

**Document Delivery in the Internet
from Distributed Document Stocks
Exemplified by the JASON Project
in North-Rhine-Westphalia**

Diplom-Informatiker Rolf Rasche

**ImageWare Components GmbH, Bonn
Berliner Freiheit 36
D-53111 Bonn
Telephone (49) 228 – 96 98 5-0
Telefax (49) 228 – 96 98 5-84
E-Mail ImageWare@ImageWare.de
www ImageWare.de**

A	BACKGROUND AND PROBLEM DESCRIPTION.....	3
B	THE IDEA	5
C	THE SOLUTION	6
C.1	Replacement of the Identification Service	6
C.2	The Electronic Order	7
	Structure	7
	Process.....	7
C.3	Delivery	7
	C.3.1 Conventional Delivery.....	7
	C.3.2 Electronic Delivery	7
	Infrastructure.....	7
	Process.....	8
C.4	Facts about Jason	9
C.5	Advantages	9
C.6	Looking Ahead	10
C.7	Why Bookeye.....	11
D	FURTHER USES OF THE APPLICATION.....	12
	APPENDIX - ILLUSTRATIONS.....	15

A Background and Problem Description

"Information is the fourth productivity factor"

Information is not only essential for production, but also for science, research and training. The price of information rises the more up to date it is and the smaller the target group. This is especially true in scientific circles, where division into many highly specialised areas makes the target groups very small. The circulation of scientific books and periodicals is correspondingly low, and they are therefore expensive compared to popular magazines.

Scientific libraries, with their wide spectrum and many subject areas, are especially affected. The acquisition of periodicals for the various faculties is a major cost factor in the budgets of universities and polytechnics.

Because of these high costs, combined with general cuts in public spending which also affect academic and public libraries, North-Rhine-Westphalia (NRW) developed a special initiative for periodicals in the middle of the 1980's. This gave financial support to scientific libraries that specialised in certain subjects and loaned the titles to other libraries.

To offer a wide coverage of periodicals to universities in NRW, and to avoid duplication, the specialisation of the libraries was co-ordinated by the NRW state government. The University of Cologne, for example, is now the main library for periodicals for economic and social studies.

The new concept had to be implemented in the form of a remote lending service between libraries.

What does remote lending, or Document Delivery, mean for the receiver and supplier libraries?

Graphic 1: Classical Remote Lending Process

The user fills in a request form for the desired title and hands it in to his local library. The identification service establishes whether the periodical is in this library or in another library in the same town. If it is not, the request is returned to the reader.

He now has to fill in a remote lending request form for the same title with the code of the library that houses the periodical.

This code is rechecked by the identification service and only then is the request form forwarded to the supplier library.

Further problems can arise at the supplier library

- The periodical is not there because
 - the periodical is at the bookbinders. At the end of each year all the issues for that year are bound together. Orders that arrive during this time cannot be fulfilled.
 - The requested issue is already on loan.
- The article has been torn out (unhappily, this does happen)
- The wrong source was given, the article is not where it should have been

If the article is found a copy is made and sent back to the requesting library. If the reader is lucky he will receive the requested article in about 14 days, but it could take months.

The disadvantages of this process are obvious

- long delivery times
- missing issues cause even longer delays
- cumbersome process, and therefore not user friendly
- high personnel costs
- no error detection by the user, as mistakes on the request form are only noticed by the identification service
- no cost transparency between requester and supplier libraries

Costs: a rough estimate is about DM 100 (50) (GBP 30) per transaction

B The Idea

Based on practical experience and the increasing amount of IT equipment in libraries, the idea of supporting the periodical specialisation initiative electronically and to extend it with electronic document delivery was born. The basic components of this concept are:

- Replacement of the identification service, i.e. entry of the code in the request form directly by the user.
- Electronic request to the supplier library by e-mail.
- Extension of the document delivery with an electronic path
Document Delivery by e-mail

The conditions prevailing at the beginning of 1994 in the universities of the state of NRW supported these concepts:

- 80% of the libraries have a network that can be accessed by users.
- 80% of the libraries have access to the Internet and e-mail.
- All periodical titles are registered in the periodicals database (ZDB) of the German Libraries Institute (DBI) in Berlin.

The Jason project for electronic remote lending was started on the basis of this infrastructure.

With similar problems in mind, the Subito project was initiated by other states with support from the federal government.

C The Solution

C.1 Replacement of the Identification Service

Infrastructure

The periodicals stocks in NRW are regularly extracted from the ZDB and processed at Bielefeld University. The information (about 300MB) is then written to CD-ROM, copied, and made available to the associated libraries. When this has been read at the local library, the information is available to readers on the network for filling in request forms.

Graphic 2: Infrastructure ZDB / Library

The reader can use the local database to find the identification of the required periodical from the title, and from that see

- in which location and which library the periodical is stored and
- whether it is available.

After that, the request path is decided, i.e. the order in which the requests are processed. It may be, that in spite of correct identification/location, the title is not in the first library (for reasons see above, 5% of cases)

A maximum of 3 entries is possible in the path, with the following priority:

1. Specialised library
2. Library that delivers voluntarily
3. Library that delivers voluntarily

Additionally, the means of delivery must be entered, and this influences the price. The following are differentiated:

- conventional
 - post
 - telefax
- electronic
 - e-mail direct to the reader
 - e-mail to the local library and collection by the reader from there on a user account, a floppy disc, or as a paper printout

Finally, for calculating the charges, Jason uses a transaction number. Each transaction number has a value of DM 3.00 (GBP1.00). The user has to pay the charges for the appropriate quantity of transaction numbers in advance.

With the entry of the transaction numbers the order is complete and can be closed by the user. The result is an electronic order form stored on the server.

C.2 The Electronic Order

Structure

Graphic 3: The Electronic Order

In every library there is a dedicated PC workstation called MEDEA. The MEDEA workstation performs the following tasks among others:

- creation of electronic orders from readers' requests
- creation of e-mails, e.g. for ordering
- printing of orders for bundling in the supplier library
- sending and receiving orders
- forwarding of orders according to the defined path
- follow-up of orders

The MEDEA workstation is therefore the node that links the requester and supplier libraries with each other.

Process

The MEDEA workstation receives incoming e-mail orders and checks that they are syntactically correct. Then MEDEA generates a unique order number. Using the path, an e-mail is sent to the first library, hereafter referred to as the supplier library.

C.3 Delivery

The supplier library also has a MEDEA workstation. Incoming e-mails from various local libraries are collected there and the order forms printed. As well as the title and reader data, the order form contains the order number printed as a barcode. Depending on the type of delivery, further processing is done either conventionally with a photocopier or electronically (digitally) with a scanner.

C.3.1 Conventional Delivery

For conventional delivery the article is copied by an operator. This is delivered by post together with the order form.

For fax delivery the MEDEA workstation prints an additional fax cover sheet, and this is used to fax the copied pages to the requester.

C.3.2 Electronic Delivery

Infrastructure

Apart from the MEDEA workstation there is a Bookeye station in every supplier library. This consists of a PC (Windows 3.11, Windows 95 or Windows NT) with a Bookeye overhead scanner and the Bookeye Capturing System (BCS) as the user interface. The Bookeye Capturing System has an interface that processes the data as required by Jason and MEDEA, and transfers data to MEDEA.

Process

Graphic 4: Electronic Delivery

The operator first scans the order form with the Bookeye scanner, so that the order number is recognised from the barcode. The required articles are then scanned as a batch. The scans are automatically allocated to the order using the barcode. The set-up and starting of the next scan is done at the Bookeye, not on the PC. This means for example:

- selection of left or right page
- selection of left and right page in one scan with page separation (one scan produces two TIFFs)
- format selection A5, A4, A3,A2 (Autoformat)
- brightness
- start next scan either on the panel or with a footswitch
- termination of the scan job for the current order

The order is stored on the PC and the next order can be scanned.

Delivery of the article

The MEDEA workstation collects the scanned jobs with their indexing data (order number) at regular intervals from the Bookeye station. MEDEA matches the jobs to the pending orders using the order number. An e-mail with the scanned order is created and sent to the e-mail address of the local library. MEDEA divides larger orders into smaller packets that are reassembled by the MEDEA workstation in the local library. Reception and further processing of the incoming e-mails are done by the MEDEA station in the local library. Depending on the delivery method, the data are then

- sent as e-mails in the local network
- stored under the account of the user, or
- printed and placed in the user's post tray.

Graphic 5: Overview of the JASON System

C.4 Facts about Jason

An overview

- started in 1995
- delivery within 48 hours
- requests per specialised library > 100 per day
- unified charging model for the users

Means of delivery	University/polytechnic (NRW)	Other users
Conventional <ul style="list-style-type: none"> • Post • Fax 	DM 6,-	DM 12,-
Electronic E-mail Collection/ local library (Floppy disk, User-Account, Paper)	DM 3,-	DM 6,-

- Distribution of charges
 - 20% local library
 - 80% supplier library

C.5 Advantages

Advantages of Jason over conventional remote lending:

- delivery within 48 hours
- transparent charge structure
- automatic processing of multiple holder sites using the path data
- fast delivery, even if the document is not found in the first library
- no billing, but payment in advance using transaction numbers.
- better distribution of costs and charges between local and supplier libraries

C.6 Looking Ahead

Jason has proven itself in practice. Exactly for this reason, the usage should be extended and simplified further for the user. The following extensions are planned

- Web interface for the ZDB or for the ZDB extract for NRW.
Advantage: less effort needed for user training, more up to date information, no distribution of the data to associated libraries on CD. → This has now been implemented.
- Instead of delivering the article by e-mail, send an e-mail containing only the URL of the scanned article on a web server.
Advantage: lower data communications costs for the libraries. No need to break large orders down into packets. Easier transfer to the user.
- Conversion from TIFF to GIF .
Advantage: the user can access the information directly with his browser.
- Billing / payment by credit card.
Advantage: more customers outside the universities can be gained.
- Extension of the remote lending beyond the periodical stocks.
Advantage: books from which only a few pages are needed can be included.

C.7 Why Bookeye

At first, flatbed scanners were used, resulting in the following problems:

- many employees who had previously photocopied periodicals for the classical remote lending system had difficulty using the PC, Windows, a mouse and the scanner. This was particularly true of the constant changing between scanner and PC.
- Scanning bound originals with a flatbed scanner is cumbersome and time consuming.
- Complete processing of a scan job at the scanner without operating the PC was not possible because the scanners have no panel.
- There is no selection of left page / right page etc. on a flatbed scanner.

To overcome these problems, overhead scanners were investigated. Of the three systems available on the market, Bookeye was chosen, among others for the following reasons:

- Overhead principle „scan, what you see“
- Complete processing of a scan job at the scanner - no operating at the PC
- Format selection (A5, A4, A3, A2) at the scanner
- Support for special needs of bound originals, such as left/right page, that can be selected on the Bookeye panel
- Scanning of left and right page in one scan with separation into 2 images (left/right page) by the Bookeye
- Throughput and speed in practice (cycle time)
- Robustness
- Upgrade to A2 scanner, i.e. existing A3 machines can be upgraded
- Price
- Open interfaces
- Interfaces to other delivery systems such as Subito/Ariel
- The optional software Bookeye Capturing System has interfaces that permit direct links to Jason and Subito without further adaptation or programming
- Bookeye and the Bookeye Capturing System run on all relevant system platforms in use at universities (Windows 3.1..., Windows 95 and Windows NT)

D Further Uses of the Application

Document delivery is also a typical problem in industry. Let us detach ourselves from the term "library", and replace it by "archive", "filing cabinet", or "central documentation". This automatically brings us to industry, commerce, and administration. Here are a few examples:

- **Mechanical / Plant Engineering.**
The documentation of industrial plant, especially the current version, is usually kept centrally. Copies of the documentation are often kept on the customer site, but these are not always up to date or easily accessible when maintenance is called for. For maintenance or in emergencies the documentation must be obtained from the head office.
- **Banking / Insurance**
To assess credit risks, e.g. for real estate deals, comprehensive documentation is required. This normally comes in highly variable formats, some in the normal DIN A4, but also balance sheets with bound A3 annexes, and plans in A3, A2 or larger sizes. These documents are usually kept in one place (head office or branch office), but are needed equally often in other locations, e.g. for review or countersigning the credit involvement, auditing, etc. The classical copying and faxing then has to be performed.
- **Utilities and Energy Suppliers**
The original plans are needed regularly during building and renovation of utility supplies to buildings. If unexpected problems arise on the site (e.g. wrong or outdated plans) corrected plans and additional files must be requested from the central office. These are currently collected or delivered by a courier, and the building work comes to a standstill during this time.

If we examine these cases and our own field, the situation of document delivery, or more generally transferring information to others, appears as follows:

- local delivery is by courier
- delivery further afield is by post or courier service
- where time is of the essence, the original is copied onto standard A4 format and then sent manually by fax.

Every case incurs a substantial overhead, independently of the procurement and delivery of the information, e.g.

- copying / reducing to A4
- returning the unneeded and/or copied originals to the files
- packing and labelling the documents for delivery
- checking on the return of loaned originals

Graphic 6: Current Situation: Manual Document Delivery

The decisive factor with reproduction is not the 5 - 10 Pf for the copy or the 12 Pf for the fax, but the overheads - the effort of personnel to remove and return pages, stand and wait at the copier and/or fax machine, and dispatch the documents.

Finally, there are not usually any historical data or statistics on how often requests come and what costs are incurred. Now many users, and also suppliers of DMS systems, dream of digitising their entire stocks to enable all queries to be answered at the PC and dispatched from there.

However, it is highly questionable whether this is economical for large document stocks where the access rate is measured in thousandths or less. How long would it take to digitise the stocks and provide the necessary capability?

Also the longevity of such documents (real estate, industrial plant and machinery, financing, etc.) is grossly incompatible with the short life of the data media. What is state of the art today is obsolete tomorrow (or in 5 years at the most), and not maintainable in 10 years. This forces the administrator of such document stocks to migrate the digital archive every 5-7 years to new media/jukeboxes, system software, etc. at a correspondingly high cost.

These economic considerations prohibit the conversion of large document stocks with relatively low access rates.

What should the user do? First he must decide whether the assumptions apply to his document stocks, i.e.

- size of the document stocks
- low but regular access rate
- heterogeneous stocks, such as bound and/or stapled documents, mixed formats A4, A3, A2

Furthermore the following questions must be answered

- current cost of information delivery
- how many workstations will be needed for the delivery
- how many users/customers receive the information
- what is the current (!!) infrastructure at the workplace, at the information desk, and at the recipient - does it have PCs, network and telefax or is there e-mail and Internet access?

If the above assumptions hold for your application and the necessary infrastructure (e.g. at least a PC with fax connection) is available at the workplace, find out about adding a scanner with connection to the telefax and/or e-mail for implementing a document delivery system. For the choice of a scanner the following decision matrix might be helpful.

Decision Matrix

- For standard originals, not clipped / stapled / bound and/or less than 50 requests per day, a standard flatbed scanner (format up to A3) should be sufficient.
- For bound/stapled originals, originals in files, formats larger than A3, and/or more than 100 requests per day, you should consider overhead scanners. The personnel effort for

pre and post-processing is much lower for overhead scanners than for flatbed scanners.

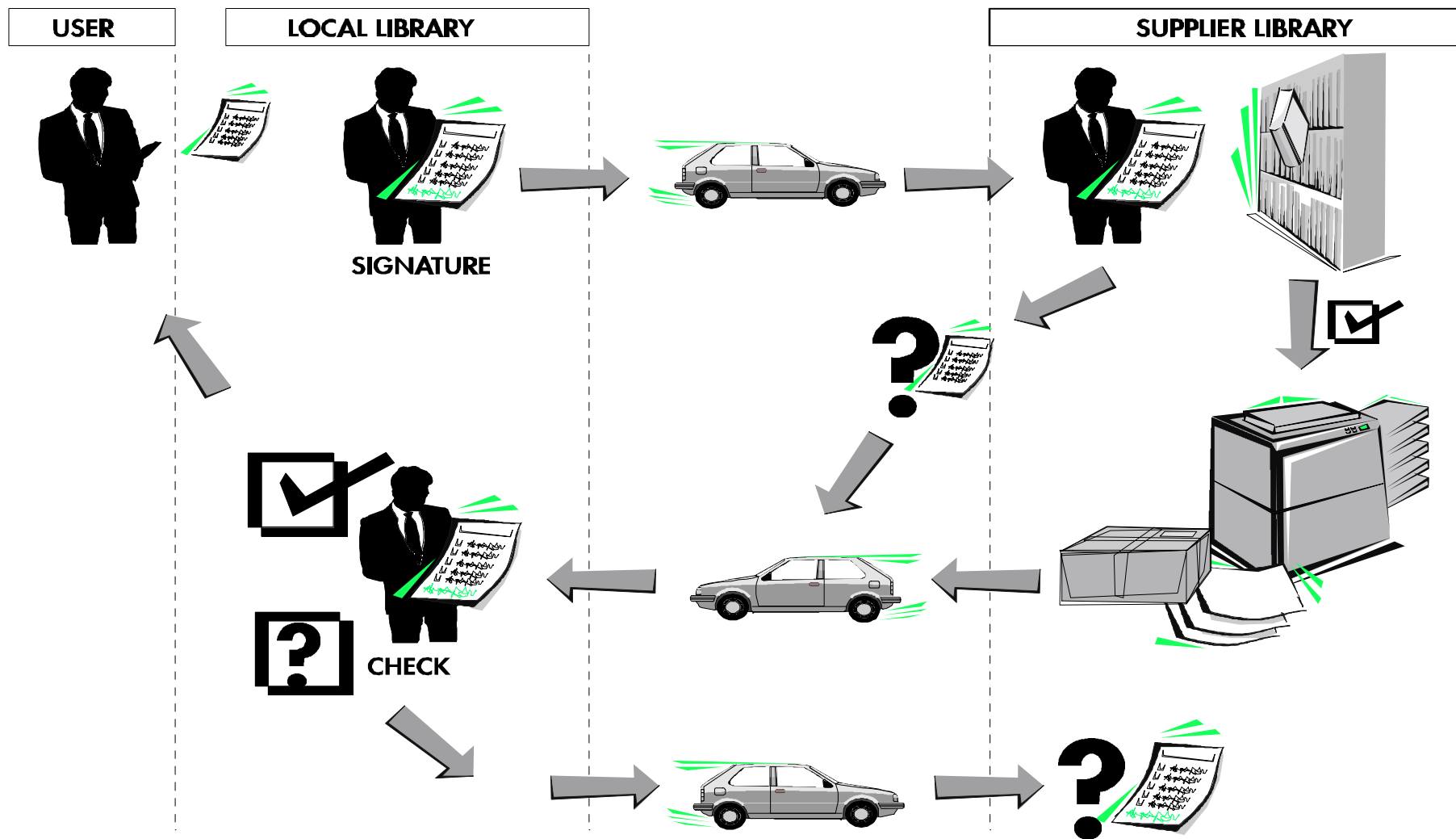
Graphic 7: Goal: Electronic Document Delivery

Examine your document delivery application for cost, speed and overheads. Compare this with a scanner solution. The infrastructure is there (PC, fax, network, e-mail, mobile phone). The costs are manageable and the risk of failure or time and budget overspend is very limited compared to large projects. The implementation happens fast and the interval for assessing the success is correspondingly short.

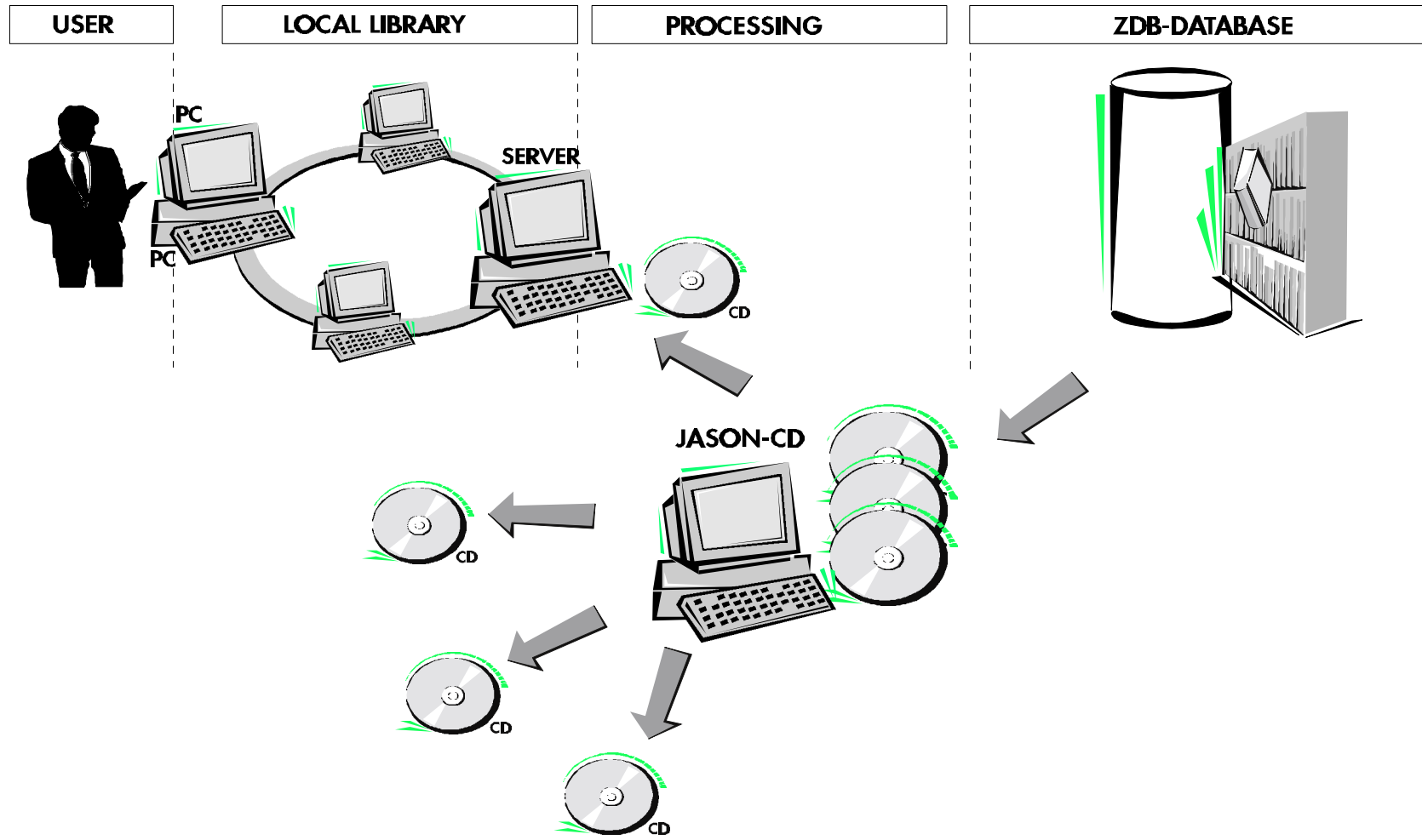
Appendix - Illustrations

- G 1. Classical Remote Lending Process
- G 2. Infrastructure ZDB / Library
- G 3. The Electronic Order
- G 4. Electronic Delivery
- G 5. Overview of the JASON System
- G 6. Current Situation: Manual Document Delivery
- G7. Goal: Electronic Document Delivery

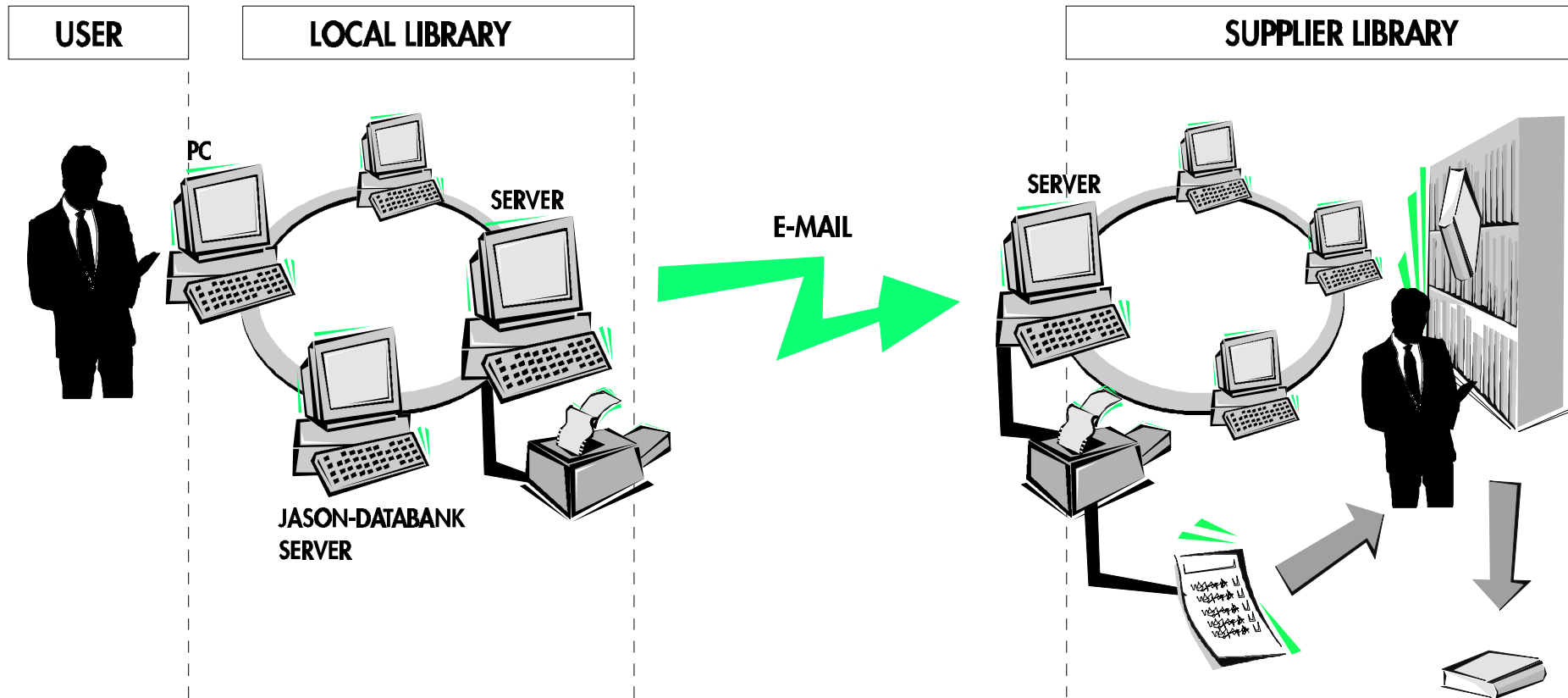
G1. Classical Remote Lending Process



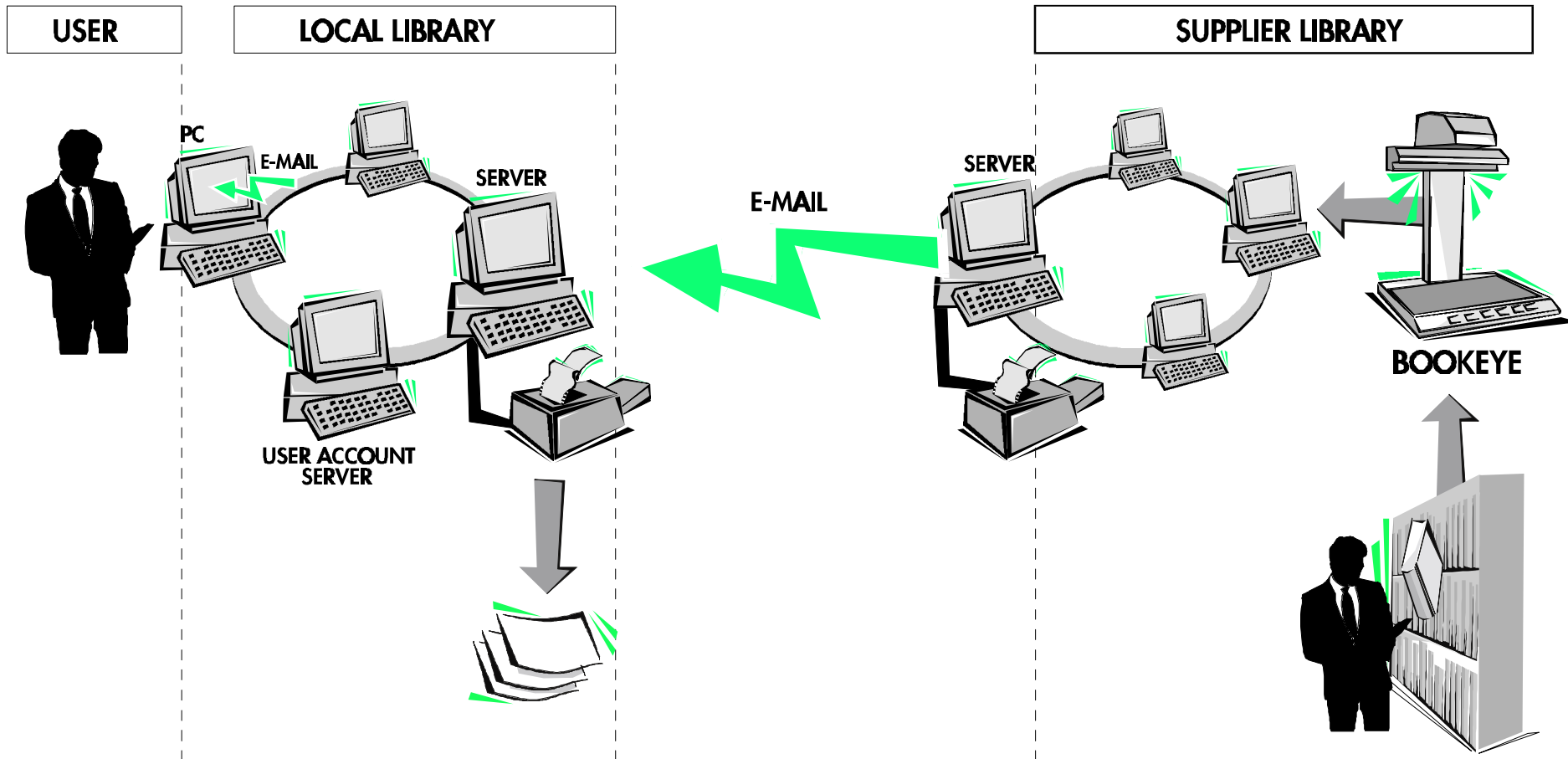
G2. Infrastructure ZDB / Library



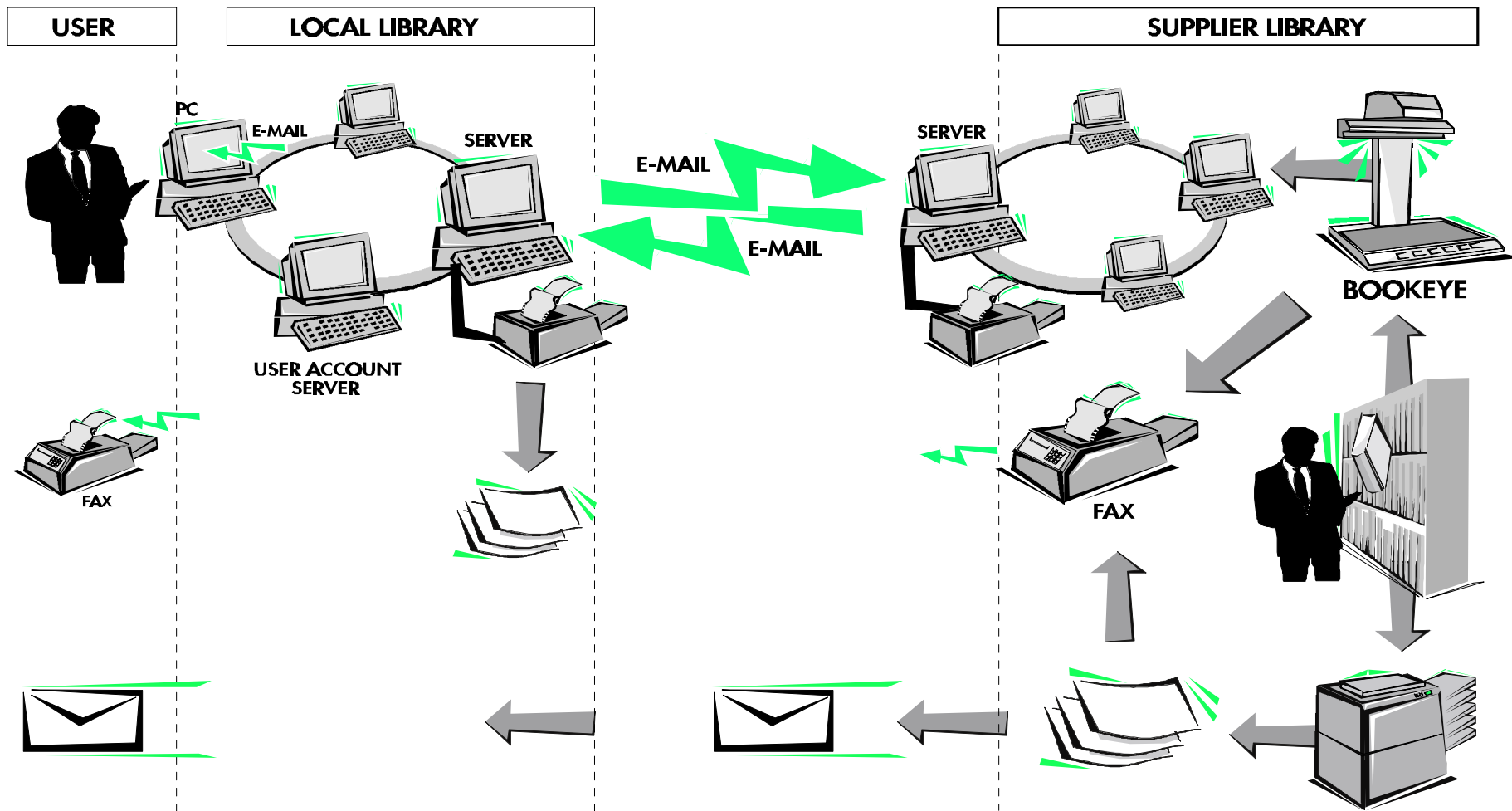
G3. The Electronic Order



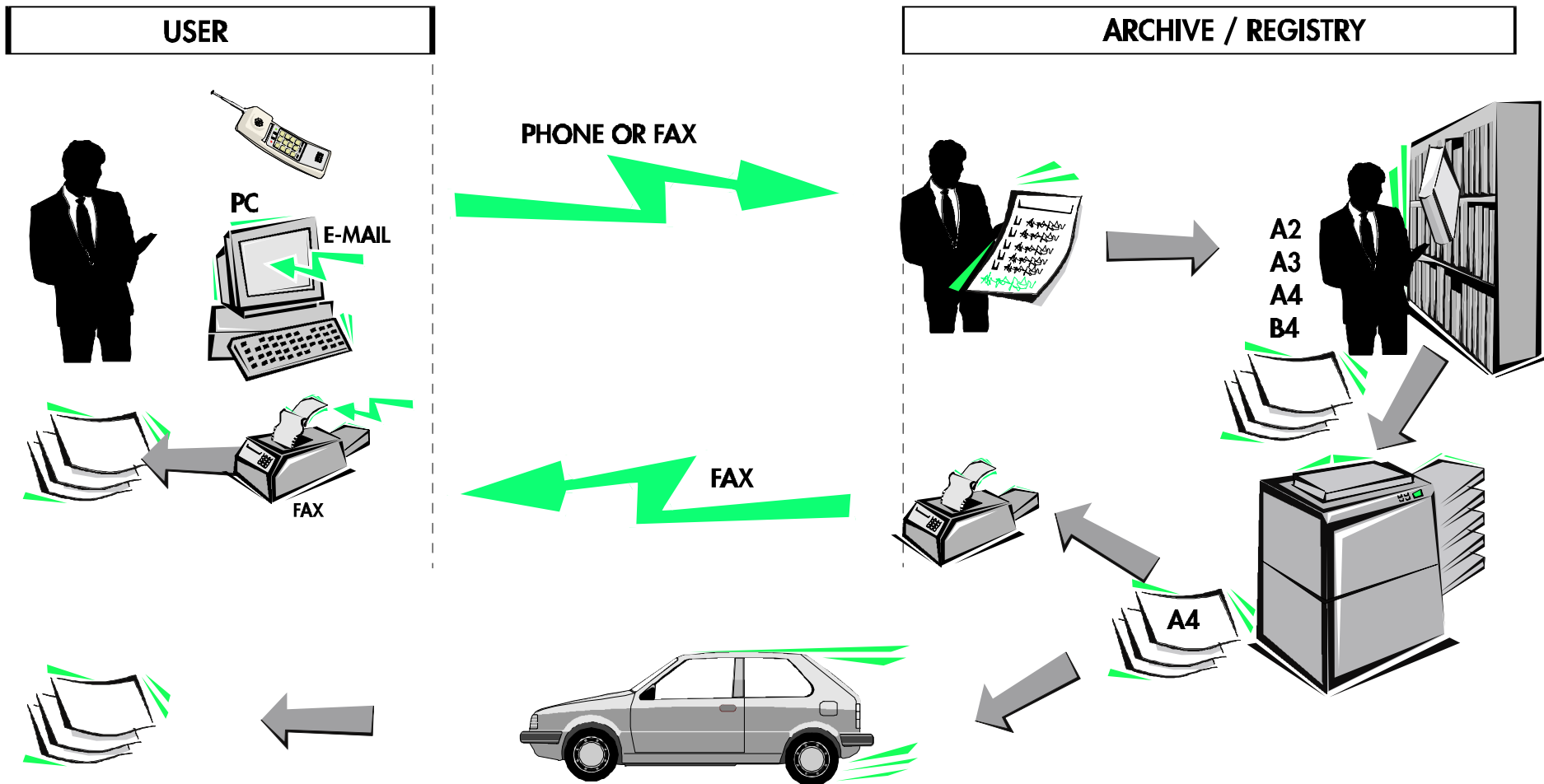
G4. Electronic Delivery



G5. Overview of the Jason System



G6. Current Situation: Manual Document Delivery



G7. Goal: Electronic Document Delivery

