

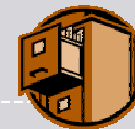
Session Five: The Frontiers of Search

Searching Large Email Collections: the Next Challenge

Diederik van Braam van Vloten
ZyLAB, The Netherlands

THE #1 SOLUTION

TO INVESTIGATE, MANAGE AND
SHARE YOUR INFORMATION



E-mail analysis and archiving

How expensive can searching and managing email be? We all know the answer: **VERY EXPENSIVE!**

But, things can be different.... For over 20 years, we've been working in partnership with companies and government agencies who require the best and most cost-effective tools for conducting detailed, large-scale data investigations.



Investigate & Share Information

- Find "smoking guns" faster, at lower costs and with a smaller team.
- Share knowledge more effectively.
- Capture, organize and disclose your data in the most efficient and secure manner possible.
- Benefit from open technology and data storage, suitable in standard IT environments.
- Implement a complete, front-to-back single-vendor solution: from capture to disclosure.



Structure

- Short Introduction to ZyLAB
- Problem: Searching large e-mail collections
- Why existing Solutions do not always work
- Alternative Strategies: Best Practices
- Additional Technological Innovations to Fit Specific Processes
- Conclusions and overview



ZyLAB markets: all investigators



Law enforcement



Legal



Knowledge
management



Financial /
Office
Management



20 years of award winning technology with 7,500 installations worldwide

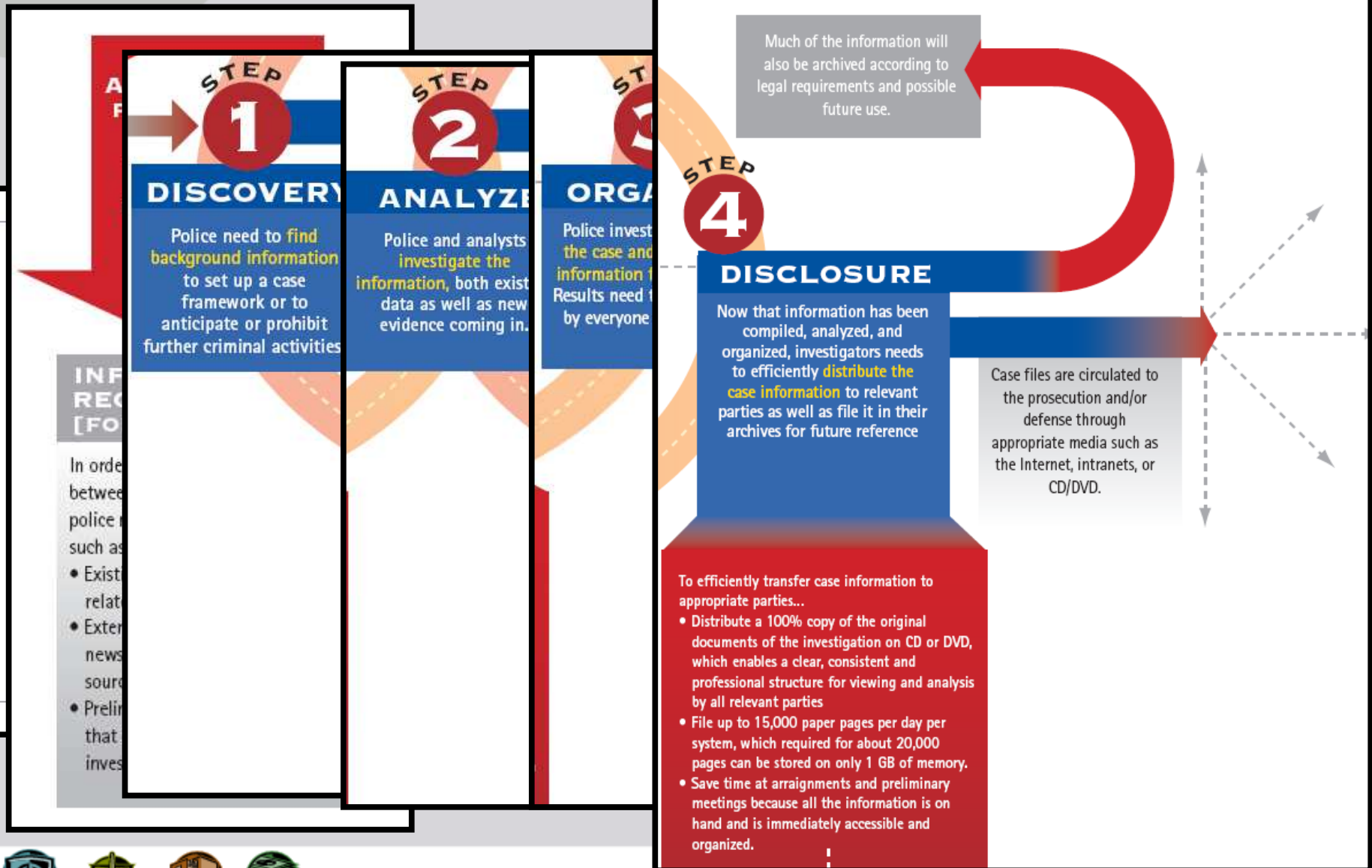


Problem: searching large e-mail collections

- **Collection / Discovery:**
 - Different formats, PST, Lotus Notes, Exchange, MIME, etc.
- **Analyze:**
 - Language Issues
 - Lots of Duplicates + Repetition of Information
 - Attachments
- **Organize:**
 - Multiple formats and sources that need to be managed
 - Inherent Metadata of various Email Formats
- **Disclosure:**
 - Multi lingual, multiple formats and selection of data necessary for disclosure



Problem: searching large e-mail collections



Why existing solutions do not always work

Technological issues:

- Emails contain bitmaps and other images
- Multi Lingual
- Other Algorithms don't work either:
 - Relevance Ranking
 - Advanced Language Technology
- 100% Recall -> Low Precision



Why existing solutions do not always work

- When people are searching large Email collections
 - Are not formally trained in advanced full-text retrieval
 - May not know what they're looking for
 - Must prove that all evidence/info has been considered in their decisions, verdicts, and so on
 - Cannot miss *any* relevant documents
 - They require **100% recall!**



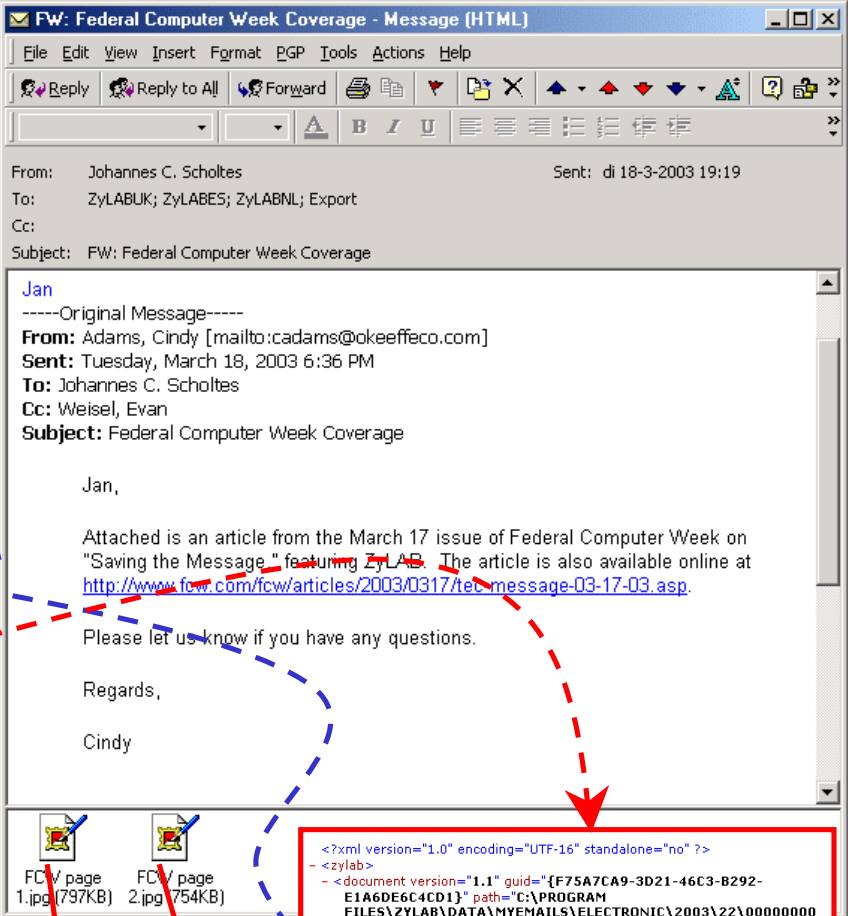
- Convert Email to XML
- Specific Email Technologies
- 100% Recall -> overcome Low Precision by using:
 - Hit Highlighting
 - Hit Navigation
 - Fuzzy
 - Visualization
 - Refine Results
 - KWIC
 - Text mining
 - Transliteration
 - Translation
- Other formats: paper, electronic, multimedia
- Annotation with eDiscovery (cross references)



Email Conversion to XML

```

- <zylab>
- <document version="1.1" guid="{9AA600D5-9098-4398-BBF1-DF9E02828D32}"
  path="C:\PROGRAM FILES\ZYLAB\DATA\MYEMAILS\ELECTRONIC\2003\22\00000000\
  name="20030318191838_A44E2000.HTM">
- <fields>
  <field id="emxsnd">Johannes C. Scholtes</field>
  <field id="emxrcx">ZyLABUK; ZyLABES; ZyLABNL; Export</field>
  <field id="emxrcp">ZyLABUK, ZyLABES, ZyLABNL, Export</field>
  <field id="emxstd">20030318</field>
  <field id="emxrdt">20030310</field>
  <field id="emxsub">FW: Federal Computer Week Coverage</field>
- <field id="emxbdy">
  <|>FYI,</|>
  <|>Jan</|>
  <|>-----Original Message-----</|>
  <|>From: Adams, Cindy [mailto:cadams@okeeffeco.com]</|>
  <|>Sent: Tuesday, March 18, 2003 6:36 PM</|>
  <|>To: Johannes C. Scholtes</|>
  <|>Cc: Weisel, Evan</|>
  <|>Subject: Federal Computer Week Coverage</|>
  <|>
  <|>Jan</|>
  <|>
  <|>Attached is an article from the March 17 issue of Federal Computer Week on
  <|>"Saving the Message," featuring ZyLAB. The article is also available online at
  <|>http://www.fcw.com/fcw/articles/2003/0317/tec-message-03-17-03.asp.</|>
  <|>
  <|>Please let us know if you have any questions.</|>
  <|>
  <|>Regards</|>
  <|>
  <|>Cindy</|>
  <|>
  </field>
- <field id="emxdoc">
- <documents>
  <document name="FCW page 1.jpg" type="jpg" guid="{1B2F8D3B-47AD-4250-9FEB-
  1F8A50D295BB}" />
  <document name="FCW page 2.jpg" type="jpg" guid="{F75A7CA9-3D21-46C3-B292-
  E1A6DE6C4CD1}" />
  </documents>
  </field>
  <field id="emxmsg" />
</fields>
</document>
</zylab>
  
```



```

<?xml version="1.0" encoding="UTF-16" standalone="no" ?>
- <zylab>
- <document version="1.1" guid="{1B2F8D3B-47AD-4250-9FEB-
  1F8A50D295BB}" path="C:\PROGRAM
  FILES\ZYLAB\DATA\MYEMAILS\ELECTRONIC\2003\22\00000000\
  name="20030318191838_FCW PAGE 1.JPG">
- <fields>
  <field id="emxmsg">
  <document description="FW: Federal Computer Week
  Coverage, 2003-03-18 19:18:38" type="outlook"
  guid="{9AA600D5-9098-4398-BBF1-DF9E02828D32}" />
  </field>
  </fields>
</document>
</zylab>
  
```

```

<?xml version="1.0" encoding="UTF-16" standalone="no" ?>
- <zylab>
- <document version="1.1" guid="{F75A7CA9-3D21-46C3-B292-
  E1A6DE6C4CD1}" path="C:\PROGRAM
  FILES\ZYLAB\DATA\MYEMAILS\ELECTRONIC\2003\22\00000000
  \name="20030318191838_FCW PAGE 2.JPG">
- <fields>
  <field id="emxmsg">
  <document description="FW: Federal Computer Week
  Coverage, 2003-03-18 19:18:38" type="outlook"
  guid="{9AA600D5-9098-4398-BBF1-DF9E02828D32}" />
  </field>
  </fields>
</document>
</zylab>
  
```



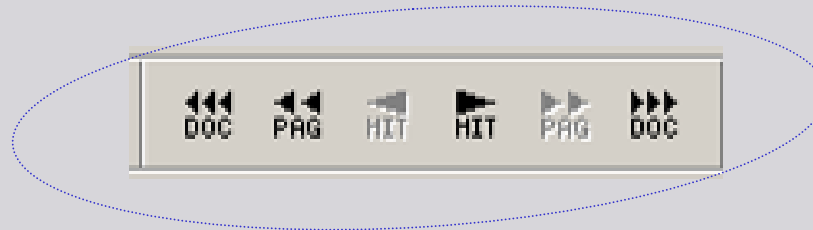
Specific Email Technologies

- Double message removal (with HASH)
- Automatic Bitmap OCR in the background
- Attachment Indexation



Hit Navigation / Hit Highlighting

- Document & Hit Navigation



- Hit Highlighting on the Image

IN WITNESS WHEREOF, this Agreement

CONT

CONTRACTOR'S NAME (If other than an individual, st

Enterprise Rent A Car

BY (Authorized Significe)

Susie M. Irwin

PRINTED NAME AND TITLE OF PERSON SIGNING

Susie Irwin

ADDRESS

4515 Auburn Blvd., Sacramento, CA 95841

Fuzzy Search increases recall, but lowers precision.

Search on:

“Syimage and Wordscan and Carlin and Calera”

In low-quality OCR text with transliterations and spelling errors.

ZyImage 3.0 will facilitate distribution on CD-ROMs

Boasts integration with **WordScan** OCR software

BY WILLEM KNIBBE

ZyLab is shipping this week an upgrade to its **ZyImage image** and text retrieval software that will help users distribute their document databases on CD-ROM.

ZyImage 3.0 also features batch processing improvements and tight integration with the latest version of **Caere Recognition Systems Inc.'s WordScan** optical character recognition (OCR) software.

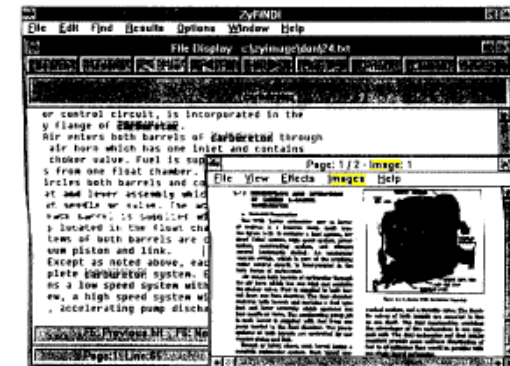
ZyImage 3.0 costs \$995. But the application is integrated with **WordScan** and cannot be run independently. **WordScan** must be purchased separately for \$595.

"Our integration is not based on toolkits or OLE 2.0, so we can offer all of **WordScan's** features," said ZyLab president **Jon Karlin**, referring to features such as page zoning, auto-page orientation, and batch processing.

ZyLab has also integrated APIs to Meridian Data Inc.'s CD-ROM recorders so that users can easily plug in the \$8,295 NetScribe or Personal Scribe recorders, which sell for \$4,195 to \$6,195.

ZyImage's new ZyPublish module lets users drag and drop to select which components — text, document **images**, indexes, and a run-time of the ZyLab search module — to publish to CD-ROM.

ZyPublish and the Meridian



ZyImage's fuzzy search capabilities make up for misspelled search terms or OCR errors by finding inexact matches.

APIs are intended to let users distribute CD-ROMs as an alternative to the write-once-read-many (WORM) optical disk **storage** that **ZyImage** already supports.

Other new features in **ZyImage** include Messaging API support and **Image Coding**, which speeds batch processing by letting users wait until a group of documents have been scanned before viewing document **images** and entering key **words**.

ZyImage also includes an updated version of its search module, which includes fuzzy searching capabilities that finds **words** close to the search word but not an exact match.

"Fuzzy search capabilities are

especially important for OCR'd text to make up for inaccuracies," **Karlin** said. "It's also important to let users misspell search **words** and still find the text."

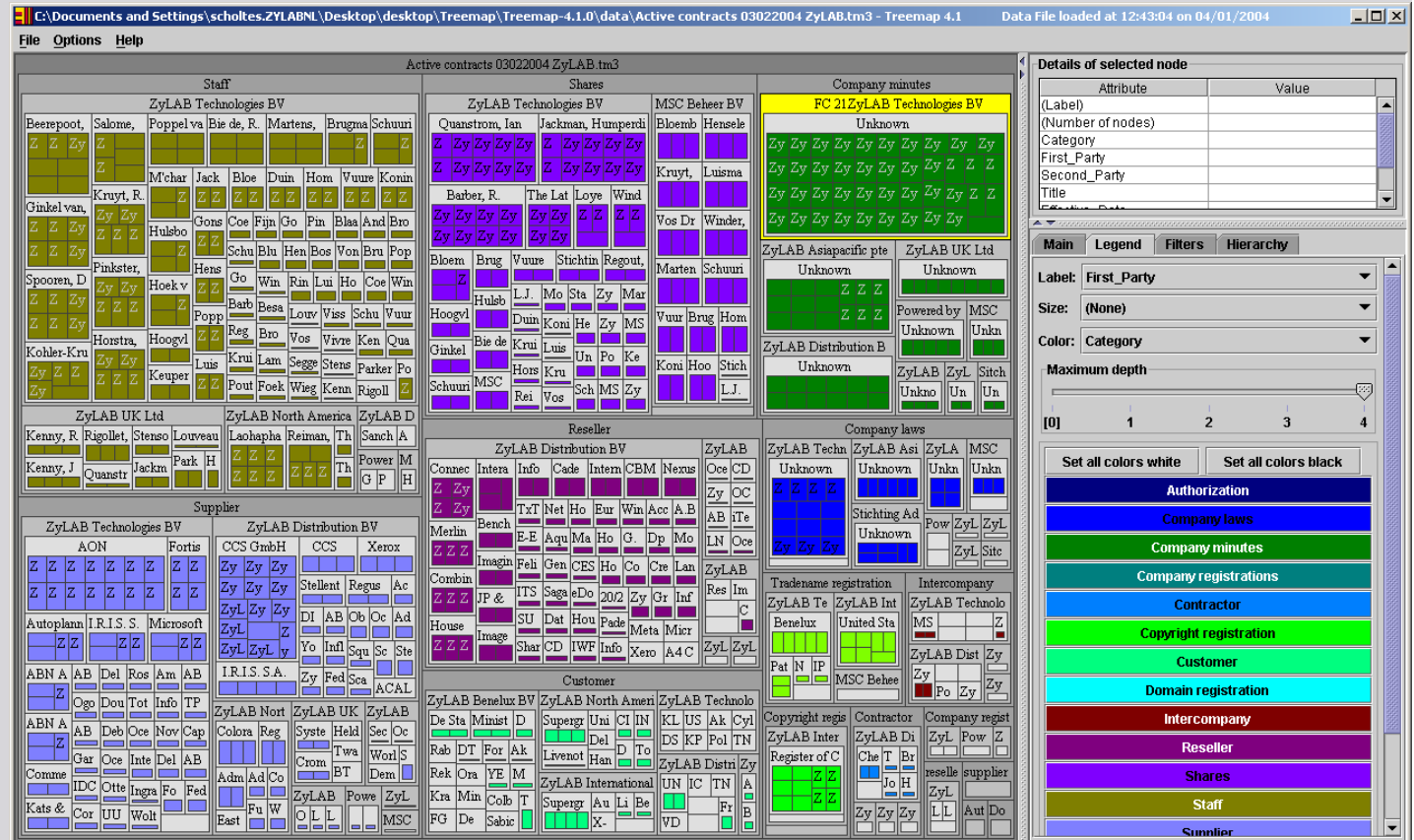
By the end of the year, the company will ship the fuzzy search capabilities in an update to its \$395 ZyIndex text indexing and retrieval product.

ZyLab is also working on a \$500 single-user version of **ZyImage called ZyImage Personal Edition**, due for release in November through ZyLab's VAR channel, according to **Karlin**.

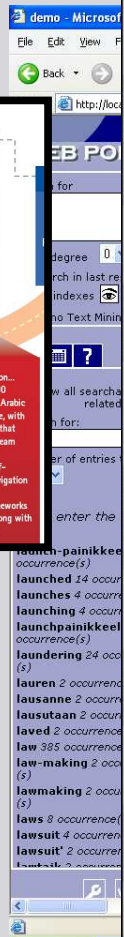
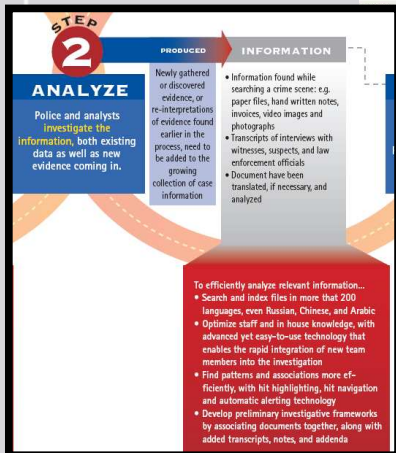
ZyLab, a division of ZyCo International, is in Gaithersburg, Md., at (800) 544-6339.



Visualization: The Treemap



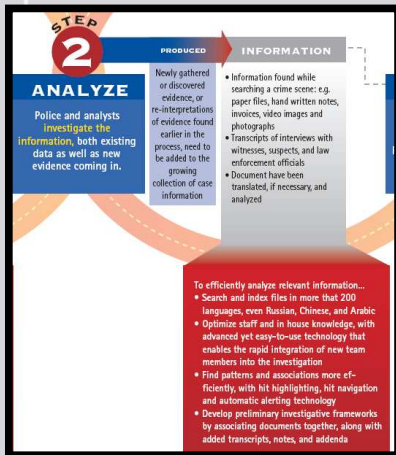
Analyze using Keywords in context



- 1 [0000002E.xml:Document World: Review,donderdag 1 april 1999 \(0000002E.xml\)](#): **6 hits, hit density 0,72**
of special European **law** enforcement agencies ... provide security and **law** enforcement agencies with ... query multiple European **law** enforcement databases in ... algorithms, international intelligence, **law** enforcement agencies, lingual
- 2 [0000013J.xml:South African Law Commission: Annual Report,maandag 20 november 2000 \(0000013J.xml\)](#): **116 hits, hit density 0,67**
<Sous_titre></Sous_titre> SOUTH AFRICAN **LAW** COMMISSION 1995 ANNUAL ... the South African **Law** Commission Act, 1973 ... the South African **Law** Commission are situated ... Secretary South African **Law** Commission Private Bag
- 3 [2274_01.xml:ZyLAB: Review,dinsdag 29 november 1994 \(2274_01.xml\)](#): **25 hits, hit density 0,51**
to suggest that **law** firm work product ... was reminded that **law** firms, except for ... ones which hire **law** office management consultants, ... best program for **law** office use. (InfoWorld
- 4 [3F_01.xml:Kofax: Press Release,maandag 15 april 1996 \(3F_01.xml\)](#): **2 hits, hit density 0,18**
company markets to **law** firms, public utilities, ... firms, employment agencies, **law** enforcement agencies and
- 5 [1715_01.xml:Fujitsu: Brochure, Public Relations,woensdag 30 november 1994 \(1715_01.xml\)](#): **2 hits, hit density 0,17**
500, including FileNet, **Law** Cypress, Opti- ka ... Imagelink, Intrafed, Kofax, **Law** Cypress, M3099A, MAGAZINE,
- 6 [3G_01.xml:Wang/ZyLAB International: Press release,maandag 15 april 1996 \(3G_01.xml\)](#): **2 hits, hit density 0,14**
and the Internet. **Law** firms, public utilities, ... employment agencies and **law** enforcement agencies have
- 7 [0000008O.xml:ZyLAB: Press Release,woensdag 8 november 2000 \(0000008O.xml\)](#): **1 hits, hit density 0,11**
are many security, **law** enforcement and intelligence
- 8 [0000008P.xml:ZyLAB: Press Release,woensdag 8 november 2000 \(0000008P.xml\)](#): **1 hits, hit density 0,10**
are many security, **law** enforcement and intelligence
- 9 [0000008N.xml:ZyLAB: Press Release,woensdag 8 november 2000 \(0000008N.xml\)](#): **1 hits, hit density 0,10**
are many security, **law** enforcement and intelligence
- 10 [0000008M.xml:ZyLAB: Press Release,woensdag 8 november 2000 \(0000008M.xml\)](#): **1 hits, hit density 0,10**
are many security, **law** enforcement and intelligence
- 11 [39_01.xml:ZyLAB UK: Brochure,maandag 15 april 1996 \(39_01.xml\)](#): **2 hits, hit density 0,09**
research materials • **Law** firms use it ... and research • **Law** enforcement agencies use
- 12 [2257_01.xml:ZyLAB: Newsletter, Public Relation,dinsdag 29 november 1994 \(2257_01.xml\)](#): **3 hits, hit density 0,09**
metropolitan area" DC **LAW** FIRM'S INFO MANAGEMENT ... LEGAL RESEARCH Historically, **law** firms have been ... CD-ROMs, CDs, DC **LAW** FIRM, Fuzz, Fuzzy
- 13 [0000008U.xml:ZyLAB: Press Release,woensdag 8 november 2000 \(0000008U.xml\)](#): **1 hits, hit density 0,09**
are many security, **law** enforcement and intelligence



Analyze using Refine Results



demo - Microsoft Internet Explorer provided by ZyLAB Technologies BV

Address: http://localhost/Exec/ZYNET.exe?Client=demo&nit=1&ZyAction=ZyAction

WEB PORTAL

Search for: Law

Fuzzy degree: 0

Search in last results

Select indexes: [checked]

Demo Text Mining: [checked]

View all searchable words related document

Search for: law

Number of entries to show: 15

Please enter the word to

launch-painikkeella 2 occurrence(s)

launched 14 occurrence(s)

launches 4 occurrence(s)

launching 4 occurrence(s)

launchpainikkeella 2 occurrence(s)

laundering 24 occurrence(s)

lauren 2 occurrence(s)

lausanne 2 occurrence(s)

lausutaan 2 occurrence(s)

laved 2 occurrence(s)

law 385 occurrence(s)

law-making 2 occurrence(s)

lawmaking 2 occurrence(s)

laws 9 occurrence(s)

lawsuit 4 occurrence(s)

lawsuit 2 occurrence(s)

Search results for "law" (1 to 20 of 37)

results page: 1 - 2

Select	Description
<input checked="" type="checkbox"/>	1 0000002E.xml:Document World: Review_donderdag 1 ap density 0,12 of special European law enforcement agencies ... provid with ... query multiple European law enforcement datab intelligence, law enforcement agencies, lingual
<input checked="" type="checkbox"/>	2 00000131.xml:South African Law Commission: Annual Re (00000131.xml): 116 hits, hit density 0,67 <Sous_titre></Sous_titre> SOUTH AFRICAN LAW COMM African Law Commission Act, 1973 ... the South African Secretary South African Law Commission Private Bag
<input checked="" type="checkbox"/>	3 2274_01.xml:ZyLAB: Review_dinsdag 29 november 1994 to suggest that law firm work product ... was reminded t hire law office management consultants, ... best program
<input checked="" type="checkbox"/>	4 2F_01.xml:Kofax: Press Release_maandag 15 april 1996 company markets to law firms, public utilities, ... firms, agencies and
<input checked="" type="checkbox"/>	5 1745_01.xml:Fujitsu: Brochure, Public Relations.woensda hits, hit density 0,17 S00, including Filelet, Law Cypress, Opti- ka ... Imagel M3099A, MAGAZINE,
<input checked="" type="checkbox"/>	6 2G_01.xml:Wang/ZyLAB International: Press release.ma hit density 0,14 and the Internet, Law firms, public utilities, ... employe agencies have
<input checked="" type="checkbox"/>	7 00000080.xml:ZyLAB: Press Release.woensdag 8 novem density 0,11 are many security, law enforcement and intelligence
<input checked="" type="checkbox"/>	8 0000008P.xml:ZyLAB: Press Release.woensdag 8 novem density 0,10 are many security, law enforcement and intelligence
<input checked="" type="checkbox"/>	9 0000008H.xml:ZyLAB: Press Release.woensdag 8 novem density 0,10 are many security, law enforcement and intelligence
<input checked="" type="checkbox"/>	10 0000008M.xml:ZyLAB: Press Release.woensdag 8 novem density 0,10 are many security, law enforcement and intelligence
<input checked="" type="checkbox"/>	11 39_01.xml:ZyLAB UK: Brochure.maandag 15 april 1996 research materials • Law firms use it ... and research •
<input checked="" type="checkbox"/>	12 2257_01.xml:ZyLAB: Newsletter, Public Relation_dinsdag hits, hit density 0,09 metropolitan area* DC LAW FIRM'S INFO MANAGEMENT firms have been ... CD-ROMs, CDs, DC LAW FIRM, Fuzz
<input checked="" type="checkbox"/>	13 0000008U.xml:ZyLAB: Press Release.woensdag 8 novem density 0,09 are many security, law enforcement and intelligence
<input checked="" type="checkbox"/>	14 0000008K.xml:ZyLAB: Press Release.woensdag 8 novem density 0,09 On-line Technical Publications, Law Enforcement and Cor
<input checked="" type="checkbox"/>	15 0000008V.xml:ZyLAB: Press Release.woensdag 8 novem density 0,08 are many security, law enforcement and intelligence
<input checked="" type="checkbox"/>	16 2000003C.xml:ZyLAB: Brochure.woensdag 1 januari 199 density 0,07

Refine your results

Source

- Document World (1)
- Fujitsu (1)
- House of Inventions (1)
- Kofax (1)
- PC Magazine (1)
- South African Law Commission (1)
- Wang/ZyLAB International (1)
- www.persberichten.com (1)
- ZyLAB (22)
- ZyLAB Europe (3)
- ZyLAB International (3)
- ZyLAB UK (1)

Date

Document_type

- Annual Report (1)
- Brochure (13)
- Newsletter (1)
- Press Release (12)
- Public Relation (1)
- Public Relations (1)
- Review (6)
- White Paper (4)

COUNTRY

- Australia (2)
- Belgium (1)
- Canada (1)
- Estonia (1)
- France (10)
- Germany (10)
- Holland (1)
- Luxembourg (1)
- REPUBLIC OF SOUTH AFRICA (1)
- Singapore (22)
- South Africa (1)
- Spain (9)
- The Netherlands (12)
- U.S. (11)
- UK (1)
- United Kingdom (10)
- United States (2)
- US (4)
- USA (2)



Transliteration

Sophisticated tooling searches between different languages, even those with non-Latin characters (e.g. Cyrillic, Chinese and Arabic).

Example: Search will find spellings of “Gorbachov” in both Latin and Cyrillic characters.

Transliteration Assistant

Edit | History

Latin/Diacritic search criteria: G o r b a c h o v

Cyrillic search criteria: Г о р б а ц х о в

Capital letters: А Б Ц Ъ Ч Д Ъ Ц Е Ф Г Х И Ј К Л Љ М Н Њ О П Р С Ш Т У В З Ж Џ Ї Љ Ђ Ѓ Є Ѕ Ž

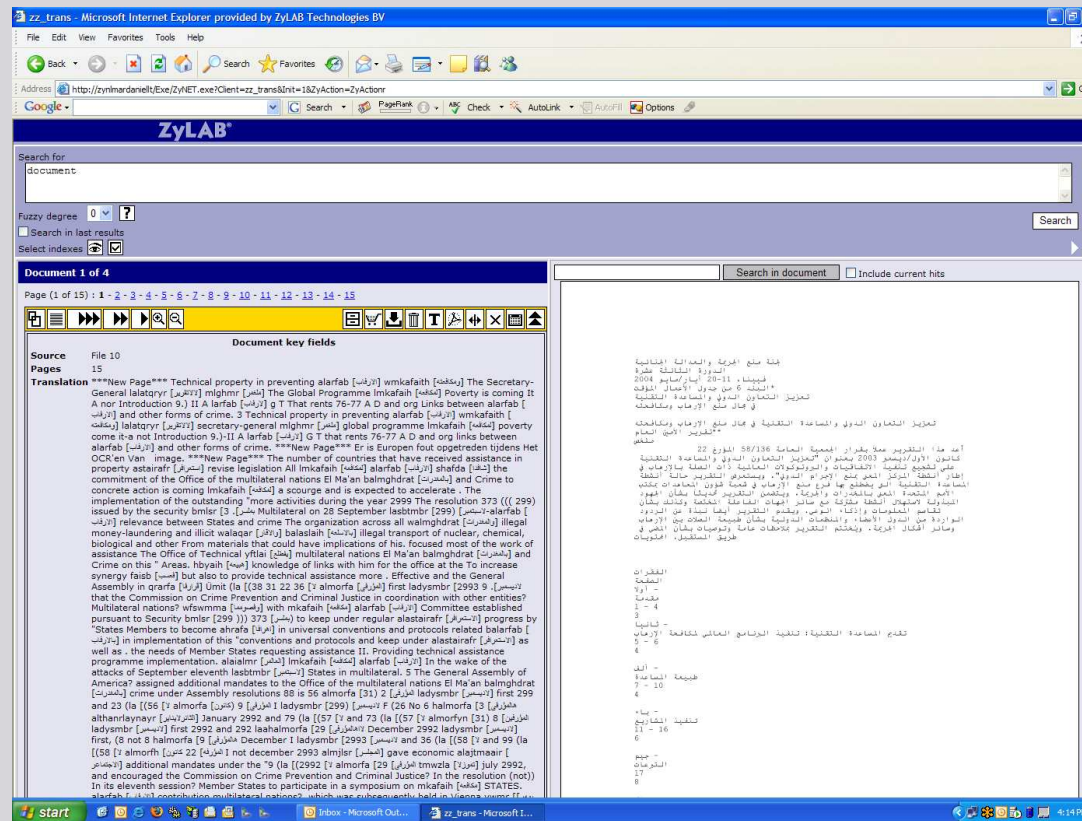
Small letters: а б ц њ ч д њ ц е ф г х и ј к л љ м њ о п р с ш т у в з ж џ ѝ ѡ Ѣ ѣ ѥ ѧ Ѩ

Latin/Diacritic Query Cyrillic Query Latin/Diacritic and Cyrillic Query

Select Language Clear Insert

Automatic translation software

- Support for Russian, Arabic, Farsi and many relevant languages.
- Works with paper, e-mail and electronic files.



Electronic Files in E-mail

- Native file format indexing of over 370 electronic file formats (PDF, Word, XLS, PPT, ...).
- Fast full-text retrieval and alerting.
- Hit highlighting and hit navigation.
- Convert e-mail to open XML data.
- Identify duplicates by using SHA1 hashes.

The screenshot displays the ZyView application window. The title bar indicates the file path: C:\PROGRAM FILES\ZYLAB\INDEX DATA\ZZ_DEM_I\ELECTRONIC\2004\00000000\Indiana Sherrif Council... The interface includes a menu bar (File, Edit, View, Navigate, Image, Actions, Redaction, Workflow, Document, Help) and a toolbar with various search and navigation icons. A table with two columns, 'Field' and 'Value', is visible, containing the following information:

Field	Value
Title	Applying the ZyIMAGE solution's competitive business advantage to I
Source	ZyLAB
Date	Tuesday, August 12, 2003
Document_tyr	Case Study
Remarks	Word Document
ZyLAB	
Product	

Below the table, a preview of the document is shown. It features a 'CASE STUDY' header, the ZyLAB logo (The Paper Filing Company), and an image of a laptop displaying a website. The main text of the document reads:

Applying the ZyIMAGE solution's competitive business advantage to Indiana sheriffs' real-time information retrieval needs

In addition to **law-enforcement** officials' more obvious responsibilities, such as being on the street catching suspected criminals, managing crime scenes, interviewing witnesses, and so on, a large part of their time is spent handling **extensive** administrative activities. These activities, while seemingly mundane, support all the other duties law **enforcement** officials perform. Having all the facts available and the necessary paperwork in order can mean the difference between pursuing the right or wrong course of action or **even** winning or losing a case. For these reasons, more and more law **enforcement** officials are turning to ZyLAB's newest custom solution, ZyIMAGE Investigate, to handle their special information archiving and retrieval activities.

The status bar at the bottom of the window shows: For Help, press F1 | Total Hits: 32 | Doc 12 of 25 | Query: fname Inde: //



Multimedia

- 100% original retrieval
- Linked to Multimedia object
- Add textual information to MM object.
- Ability to annotate
- Search on annotation
- Play original document

ZyView - [C:\PROGRAM FILES\ZYLAB\INDEX DATA\ZZ_MULTI\ELECTRONIC\2003\00000000\Eviden...

Field	Value
Title	Evidence Case 1234: Confiscated on suspect #1
Source	ZyLAB
Date	Tuesday, August 12, 2003
Document_type	Picture
Remarks	JPEG File

ZyView - [C:\PROGRAM FILES\ZYLAB\INDEX DATA...

Field	Value
Title	Surveillance: Suspect #3 at parking lot
Source	ZyLAB
Date	Tuesday, August 12, 2003
Document_type	Picture
Remarks	JPEG File



Example of manual structuring data

- Create annotations on documents and search the annotations later.

The screenshot displays the ZyLAB software interface. The main window, titled "Case 1234 - ZyFIND", shows a search query field and a list of documents. A document titled "Comment" is selected, showing a list of annotations: <MEMOTEXT>Signed contract by, <MEMOTEXT>Look at this signature, and <MEMOTEXT>Look at this signature. A red circle highlights the "Case Management" and "Case Management search" buttons in the left sidebar. A red arrow points from the annotations to a "Case management - [Demonstration Contract Management] [Re:] [R...]" window. This window shows a text input field with "Now I see it, good catch!" and a "Type" dropdown set to "Evidence". The "Author" field is set to "LISA" and the "Created" date is "Thursday, 31 March, 2005". A red arrow also points from the annotations to a document preview window showing a scanned document with a blue highlighted section and an "APPROVED" stamp.



Additional technological innovations to fit specific processes

- Redaction
- Disclosure via DVD
- Disclosure via WEB



Redaction and ZyIMAGE

- Redaction is a ZyFIND tool.
- Redacted data can be exported including redactions to 3rd parties or EDS

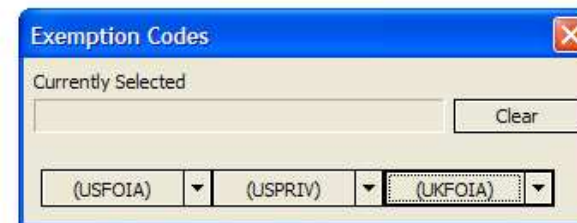
State departments are strongly encouraged to rent alternative fuel vehicles when



(UKFOIA)Section 22 [REDACTED], Interim Director
Department of General Services

CH:JH:DC

Attachment



Exemption Codes

Currently Selected

Clear

(USFOIA) (USPRIV) (UKFOIA)

ZyIMAGE CD/DVD Publishing

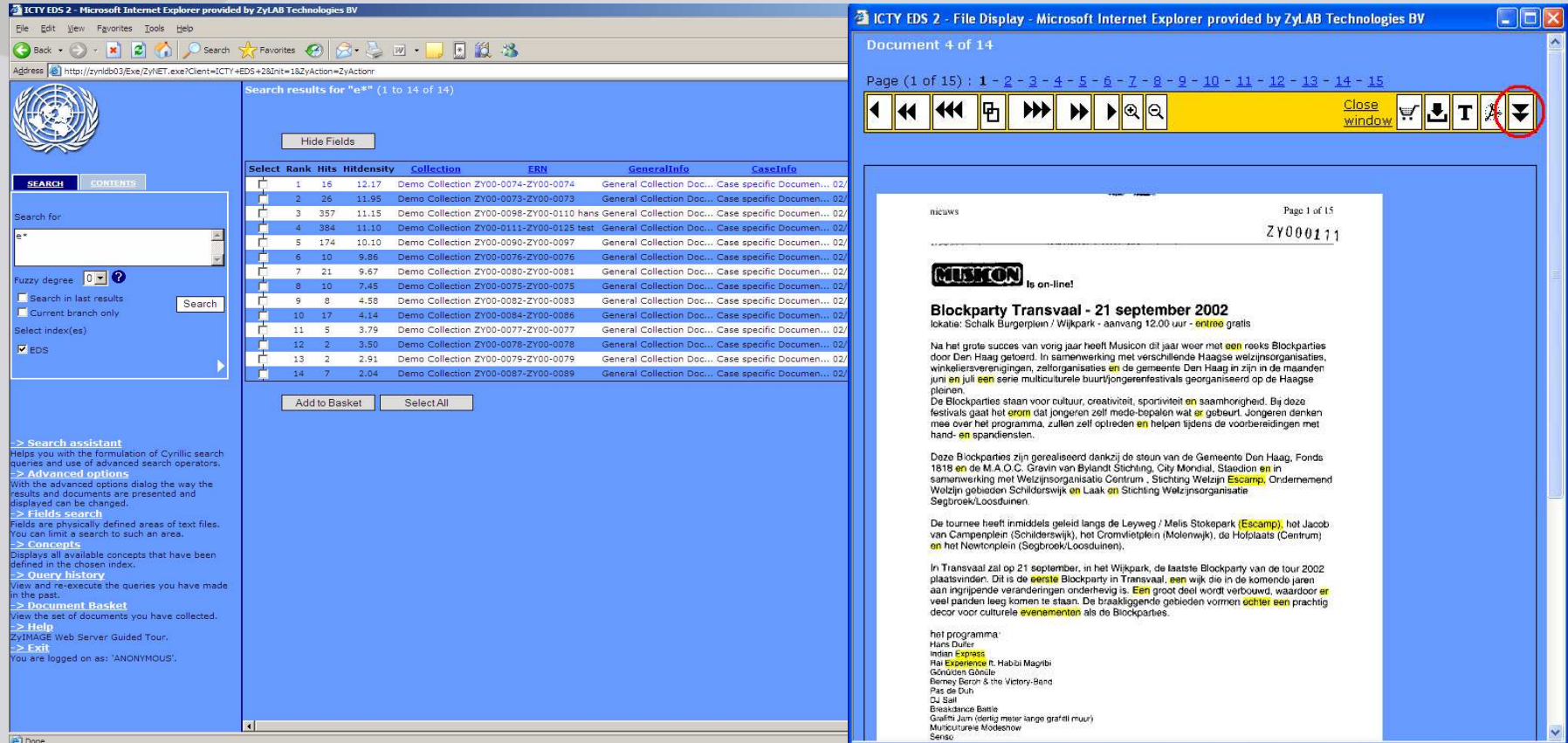
- Copy data to CD and DVD.
- Include a copy of the Windows search engine.
- Archive data.
- Share data with any 3rd party.

The screenshot shows the Zyview software interface. On the left, a 'CASE STUDY' window displays the title 'Applying the ZyIMAGE solution's competitive business advantage to Indiana time information retrieval'. On the right, a contract form is visible with the following text:

IN WITNESS WHEREOF, this Agreement
CONT
CONTRACTOR'S NAME (If other than an individual, st
Enterprise Rent A Car
BY (Authorized Signatory)
Susie Irwin
PRINTED NAME AND TITLE OF PERSON SIGNING
Susie Irwin
ADDRESS
4515 Auburn Blvd., Sacramento, CA 95841



UN Electronic Disclosure System (EDS)



The screenshot displays the UN EDS web interface in Microsoft Internet Explorer. The left pane shows search results for the query "e*" (1 to 14 of 14). The results table includes columns for Rank, Hits, Hitdensity, Collection, ERN, GeneralInfo, and CaseInfo. The right pane shows a document preview for "Document 4 of 14", which is page 1 of 15. The document content is in Dutch and discusses the "Blockparty Transvaal - 21 september 2002". A red circle highlights a download icon in the document viewer's toolbar.

Select	Rank	Hits	Hitdensity	Collection	ERN	GeneralInfo	CaseInfo
<input type="checkbox"/>	1	16	12.17	Demo Collection ZY00-0074-ZY00-0074	General Collection Doc...	Case specific Documen...	02/
<input type="checkbox"/>	2	26	11.95	Demo Collection ZY00-0073-ZY00-0073	General Collection Doc...	Case specific Documen...	02/
<input type="checkbox"/>	3	357	11.15	Demo Collection ZY00-0098-ZY00-0110 hans	General Collection Doc...	Case specific Documen...	02/
<input type="checkbox"/>	4	384	11.10	Demo Collection ZY00-0111-ZY00-0125 test	General Collection Doc...	Case specific Documen...	02/
<input type="checkbox"/>	5	174	10.10	Demo Collection ZY00-0090-ZY00-0097	General Collection Doc...	Case specific Documen...	02/
<input type="checkbox"/>	6	10	9.86	Demo Collection ZY00-0076-ZY00-0076	General Collection Doc...	Case specific Documen...	02/
<input type="checkbox"/>	7	21	9.67	Demo Collection ZY00-0080-ZY00-0081	General Collection Doc...	Case specific Documen...	02/
<input type="checkbox"/>	8	10	7.45	Demo Collection ZY00-0075-ZY00-0075	General Collection Doc...	Case specific Documen...	02/
<input type="checkbox"/>	9	8	4.58	Demo Collection ZY00-0082-ZY00-0083	General Collection Doc...	Case specific Documen...	02/
<input type="checkbox"/>	10	17	4.14	Demo Collection ZY00-0084-ZY00-0086	General Collection Doc...	Case specific Documen...	02/
<input type="checkbox"/>	11	5	3.79	Demo Collection ZY00-0077-ZY00-0077	General Collection Doc...	Case specific Documen...	02/
<input type="checkbox"/>	12	2	3.50	Demo Collection ZY00-0078-ZY00-0078	General Collection Doc...	Case specific Documen...	02/
<input type="checkbox"/>	13	2	2.91	Demo Collection ZY00-0079-ZY00-0079	General Collection Doc...	Case specific Documen...	02/
<input type="checkbox"/>	14	7	2.04	Demo Collection ZY00-0087-ZY00-0089	General Collection Doc...	Case specific Documen...	02/

- External access to court documents.
- Security per case, full audit included.



Conclusions

- Email is a different animal from those other searchable entities
- Technology alone is not enough to increase precision when you want 100% recall
- Better usability to increase precision:
 - Let the user decide what is relevant
- Optimal tools combine:
 - Full-text retrieval
 - Meta information search
 - Visualization and navigation
 - Manual classification
 - Controlled (semi) automatic classification



For Further Information

- Contact Details:

Name: Ir. Diederik van Braam van Vloten

Function: Channel Manager

E-mail: Diederik@ZyLAB.com

Tel: + 31 (0) 203119075

Mobile: + 31 (0) 654247901

Fax: + 31 (0) 206965175

Website: www.ZyLAB.com

