Chipless ID for Everyday Documents

Daniel Lopresti
Computer Science & Engineering
Lehigh University
Bethlehem, PA 18015
dal9@lehigh.edu

George Nagy
Electrical, Computer, and Systems Engineering
Rensselaer Polytechnic Institute
Troy, NY 12180
nagy@ecse.rpi.edu

Overview

Our interest: applying chipless ID to tag information as embodied in everyday documents (as opposed to physical objects).

- Motivation
- PanaMarks and "Smart" Paper
- Perfect Copier
- Other Applications
- Related Work
- New Opportunities Using Chipless ID

Goal: generating discussion and possible joint R&D collaborations within proposed RPI Center on Chipless ID.

History

In 1991, Matsushita Electric Industrial Co. (a.k.a. Panasonic) established MITL, a new U.S.-based research lab in Princeton, NJ.

Dan Lopresti led development of novel technologies for document management at MITL from 1991-1997:

- New 2-D barcode symbology
- Notion of assigning each piece of paper a unique ID when printed
- Enabling new paradigm for office equipment, e.g., Perfect Copier

All based on premise that so-called "paperless office" is a myth.

Note: certain intellectual property described in this presentation is owned by MEI.

A Paperless Society??

"Everyone knows computers are replacing paper."

"Soon, all information, including books, magazines, and newspapers, will be produced, distributed, and 'consumed' electronically."

While statements such as above reflect popular wisdom, it seems more likely that paper and computers will continue to co-exist.

> *"... despite the deployment of 25 million personal computers in this country, there appears to be a law of paper hyper-growth: for each additional billion dollars of gross national product yearly, there are 80,000 tons of paper."*
>
> – *New York Times, May 30, 1993*

Slide circa 1995. Certain material contained herein is owned by MEI.

A Paperless Society??

"Many people talk about their concerns for the 'last mile' – for the delivery (of service) into every home. I'm concerned about the last yard. We can easily move information from one computer to another, but how do we get it from the computer to the human being in the proper format?

"Not all information is suited to electronic use. Think of the auto repair manuals that you drag under the car and drip oil on. Think of children's books, with their drool-proof pages.

"There are times when I can only conclude that we have been gripped by some strange madness.

"I have fantasies of kidnapping the entire membership of the administration's Information Infrastructure Task Force and tying them down in front of 14-inch screens with really bad flicker and forcing them to read the whole of Project Gutenberg's electronic copy of Moby Dick.

"Maybe then we'd get some concern about the last yard."

– Librarian Karen Coyle at the 1994 meeting of Computer Professionals for Social Responsibility

Slide circa 1995. Certain material contained herein is owned by MEI.

A Paperless Society??

Fast-forward to 2002 ...

Sellen and Harper assert that "paper [remains] the medium of choice for reading, even when the most high-tech technologies are to hand."

They suggest the following reasons:

- *paper allows "flexible [navigation] through documents;"*
- *paper assists "cross-referencing" of several documents at one time;*
- *paper invites annotation; and*
- *paper allows the "interweaving of reading and writing."*

– The Myth of the Paperless Office, MIT Press, 2002

A Paperless Society??

... and on to June 2004 ...

"With each new wave of technology and legislation, fresh predictions are offered on the demise of paper documents as the predominant medium for business-to-business transactions. The eSign Act and the Health Insurance Portability and Accountability Act (HIPAA), for example, were once hailed as harbingers of a new electronic era, yet users report that projections haven't panned out.

"Paper-based claims are growing at BlueCross BlueShield of Tennessee (BCBST), says Deanna Quinn, manager, creative business solutions, who asserts that HIPAA has only made matters worse.

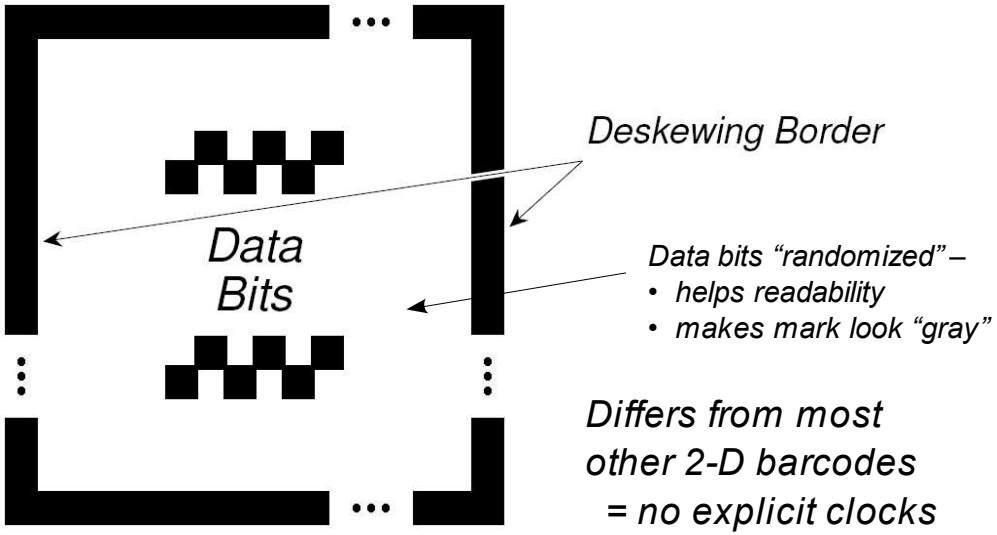
"A lot of people thought they weren't ready to comply with HIPAA standards for electronic submissions, so they went back to paper," Quinn says.

"Some cite confusion surrounding compliance mandates as one explanation for paper documents not going the way of typewriters. Others bring up signature requirements ..."

– Toward the Digital Mailroom by Ralph Gammon

"Smart" Paper

Since computers can't replace paper, paper should be made "smarter." One approach for achieving this is the PanaMark:



At MITL, we developed our own proprietary 2-D barcode symbology. However, other encoding schemes are possible.

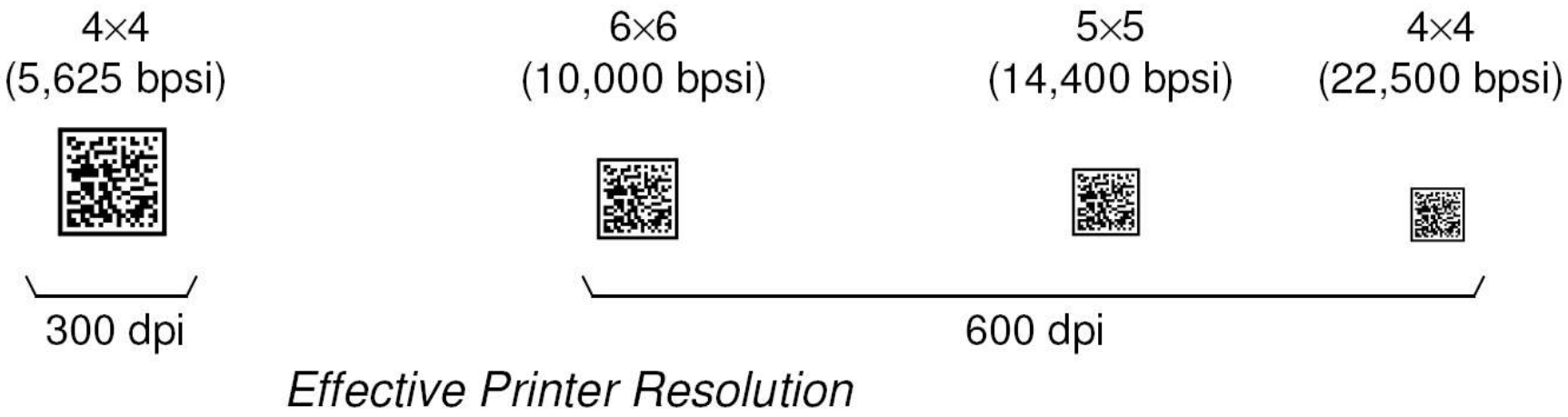
Note: PanaMark is a trademark of Matsushita Electric Industrial Co., Ltd.

Slide circa 1995. Certain material contained herein is owned by MEI.

PanaMark Overview

PanaMarks are 2-D barcodes created at the time a page is printed.

Different encoding densities are possible depending on the number of pixels used to encode each "bit" of the mark:



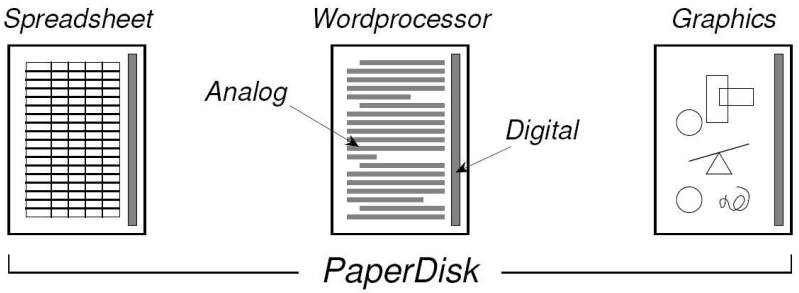
Effective Printer Resolution

Slide circa 1995. Certain material contained herein is owned by MEI.

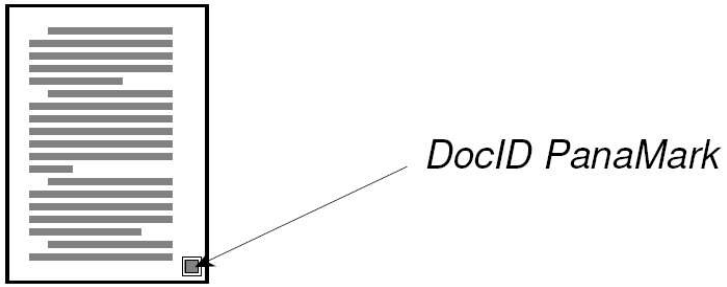
PanaMark Overview

At MITL, we identified several fundamental types of PanaMarks.

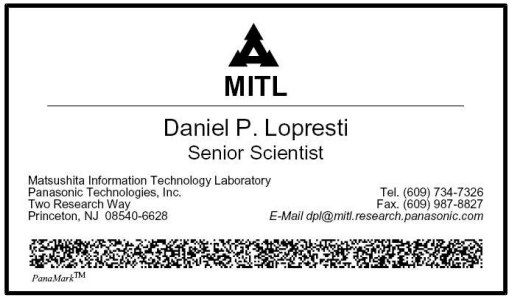
Content-encoding:



Document Identifiers (pointers):



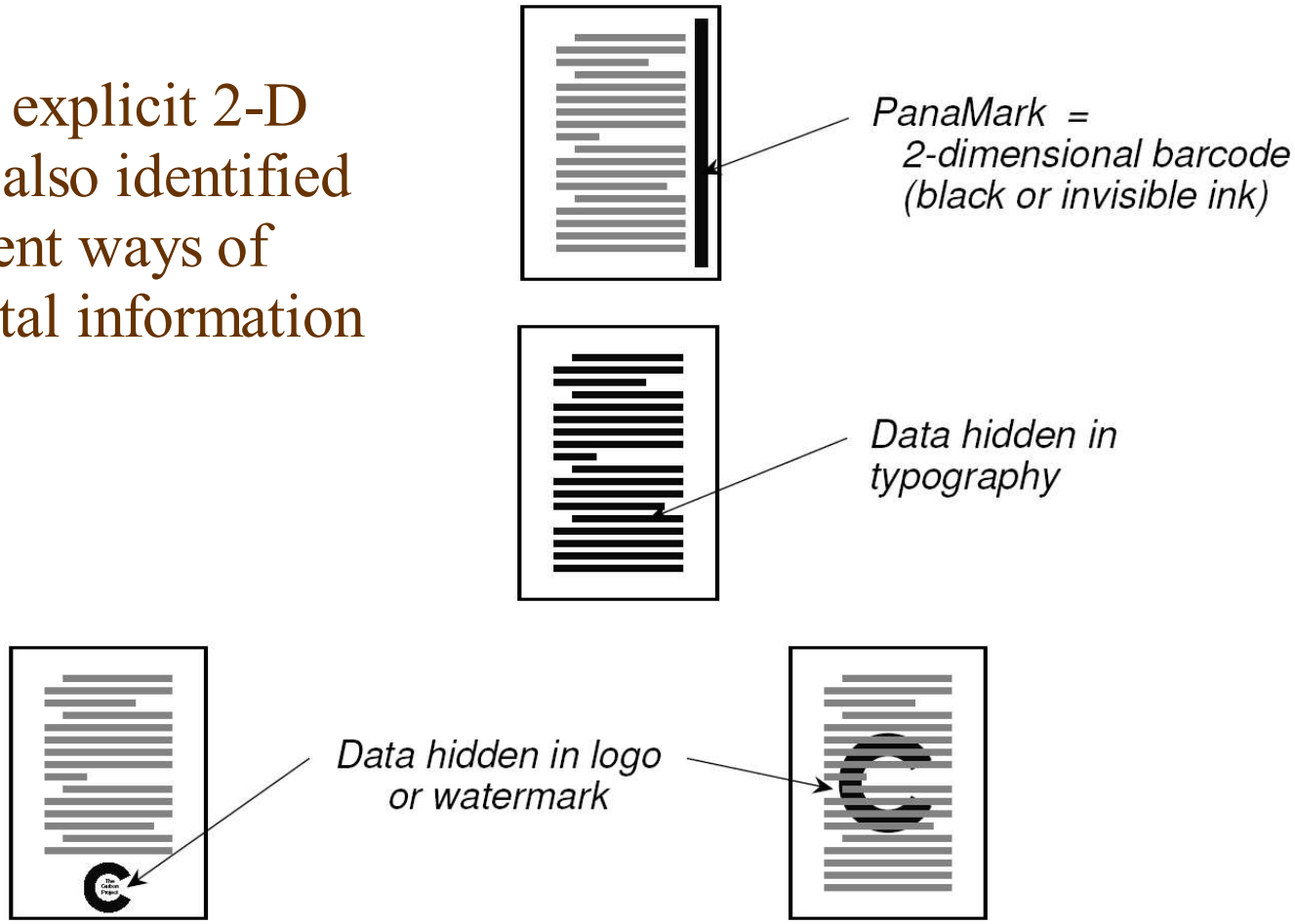
Special-purpose:



Slide circa 1995. Certain material contained herein is owned by MEI.

PanaMark Overview

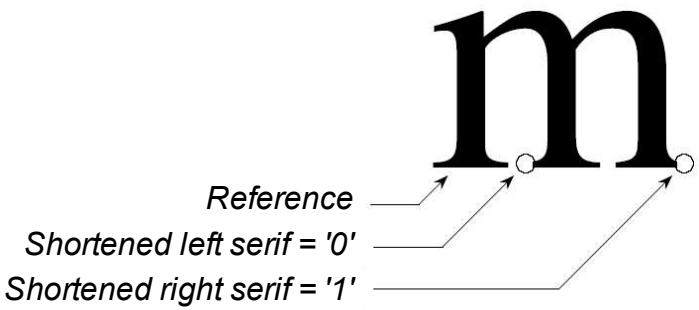
Beyond using explicit 2-D barcodes, we also identified several different ways of encoding digital information on a page:



Slide circa 1995. Certain material contained herein is owned by MEI.

Alternate Encodings

More radical encoding schemes based on typography:



Encoding Scheme	*Approximate Bits per Page*
Inter-paragraph spacing	4
Inter-line spacing	40
Inter-word spacing	600
Inter-character spacing	3,000
Word baseline	600
Character baseline	3,000
Ragged-right line ending	40
Serif variation	3,000

(Similar to concept of steganography in images.)

Slide circa 1995. Certain material contained herein is owned by MEI.

DocID PanaMarks

Document Identifier (DocID) PanaMarks encode document naming information:

- User identification (56 bits)
- Document number (22 bits)
- Page number (10 bits)

Each DocID PanaMark also includes ECC and redundancy for a total of 400 bits.

Slide circa 1995. Certain material contained herein is owned by MEI.

DocID PanaMarks

	Config. 1	*Config. 2*	*Config. 3*
Printer Resolution	300 dpi	400 dpi	600 dpi
Scanner Resolution	300 dpi	300 dpi	600 dpi
Bits / Square Inch	5,625	6,400	~20,000
DocID Size	0.1 sq. in.	0.09 sq. in.	0.03 sq. in.
	Tested Extensively		*Future*

Is 88 bits enough?

2^{56} = 72,057,594,037,927,936 different users

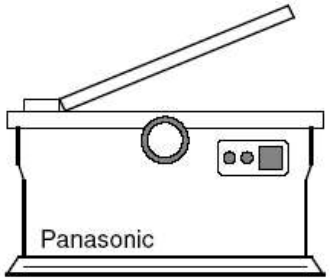
2^{22} = 4,194,304 documents per user

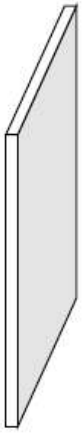
2^{10} = 1,024 pages per documents

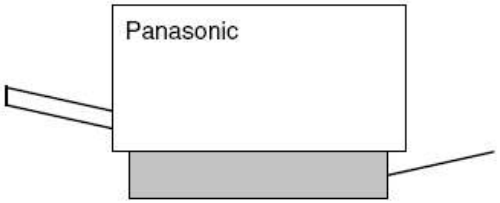
Slide circa 1995. Certain material contained herein is owned by MEI.

Old-Fashioned Copiers

Barrier separating paper-handling machines





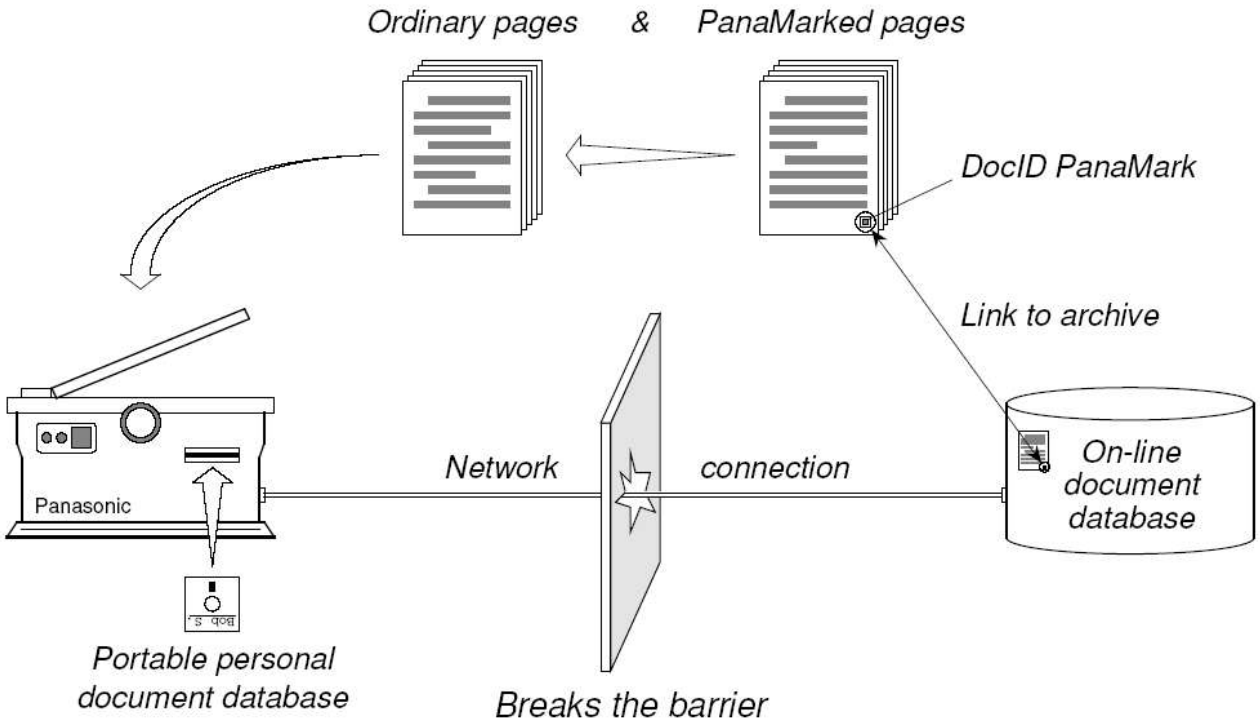


Ordinary Photocopier –
+ copies anything on paper
BUT
– only makes 2nd generation copies

Ordinary Laserprinter –
+ makes 1st generation copies
BUT
– only if you know where the file is

Slide circa 1995. Certain material contained herein is owned by MEI.

The Perfect Copier

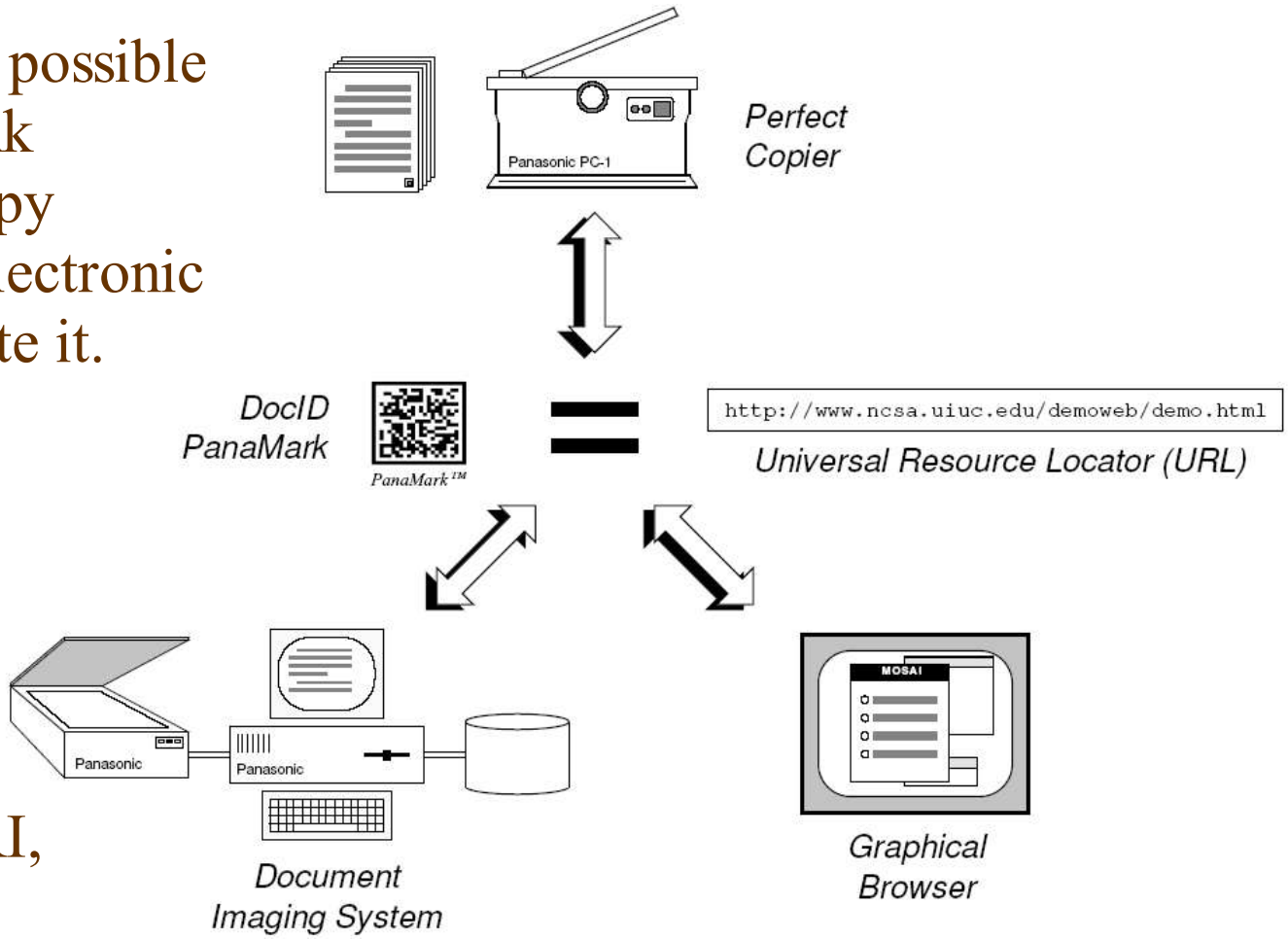


The Perfect Copier –
- looks just like an ordinary photocopier
- makes perfect copies for PanaMarked documents
- makes ordinary copies for other documents

Slide circa 1995. Certain material contained herein is owned by MEI.

DocID as a Paper "URL"

DocIDs make it possible to establish a link between hardcopy document and electronic file used to create it.

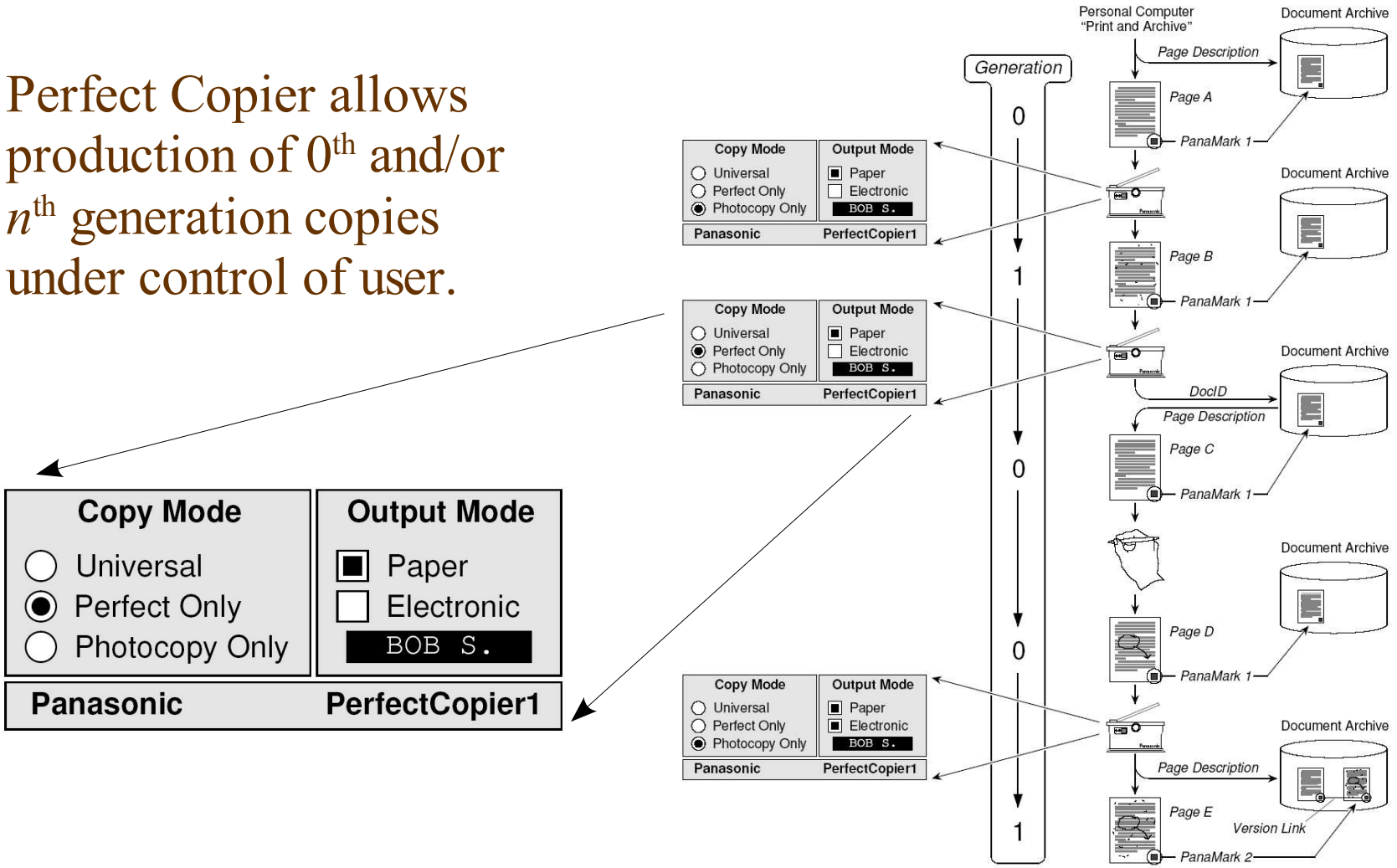


Could regard this as URL, URI, DOI, EPC, etc.

Slide circa 1995. Certain material contained herein is owned by MEI.

Example of a Document Lifecycle

Perfect Copier allows production of 0^{th} and/or n^{th} generation copies under control of user.

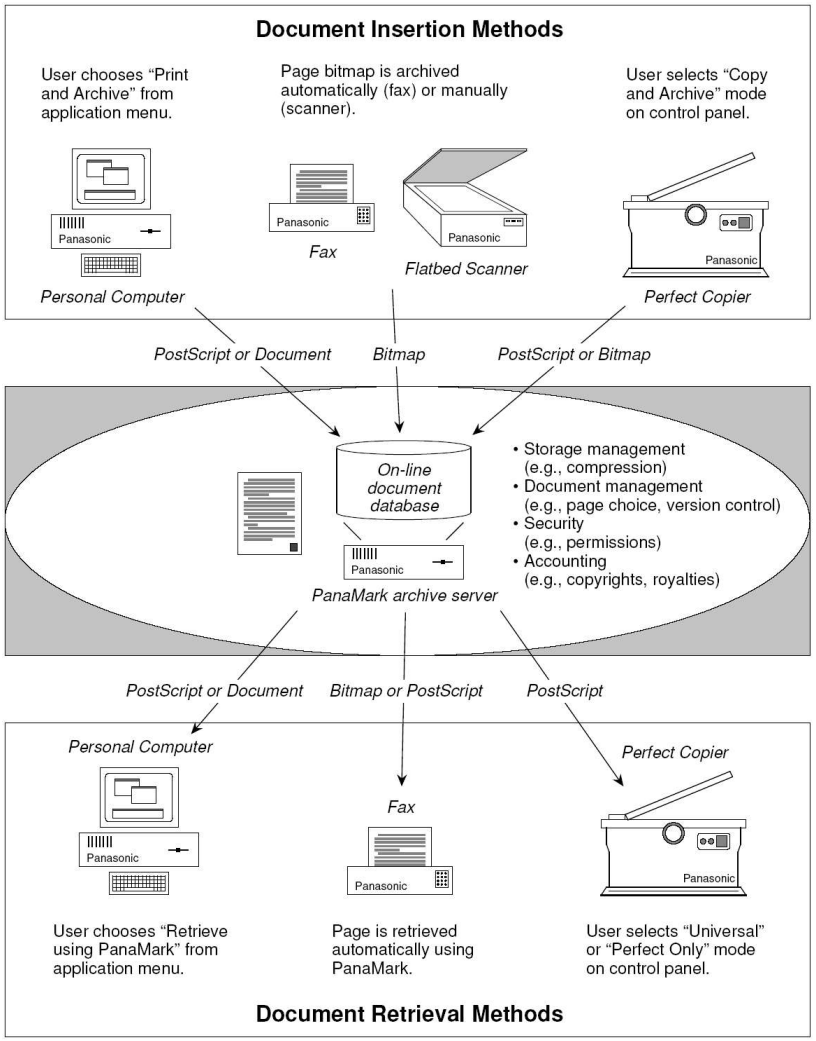


Slide circa 1995. Certain material contained herein is owned by MEI.

Document Archive

New approach centers around network-accessible archive of tagged documents.

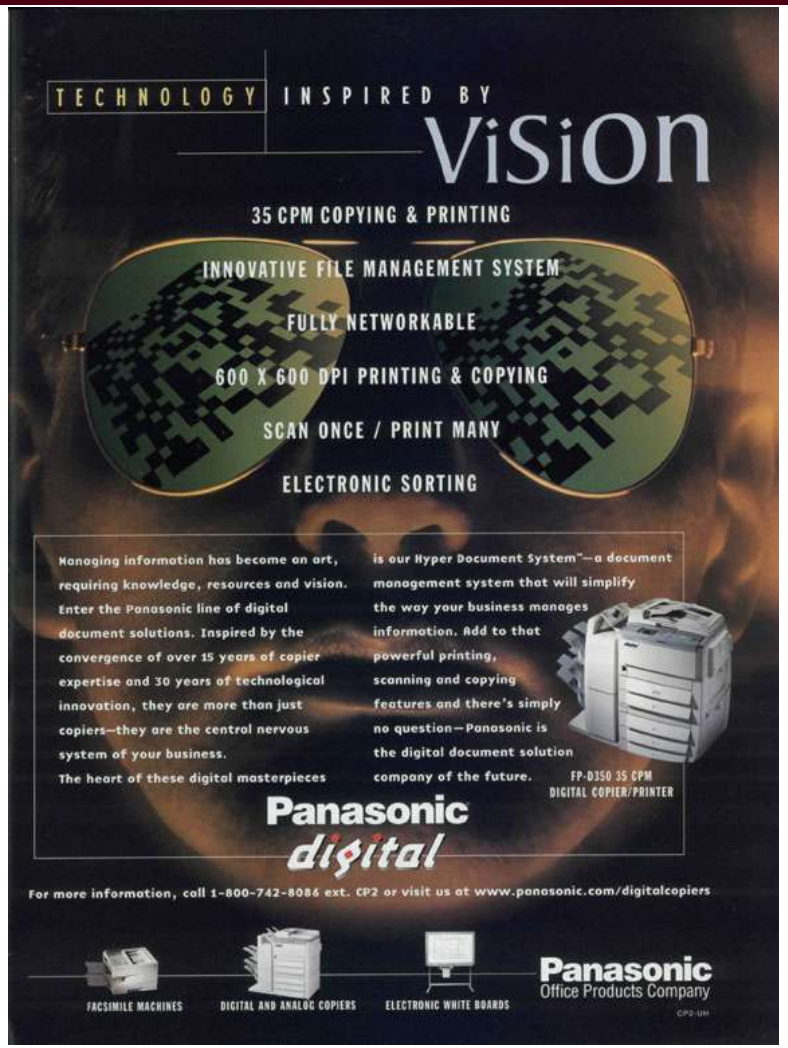
Multiple avenues for inserting and/or retrieving documents from archive.



Slide circa 1995. Certain material contained herein is owned by MEI.

Panasonic HyperDocument System

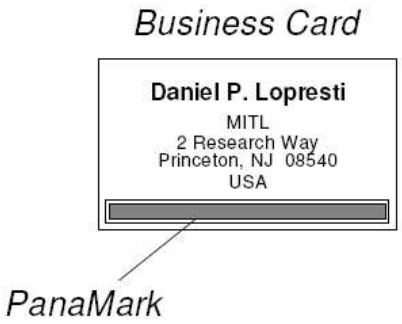
Panasonic ultimately introduced a line of advanced photocopiers based on these ideas:

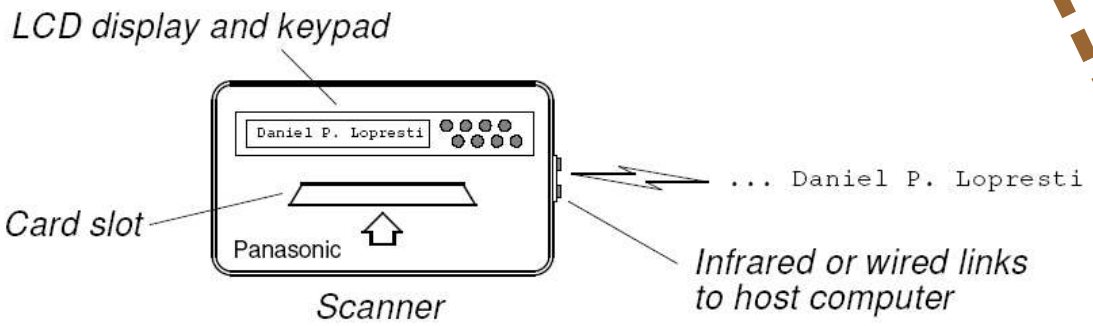


Ad from United Airlines inflight magazine.

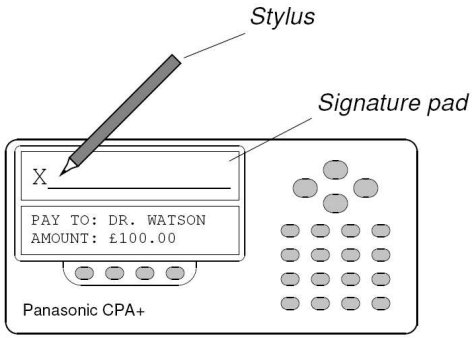
Other Applications

Business card data capture:





Forgery-resistant checks:



Slide circa 1995. Certain material contained herein is owned by MEI.

Related Patents Owned by MEI

"System and Method for Archiving Digital Versions of Documents and for Generating Quality Printed Documents Therefrom," D. P. Lopresti, J. Esakov, and J. Zhou, U.S. Patent No. 5,754,308, issued May 1998.

"Clock Free Two-Dimensional Barcode and Method for Printing and Reading the Same," D. P. Lopresti, J. Esakov, and J. Zhou , U.S. Patent No. 5,862,270, issued January 1999.

"Method of Locating a Machine Readable Two-Dimensional Barcode Within an Image," J. Zhou and D. P. Lopresti, U.S. Patent No. 5,974,200, issued October 1999.

"Clock Free Two-Dimensional Barcode and Method for Printing and Reading the Same," D. P. Lopresti, J. Esakov, and J. Zhou, U.S. Patent No. 6,115,508, issued September 2000.

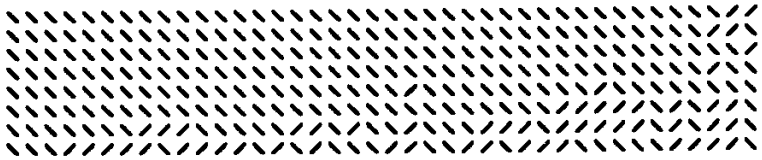
"Border-less Clock Free Two-Dimensional Barcode and Method for Printing and Reading the Same," J. Zhou, D. P. Lopresti, and A. Tomkins, U.S. Patent No. 6,201,901, issued March 2001.

"Border-less Clock Free Two-Dimensional Barcode and Method for Printing and Reading the Same," J. Zhou, D. P. Lopresti, and A. Tomkins, U.S. Patent No. 6,418,244, issued July 2002.

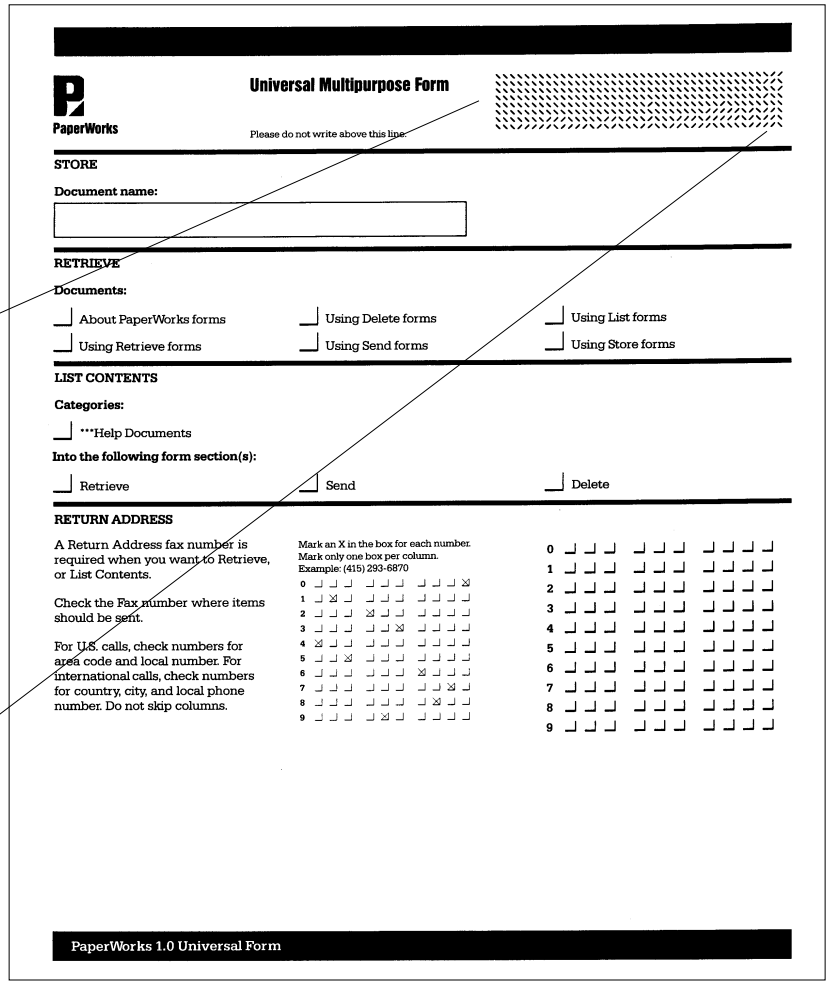
(Not necessarily an exhaustive list – only the ones Dan Lopresti was involved with.)

Related Work

Around the same time we were developing PanaMarks at MITL, Xerox introduced DataGlyphs and PaperWorks, technologies invented at Xerox PARC.



Xerox DataGlyph



PaperWorks

Universal Multipurpose Form

Please do not write above this line.

STORE

Document name:

RETRIEVE

Documents:

About PaperWorks forms | Using Delete forms | Using List forms

Using Retrieve forms | Using Send forms | Using Store forms

LIST CONTENTS

Categories:

***Help Documents

Into the following form section(s):

Retrieve | Send | Delete

RETURN ADDRESS

A Return Address fax number is required when you want to Retrieve, or List Contents.

Check the Fax number where items should be sent.

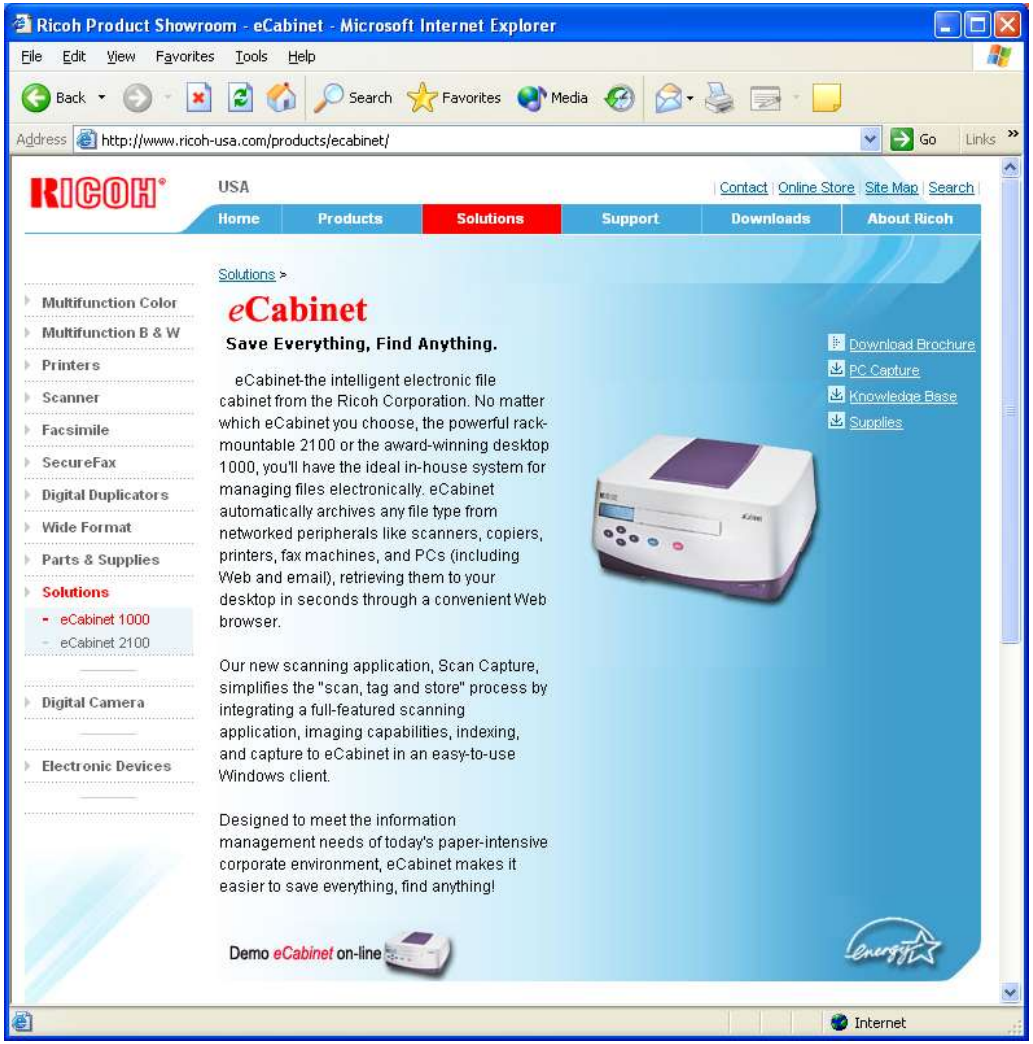
For U.S. calls, check numbers for area code and local number. For international calls, check numbers for country, city, and local phone number. Do not skip columns.

Mark an X in the box for each number. Mark only one box per column. Example: (415) 293-6870

PaperWorks 1.0 Universal Form

Related Work

Ricoh's eCabinet product incorporates similar (and in some ways more advanced) document management functionality, but, so far as we know, no notion of a unique DocID.



Chipless ID & Everyday Documents

Other chipless ID technologies can function somewhat like a printed 2-D barcode, with every piece of paper getting its own unique ID.

Advantages we see in alternative chipless ID:

- Doesn't disfigure original document like barcodes.
- May be more robust than printed barcodes.
- More flexible encoding – ID can be embedded in ink or in paper.
- Non-optical reading allows new paper handling functionality.
- Some chipless ID schemes are unique to physical page, invisible, and resist reproduction, which creates opportunities for document security, both "heavy-" and "lightweight."

  heavyweight = passport lightweight = letter of recommendation

Chipless ID & Everyday Documents

We have noted some key differences between information (as embodied in documents) and objects:

- Objects must exist in physical form. Information need not exist in physical form, but it's often desirable when it does.
- Information has long-term archival value often measured in years or decades. Objects usually don't.
- Objects always have owners. Information sometimes doesn't.
- At a given point in time, objects must exist at some specific physical location. Information can exist in many places at once.

Chipless ID & Everyday Documents

Technical challenges (potential R&D activities):

- Developing naming schemes & ID resolution services (standards).
- Integrating chipless ID reader with existing copiers / scanners.
- Addressing implementation issues (tag interference, lost links).
- Security (can it be broken?) and privacy (how to achieve?).
- Applications using lightweight security (including forms processing, digital rights management).
- Hardcopy audit trails for electronic voting (NSF interest?).

Does this fit within vision for proposed RPI Chipless ID Center?