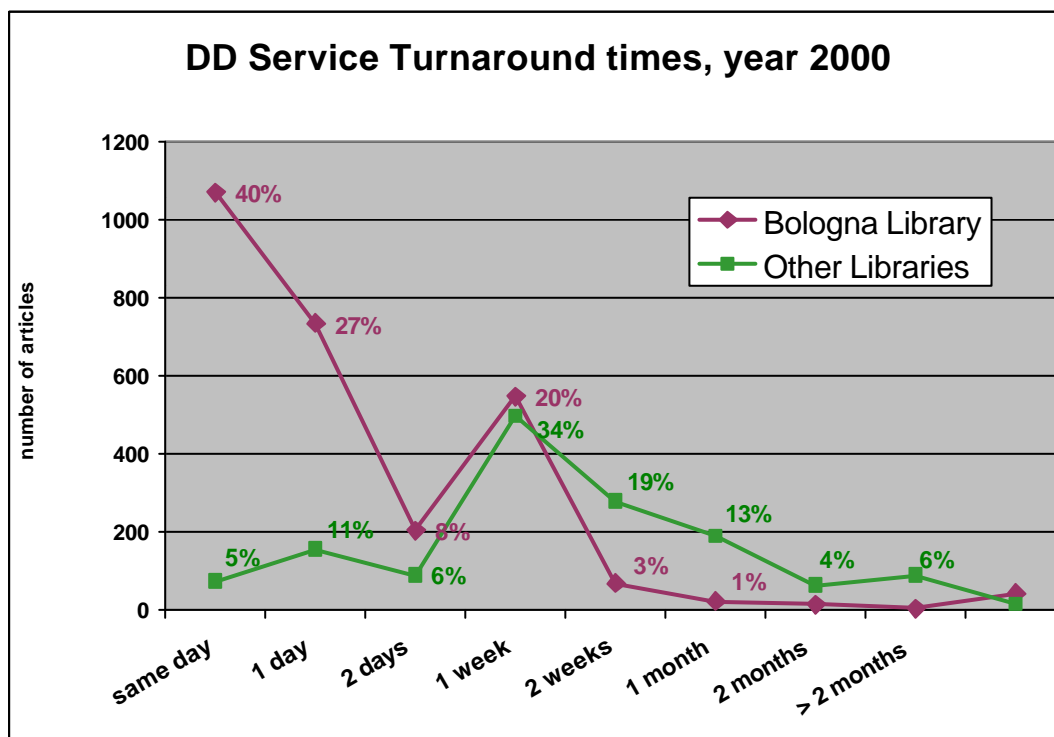


Figure 8