



NRC Publications Archive Archives des publications du CNRC

Implementing IntelliDoc: Toward a virtual library environment for CISTI Mikoski, K.

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. /
La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.

NRC Publications Record / Notice d'Archives des publications de CNRC:
<https://nrc-publications.canada.ca/eng/view/object/?id=c3d4cb9e-debd-4353-9f81-22711a521c1c>
<https://publications-cnrc.canada.ca/fra/voir/objet/?id=c3d4cb9e-debd-4353-9f81-22711a521c1c>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at
<https://nrc-publications.canada.ca/eng/copyright>
READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site
<https://publications-cnrc.canada.ca/fra/droits>
LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at
PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.



Implementing IntelliDoc: Toward a Virtual Library Environment for CISTI

A Paper Presented at the
Interlending & Document Supply
4th International Conference
Calgary, Alberta, Canada
June 11 - 14, 1995

by
Kathryn J. Mikoski

Today I have been asked to participate in the discussion of some very technical aspects of Document Delivery - specifically that of imaging workstations. I must warn you that I am not "technically" expert. I approach this subject with a manager's perspective which I hope will be of value to some of you. I am fortunate to have been given the opportunity of having with me, two technical experts in Clare MacKeigan, IntelliDoc Project Manager and Andre Meragh from Network Support Inc. I will turn to these people to assist in answering any technical questions my presentation may invite.

I also invite you, if you have not already done so, to have a look at the workstation we have set up here as it will help you to visualize better what I am going to describe.

What is IntelliDoc and Why has it been Developed

#1 CISTI Building

CISTI houses Canada's major scientific, technical & medical collection. As the library for the National Research Council (Canada's largest scientific research establishment) and North America's largest document supply source, CISTI's collection is already heavily used. However, this enormous investment in publications must be shared even **more** if we are to justify the continued high cost of maintaining the collection.

To increase usage we must be **flexible** - accepting orders in many formats, from anyone, anywhere. We must deliver excellent **quality** - drawings, photographs and colour are important features of scientific literature. We must provide a **fast** service and we must **price right**.

#2 Document Supply Process at CISTI

From ordering options to sending of invoices and producing management information, we have to make sure everything is as automated and as efficient as possible.

Two years ago we had 38 staff dedicated to the processes of order receipt, call numbering, retrieval, copying, packaging, billing and statistics keeping. These staff

processed approximately 1200 orders per day and achieved a 3 day turnaround on 90% of the orders supplied from the CISTI collection.

#3 The Challenge

Our five year objective and our challenge is to more than double our volume of daily orders to 3,000; to achieve 24 hour processing of supplied documents (50% within the same day). **And**, because we must maintain or lower our costs, we must achieve this with a maximum of the same operational budget and same number of staff.

#4 IntelliDoc

This overhead is a model of "IntelliDoc". At its core are the client and the document order. Beginning at 12 o'clock and moving clockwise around the core are the processes which go into fulfilling a document order from order receipt to call numbering to retrieval, scanning, delivery, billing, etc. Encircling these processes and binding them together are the standards and protocols that allow things to function in an integrated way.

Some of the pieces needed to achieve our objectives were already in place. Their further integration and enhancement, along with the development of new pieces is **"IntelliDoc"**

We had (represented in white on the model):

- electronic ordering facilities (2 systems)
- auto call numbering

We developed (in yellow):

- a client registration system which is the heart of IntelliDoc
- expanded electronic ordering facilities (now 8 systems plus order transcription for telephone, mail, fax orders)
- a copyright tracking system
- automated billing
- a contract with an automated courier company for paper delivery
- a scanning workstation network to replace copiers
- a transaction database and workflow management software for statistics and management information

We must still (in red):

- integrate electronic publications
- enhance & develop end-user interfaces for ordering and document receipt
- automate circulation control

- add bibliographic search services, messaging and ILL protocols
- integrate table of contents services

My focus this afternoon is on the development of the scanning workstation component or the "Electronic Document Delivery System Project" which began with the issuing of an RFP in the spring of 1993. I will not spend a great deal of time on the background but rather refer you to a couple of published articles on the subject:

6 Background

Feliciter Vol. 41, # 2 February 1995 "IntelliDoc transforms Document Delivery" by Clare MacKeigan and Elizabeth Katz

and

Interlending & Document Supply Vol. 22, #1 1994 "Electronic Document Delivery - Towards the Virtual Library" by Clare MacKeigan and Michael Brandreth

Reprint copies of the Feliciter article are available at the back of the room.

7 Workstation Objectives

Our objectives as extracted from the RFP document were to replace photocopy technology with scanning, to replace traditional mail delivery of paper copies with electronic transmission options, and, as more and more documents were made available in electronic form, to replace scanning of paper originals with the importation and conversion of electronic data.

We wanted to accomplish this transition with:

- no degradation in productivity (and for CISTI that meant a total time of 5 minutes to retrieve, scan, compress, view and reshelve a 10 page document)
- provision of a more visually attractive product
- reduced response times on document orders

We had experience with Ariel - we had purchased two stations for servicing one of our Branch libraries in western Canada. We recognized both its potential and its limitations in our environment.

- we needed software and hardware which would be much faster and more robust
- we needed a workstation which could serve not only Ariel clients but **all** our clients no matter how they wished to receive their documents

The results we have achieved took 18 months of development. We attempted full

operation as early as October, 1994 but pulled back to problem solve and fine tune our processes. I prefer to say we became fully operational on April 1, 1995.

Although what I am about to describe is customized to CISTI needs, it is recognized that the integration of hardware and software to manage workflow, document imaging and document storage and retrieval is "generic". The component parts of our system can be changed, reorganized and applied to different environments but the principles remain the same. We have used commercial off-the-shelf products whenever possible. Software suppliers have based their products on open system principles facilitating integration.

8 EDDS components

The IntelliDoc family of integrated systems began life in an IBM mainframe environment. Our order receipt and automatic call-numbering programmes at the front end, and our copyright tracking, client registration and invoicing systems at the back end, continue to operate on the mainframe. The functions of:

- order database creation
- workflow definitions
- order distribution & printing
- scanning
- exception handling
- delivery determination
- document printing or electronic delivery

have been developed in a Novell LAN environment.

9 Hardware Components

The LAN connects 16 workstations on five stack floors with a bank of servers and two high speed printers in what I term our "control" centre.

On each stack floor will be a variable number of workstations depending on the nature of the collection and volume of demand on that floor. A workstation is comprised of:

- 486 PC
- 20 inch touch screen monitor
- customized Fujitsu 3096e scanner
- stationary barcode reader
- foot pedal
- mouse
- HP Laser jet 4 printer

In the "control centre we have:

- a Tricord superserver to manage the database and images

- a SCO Unix Server for communications such as Internet
- a Pentium fax server
- 4 486's and 1 386 to handle such functions as:
 - application start up
 - order distribution
 - document printing
 - determining delivery process
 - remote printing
 - database purging

Two high speed printers producing 14 pages per minute handle production of paper copies going out to clients preferring traditional delivery methods.

12 Software Components

A number of software packages have been integrated to provide us with the functionality we require.

As an example:

DIMS	operates the scanning process
RIMS	processes orders and documents
Flowman	directs the workflow

What Does it Actually Do?

Now you have heard the reason why, the objectives for the development, the component parts - but what have we actually accomplished with our scanning workstation?

13 IntelliDoc System Workflow

Orders are received into the system either directly from electronic formatted sources or after being keyed into the system (in the case of mail, fax and telephone requests). Using the account # supplied by the client as part of the order, client information such as:

- method of delivery
- method of billing
- rate code, etc.

from the Client Registration System is linked to the order.

Orders are automatically matched with the document shelf location by the Auto-call numbering system.

Orders are printed in shelf location order on the appropriate stack floor. Each order is printed with a unique bar-code.

Staff retrieve items from the shelves and return to the workstation.

The order is entered by reading the bar-code and pages are scanned and electronic images created. Links to the client information allow the system to determine the delivery method, convert the image to the appropriate format and direct the delivery whether it be Ariel, ftp, fax, local or remote print, etc.

If a document cannot be retrieved and scanned, a whole series of exception handling options are available to the operator for redirecting the request for further searching, circulation control, etc.

Completed orders are tracked to capture copyright payment information, billing information and other statistics of interest to management.

The image file is created in TIFFB format. Group IV compression standard is used and FTP for file transfer.

14 Results

As I speak,
1,400 orders per day are flowing through our system.

We are supplying
1,200 documents per day - 1,100 scanned.

We are achieving a productivity of
5 minutes for retrieval, scanning, compression, viewing and reshelving of each document.(the scanning/compression process is averaging 6 seconds per page)

We are producing a variety of outputs:
Laser prints at the control centre and remotely to other parts of NRC
connected to our LAN
Ariel delivery
ftp
Group III fax

We are able to control and redirect exceptional orders requiring special handling.

We are able to produce useful management information. For example:

15 Turnaround Times

Previously we monitored our processing time by means of a sample study 3 or four times a year. This involved several person/days of counting and analyzing. Now the

system can produce for me at any time and for any span of time an analysis of processed orders from time of receipt to time of delivery.

This graph was produced for me late in May. If you will recall, in the beginning I mentioned that two years ago we were providing 3 day turnaround on 90% of our documents supplied. This graph reports 24 hour turnaround on 86% of documents supplied and same day processing of 24%. Our objective is 100% and 50% respectively so we are not there yet!

16 Scanning Activity

This next chart gives me some clue as to how we can improve the turnaround times but also where I can ensure no wastage on capital investment. It shows me that there are significant periods of inactivity or low activity in the scanning process. If I shifted staff schedules and order printing times, I would be able to utilize the scanners for much higher volume per day without increasing the number of workstations and I could also get more material out the same day.

17 Order Tracking

Similarly, this chart shows the tracking of one order # SS-1993415. It shows me that this order DID NOT get out the day it was received and it might have if some changes could be made.

- the order entered the system at 11:02
- it took 30 minutes before it was printed on a stack floor
- nearly 4 hours went by before it was scanned
- another 2 hours transpired before the document was printed in the control centre for courier delivery (too late to go out that night)

There are a number of things that could be done to get that document out the same day:

- by alternating staff lunch schedules, a delay of 4 hours between printing of order and retrieval of document could be avoided
- by relieving congestion and capacity problems on the printer and servers, reduce the time lag between scan and document print
- arranging for later courier pick-up time

Another potentially valuable management tool offered by this system is the tracking of employee productivity. We have not implemented this analysis as yet but it could offer supervisors an effective way of detecting training and performance problems as well as bottlenecks and workload inequities.

Where to We Go from Here?

The objective we have set for ourselves has been a massive undertaking, proving, as many major technical projects, to be much more complicated than could be reflected in the original specifications. We are not yet where we want to be. Even as I speak we are grappling with new problems preventing us from delivering our products as consistently as we would like. But I am proud of what we have accomplished so far and confident that the difficulties we are encountering will be solved. I would like to share with you some of the experiences we have had.

18 Issues

Scanner Technology

One of our biggest frustrations was in fact a fairly basic - shall we say "mechanical" one. Scanners have not been developed with the library environment in mind. We had to modify off-the-shelf equipment by removing the cover and by removing the document feeder to provide ourselves with a "flat-bed" with no obstructions to handling heavy bound volumes. Foot pedals were added to keep operator hands free to handle documents.

Text/Image

Combinations of text and images on one page proved to be another challenge - do we dither or don't we? Perhaps we do both and provide the client with a good image followed by crisp text (but then what do we do about the page count upon which copyright and billing calculations depend). We finally resolved that problem with the help of an IPC board but not without weeks of investigation and, of course additional expense.

Capacity

I use the term capacity for lack of a better all-encompassing word. We are currently living on the edge of our ability to schedule all of the functions that have to occur throughout the day, in the required frequency. In fact, as I left for the conference, we WERE NOT meeting the demands on the system. And since we no longer have the photocopiers, we have no choice but to make it work. We have and will add power to our servers, we have tweaked the schedules and we have re-organized the data. We are hopeful that the installation of multi tasking NT software will be another part of the solution.

Management of Change

On the human rather than technical side, we have had the usual challenges of introducing change to a mixed audience. The implementation of the scanning workstation created a fundamental change in the workflow. Some staff are stimulated and excited by it, others are resistant, afraid and sure it won't work. Because we were so often breaking new ground, we were frequently in the position of having to "experiment" and ask staff to test and provide opinions. Some staff, we discovered, do not react well to this environment. They would prefer that all testing be done in a "lab" and that they be handed a finished product.

Copyright

On another level entirely, scanning technology has generated questions about copyright, and the legality of what we are doing. I do not pretend to be an expert in this area and even the experts seem to be divided on the matter. To further complicate the issue, Canadian legislation is in a state of "flux". Suffice it to say that CISTI, as a responsible government organization, and as a publisher of scientific documents in its own right, is loathe to either break the law or alienate the publishing community which it views as a partner in the world of information dissemination.

We are not:

- using OCR technology to create ASCII files
- storing scanned images beyond 48 hours to ensure satisfactory transmission

We are:

- paying copyright fees under agreement with CanCopy, a Canadian collective
- entering into direct agreements with publishers
- more than willing to explore solutions with all interested parties

19 Plans & Priorities

Having come this far, we see even a greater way to go to achieve the perfect world. Faster than we can develop a function, new requirements appear.

End-user ordering and delivery

CISTI has for many years served end-users directly rather than limiting its document delivery services to other libraries. What is changing rapidly, is the degree to which traditional ILL departments want to step out of the mainstream and have their end-users deal directly with the supplier. This requires systems which will accommodate a varied combination of mediated and unmediated ordering, delivery and billing scenarios. We are finding that every set-up is unique, end-users don't necessarily have the facilities to search and order documents or to receive electronic images. We recognize that we are going to have to put a lot of effort into building these links.

Electronic Publications

Thanks to the studies of others (INIST and National Library of Medicine in particular) and our own findings, we have concluded that we will not proceed down the road of wholesale scanning and storage for future use. This does not seem to us to be a sound business decision. Instead we view the future more in terms of providing a link to electronic publications and we are currently searching for the right kind of test case for linking our call-numbering system with automated retrieval, format conversion and electronic delivery of electronic publications whether they be in CD-ROM format or on-line.

Productivity Improvements

Of course, we must continually strive for productivity improvements and we already have a long list of enhancements we will be making to IntelliDoc to help us work more efficiently. I already mentioned NT software, we also want to link our bibliographic searching, messaging, and Table of Contents services to IntelliDoc.

Protocols

We appreciate the value of standards and protocols and support them whenever it makes good economic sense to do so. We have implemented some standards:

- TIFFB
- Group IV compression
- TCP/IP
- FTP

are standards and protocols imbedded in our systems.

Our document order data elements conform to those of the ILL protocol.

Nonetheless we recognize that there is more work to be done in this area.

Sharing

We have learned a lot over the last two years as we have developed IntelliDoc. We have moved from a strict client - consultant relationship with Network Support Inc. to the point where, as a team, we are investing in future development. It is our hope that we might be able to share the collective expertise we have built up in this area.

20 Cooperation Agreement

To that end, I am pleased to tell you that CISTI and NSI have established a

"cooperation agreement" under which we will pool our IntelliDoc "know-how" to help others develop applications for their own needs.

I am equally pleased to announce that the University of Alberta is the first organization to take advantage of this expertise to help them develop an IntelliDoc - like operation to support their programs.

This concludes my presentation on CISTI's scanning workstation. I would invite you to ask questions, to pick up literature at the back of the room and once again, to have a look at the demonstration workstation we have set up for this purpose.