united nations
educational, scientific
and cultural
organization

international atomic
energy agency

the
abdus salam
international centre for theoretical physics

40th anniversary 1964 2004

SMR.1589 - 4

Workshop on
Managing Nuclear Knowledge

8 - 12 November 2004

Searching and Accessing Information

Heinz BACHMANN
CONVERA AG
Flawilerstrasse 27
9500 Wil
SWITZERLAND

These are preliminary lecture notes, intended only for distribution to participants

CONVERA OVERVIEW

IAEA

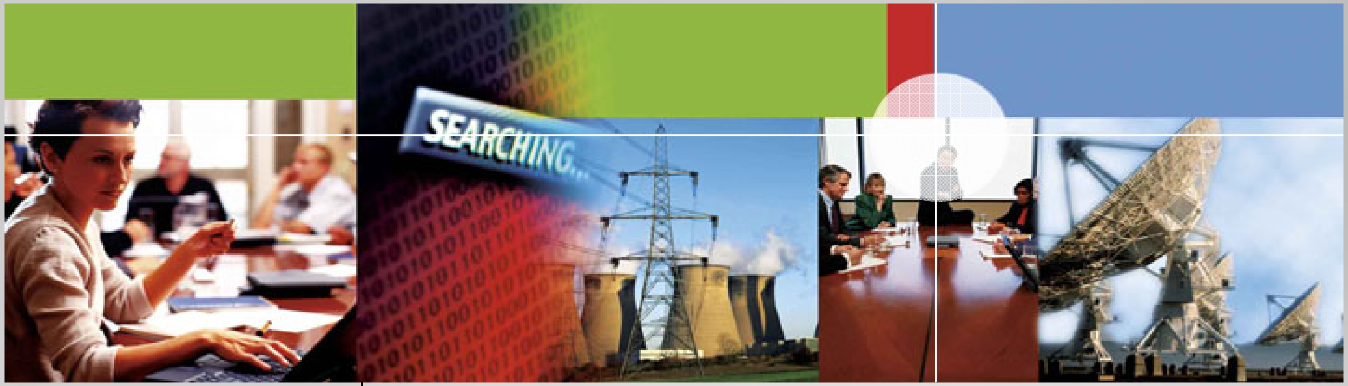
Trieste - November 2004

CONVERA

Agenda

- About CONVERA
- Introduction
- CONVERA Platform Infrastructure
- CONVERA‘s Retrieval Technologies
- Visualisation
- Automatic Categorization Technologies
- Dynamic Classification Technologies
- What we do / What‘s Third Party
- Product demonstration

Convera Corporation

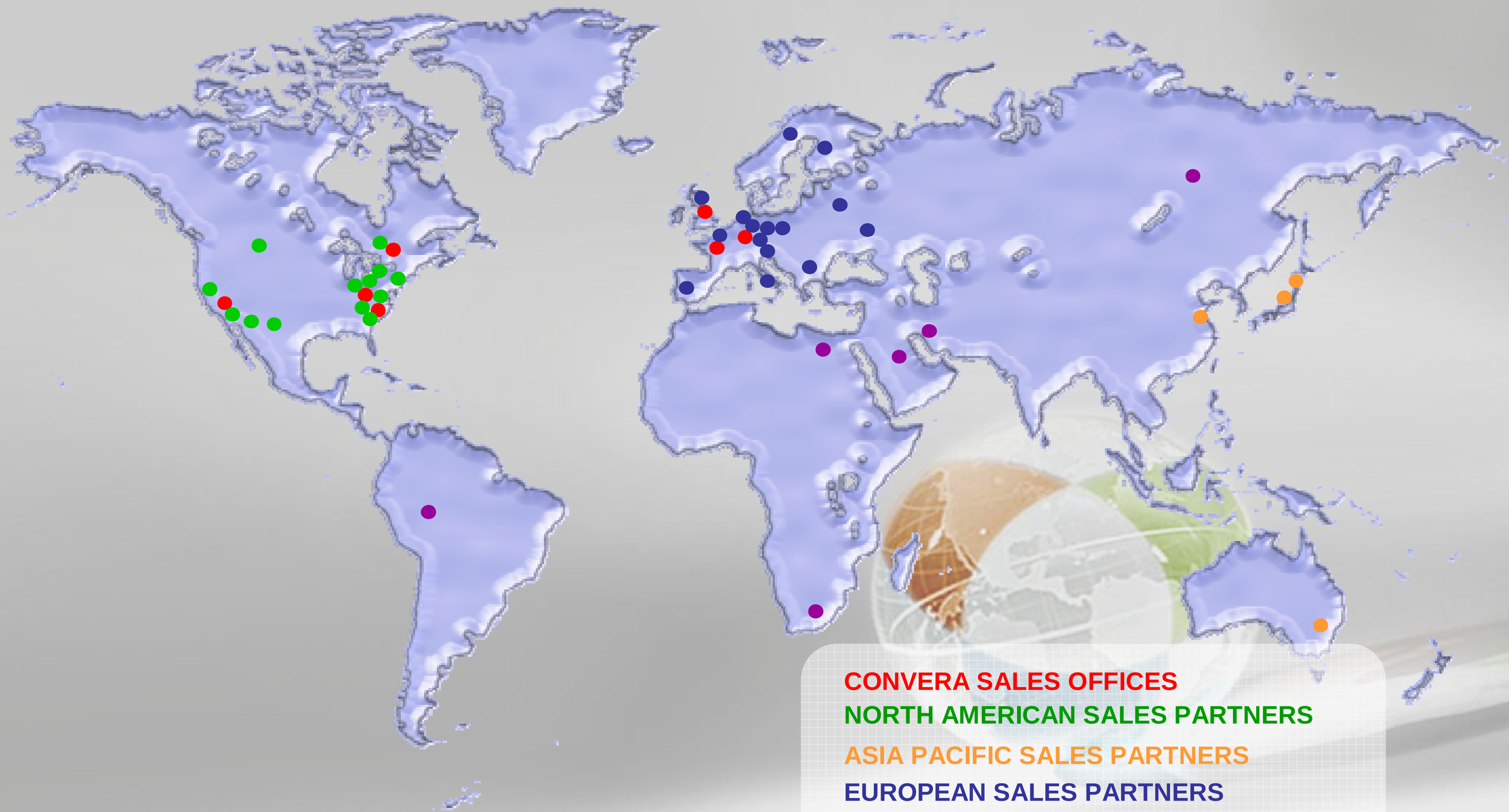


Convera is a leading provider of enterprise wide index, search and categorization software products and solutions

- 20+ years of innovation in intelligent information infrastructure
- 250 employees
- 800 customers in 29 countries
- 70+ business partners
- Publicly traded (NASDAQ: CNVR)

Convera Confidential

Convera & Partner Offices



CONVERA SALES OFFICES

NORTH AMERICAN SALES PARTNERS

ASIA PACIFIC SALES PARTNERS

EUROPEAN SALES PARTNERS

ROW SALES PARTNERS

Our Focus

LIFE SCIENCES

Johnson & Johnson

Aventis

NOVARTIS

Pfizer

AstraZeneca

Roche

gsk GlaxoSmithKline

FINANCIAL

ABN·AMRO

Goldman Sachs

HSBC

The Royal Bank of Scotland

GREENWICH CAPITAL

The Federal Reserve Board

Bloomberg

HIGH TECH

COGNOS

Netegrity

PTC

intel

intraspect

UNISYS We have a head for e-business.

symantec.

GOVERNMENT

SOCIAL SECURITY ADMINISTRATION USA

USGS science for a changing world

NASA

Sandia National Laboratories

IRS

FEDERAL TRADE COMMISSION

FDA

abc NEWS

ESPN

Discovery TLC Discovery Health online travel Animal Planet

Chicago Tribune INTERNET EDITION

NATIONAL GEOGRAPHIC TELEVISION

THOMSON

Government / Organizations

Law enforcement (Europe)

Interpol

Polisen

www.polisen.se

….and ?

Intelligence

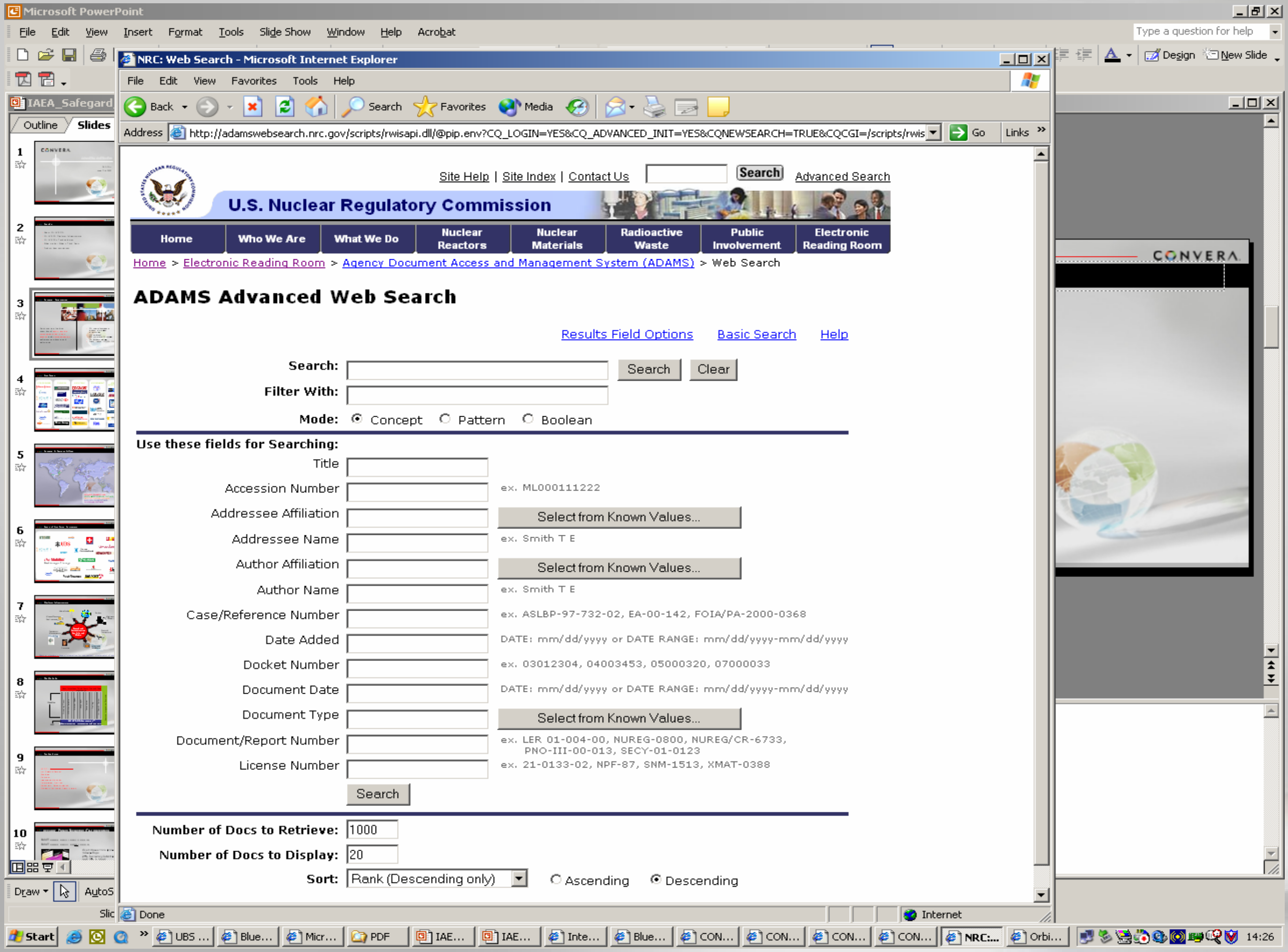
?

Organizations (Nuclear Industry)

UK AEA

Sandia National Laboratories

NRS – CONVERA RetrievalWare Public Accessible



Address http://adamswebsearch.nrc.gov/scripts/rwisapi.dll/@pip.env?CQ_LOGIN=YES&CQ_ADVANCED_INIT=YES&CQNEWSEARCH=TRUE&CQCGI=/scripts/rwis

Site Help | Site Index | Contact Us | Search | Advanced Search

U.S. Nuclear Regulatory Commission

Home > Electronic Reading Room > Agency Document Access and Management System (ADAMS) > Web Search

ADAMS Advanced Web Search

Results Field Options | Basic Search | Help

Search: [ ] Search Clear

Filter With: [ ]

Mode: ● Concept ○ Pattern ○ Boolean

Use these fields for Searching:

Title		
Accession Number		ex. ML000111222
Addressee Affiliation		Select from Known Values...
Addressee Name		ex. Smith T E
Author Affiliation		Select from Known Values...
Author Name		ex. Smith T E
Case/Reference Number		ex. ASLBP-97-732-02, EA-00-142, FOIA/PA-2000-0368
Date Added		DATE: mm/dd/yyyy or DATE RANGE: mm/dd/yyyy-mm/dd/yyyy
Docket Number		ex. 03012304, 04003453, 05000320, 07000033
Document Date		DATE: mm/dd/yyyy or DATE RANGE: mm/dd/yyyy-mm/dd/yyyy
Document Type		Select from Known Values...
Document/Report Number		ex. LER 01-004-00, NUREG-0800, NUREG/CR-6733, PNO-III-00-013, SECY-01-0123
License Number		ex. 21-0133-02, NPF-87, SNM-1513, XMAT-0388

Search

Number of Docs to Retrieve: 1000

Number of Docs to Display: 20

Sort: Rank (Descending only) ○ Ascending ● Descending

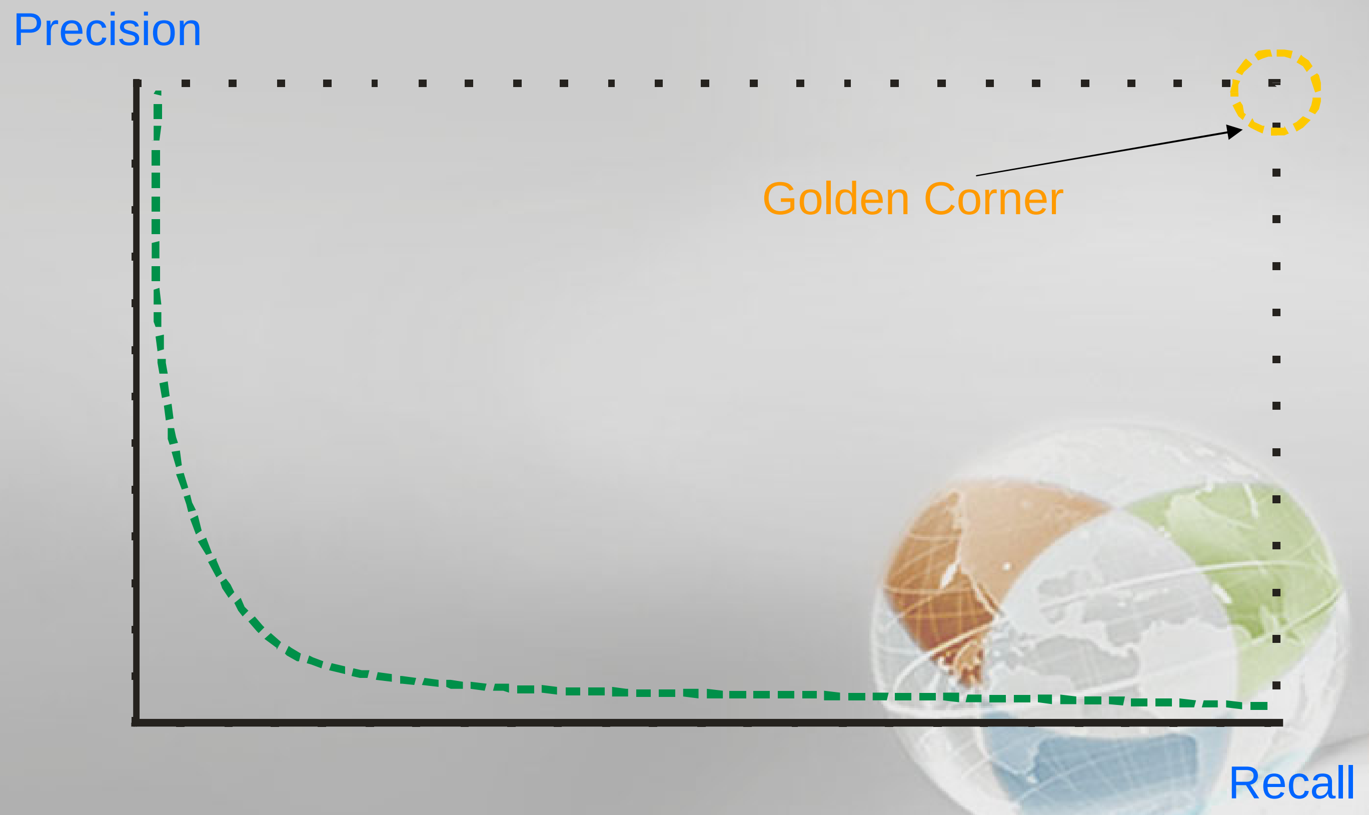
Microsoft PowerPoint
File Edit View Insert Format Tools Slide Show Window Help Acrobat
NRC: Web Search - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Go Links
Home
Who We Are
What We Do
Nuclear Reactors
Nuclear Materials
Radioactive Waste
Public Involvement
Electronic Reading Room
Done
Internet
Start
14:26

Convera Confidential

Agenda

- About CONVERA
- Introduction
- CONVERA Platform Infrastructure
- CONVERA‘s Retrieval Technologies
- Visualisation
- Automatic Categorization Technologies
- Dynamic Classification Technologies
- What we do / What‘s Third Party
- Product demonstration

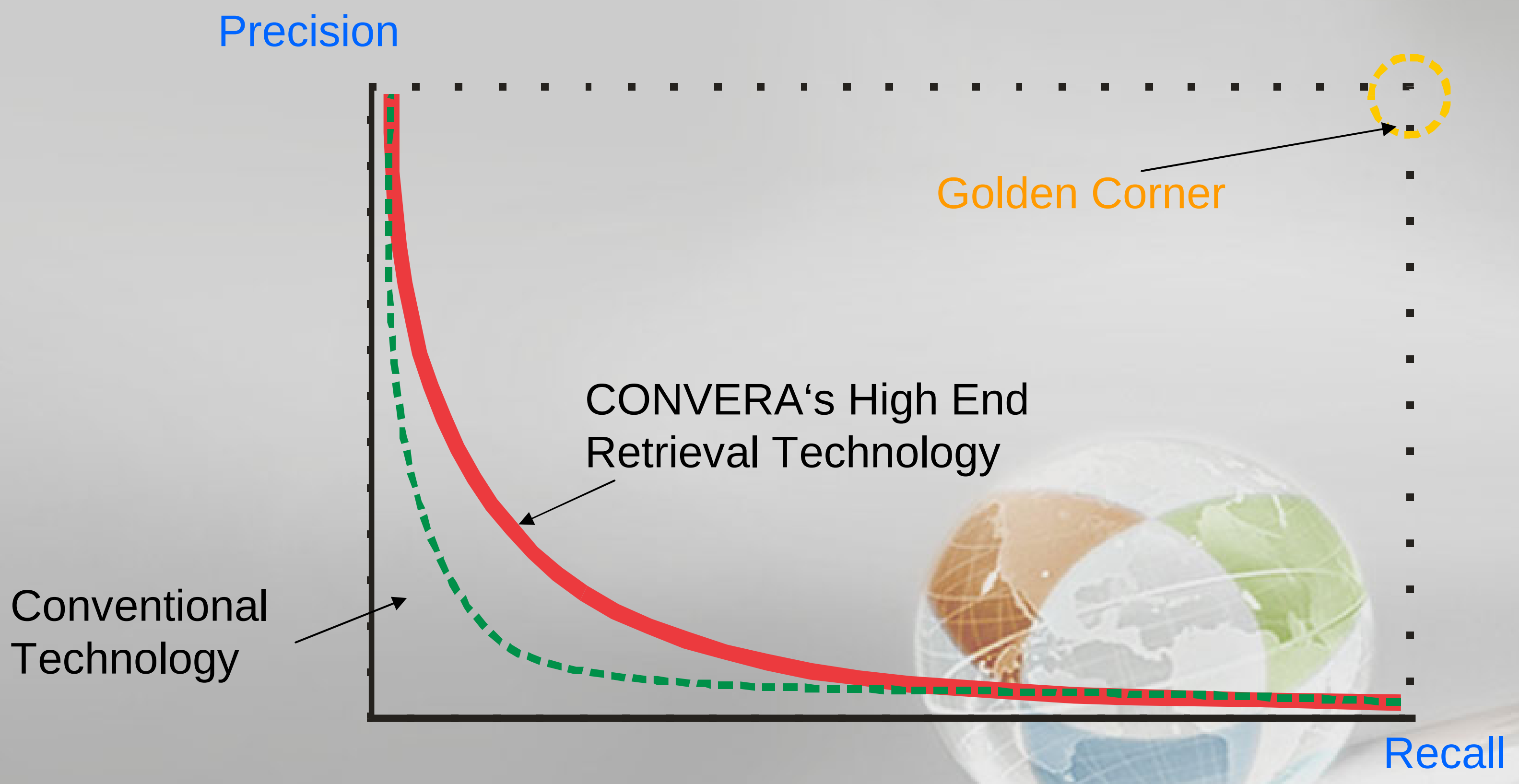
Information Science Concepts – Relevance Ranking



Relevance Ranking: Why CONVERA beats!

1. Completness
2. Proximity
3. Hit Density
4. Semantic Distance
5. Contextual Evidence

Information Science Concepts – Relevance Ranking



Convera Confidential

Information Retrieval

1. We know what we know

   Example: Today‘s Weather

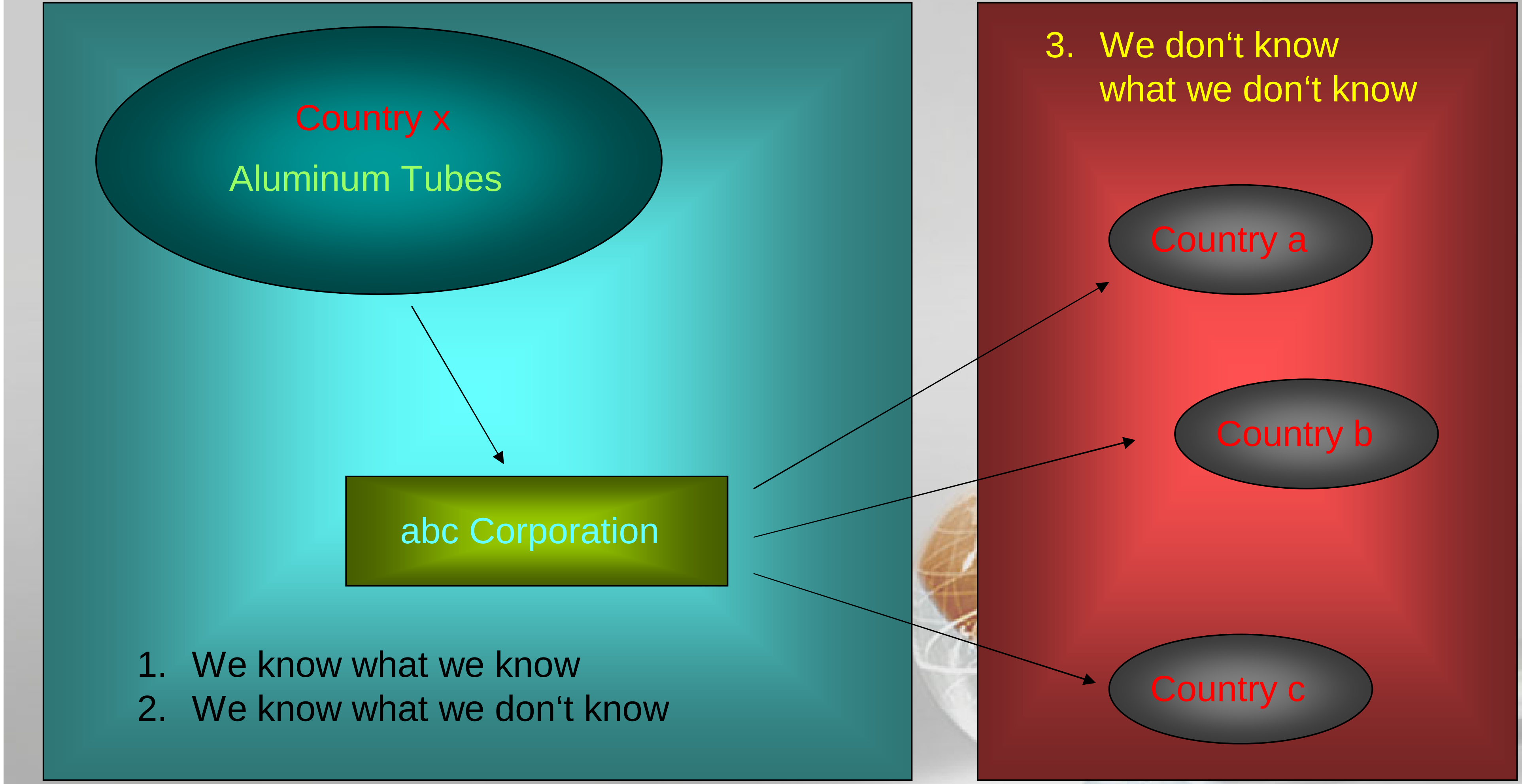
2. We know what we don‘t know

   Example: Winner of 1988 Oscars

3. We don‘t know what we don‘t know

   Example: Discovering unknow user of dual purpose Technology

Information Retrieval



Country x

Aluminum Tubes

abc Corporation

1. We know what we know
2. We know what we don‘t know

3. We don‘t know what we don‘t know

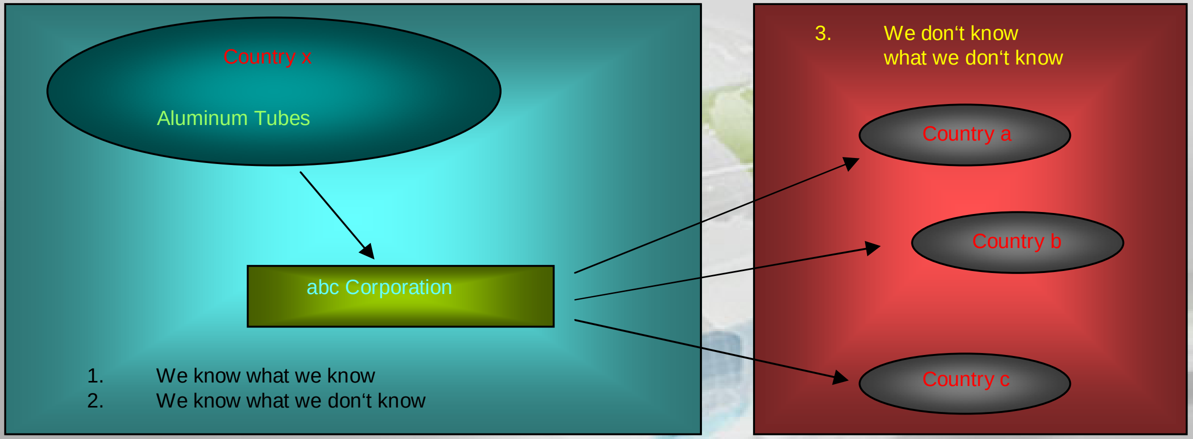
Country a

Country b

Country c

Information Retrieval – How to do that?

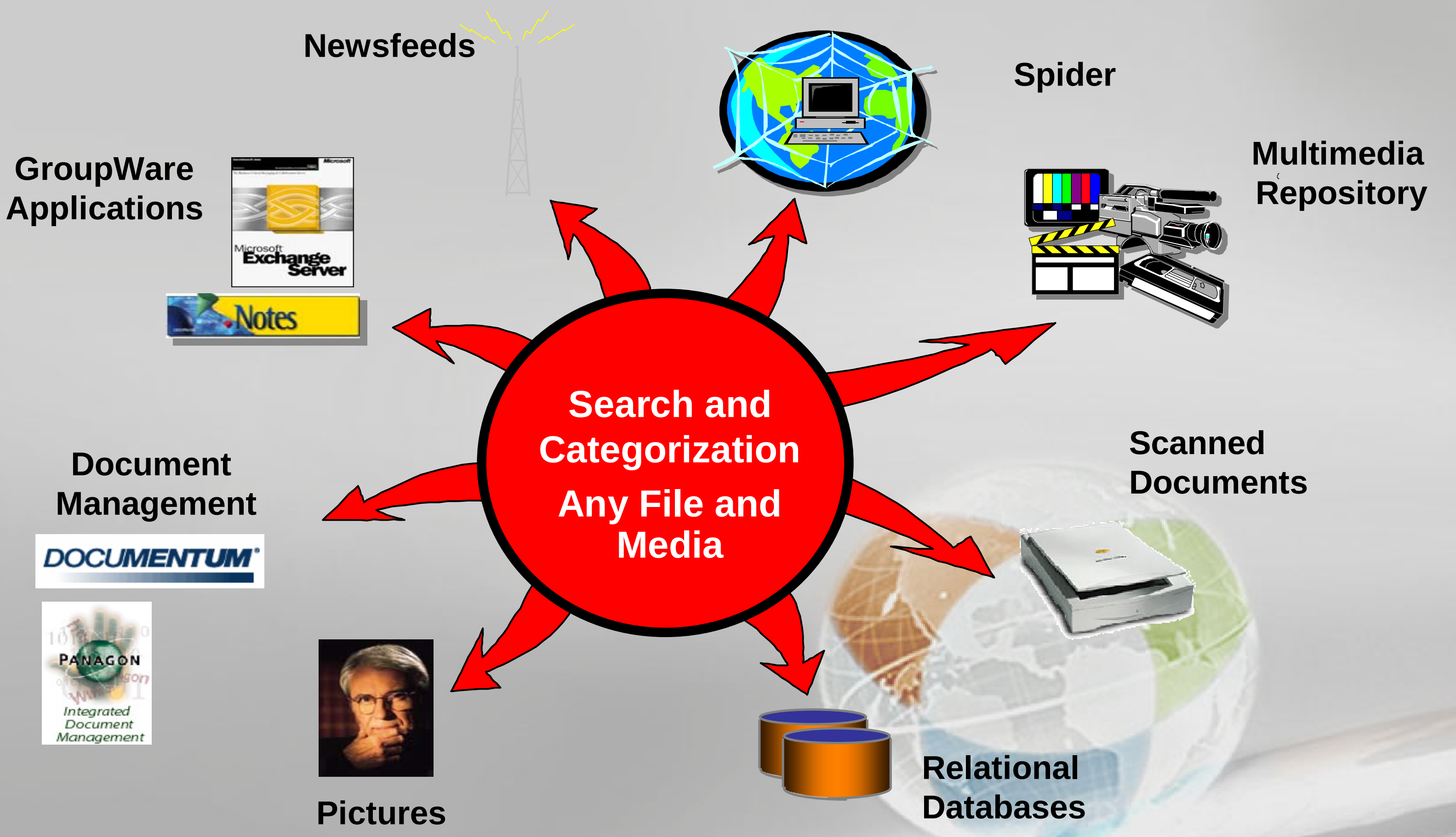
1. Same as (Relevance Feedback)
2. Fuzzy Searching (Different Spelling, OCR Errors, ...)
3. Sematic Networks (Diff Corp. Names, Abreviations,
4. Taxonomies (Automatic Indexing)
5. Dynamic Classification (Directory is the Result – not the starting point)



Agenda

- About CONVERA
- Introduction
- CONVERA Platform Infrastructure
- CONVERA's Retrieval Technologies
- Visualisation
- Automatic Categorization Technologies
- Dynamic Classification Technologies
- What we do / What's Third Party
- Product demonstration

Platform Infrastructure



Newsfeeds

Spider

GroupWare Applications

Multimedia Repository

Search and Categorization

Any File and Media

Document Management

Scanned Documents

Pictures

Relational Databases

Secure, standards-based integration for convenient, single-point of access.

Agenda

- About CONVERA
- Introduction
- CONVERA Platform Infrastructure
- CONVERA‘s Retrieval Technologies
- Visualisation
- Automatic Categorization Technologies
- Dynamic Classification Technologies
- What we do / What‘s Third Party
- Product demonstration

Technologies

- APRP – **Pattern recognition**
- List Search Server
- Boolean
- Concept
- Language Modules
- Multi Language Modules
- Cross Language Modules
- Information Profiling
- Automatic Categorization
- Dynamic Personal Classification

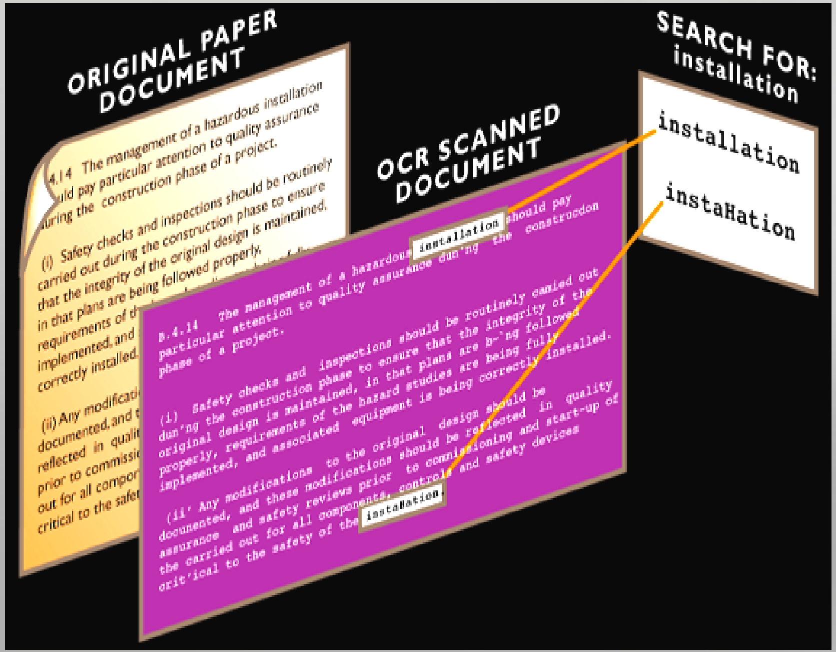
APRP solves that Problem (Mu`ammar al-Qadhafi) better

1) Muammar Qaddafi
2) Mo'ammar Gadhafi
3) Muammar Kaddafi
4) Muammar Qadhafi
5) Moammar El Kadhafi
6) Muammar Gadafi
7) Mu'ammar al-Qadafi
8) Moamer El Kazzafi
9) Moamar al-Gaddafi
10) Mu'ammar Al Qathafi
11) Muammar Al Qathafi
12) Mo'ammar el-Gadhafi
13) Moamar El Kadhafi
14) Muammar al-Qadhafi
15) Mu'ammar al-Qadhdhafi
16) Mu'ammar Qadafi
17) Moamar Gaddafi
18) Mu'ammar Qadhdhafi
19) Muammar Khaddafi
20) Muammar al-Khaddafi
21) Mu'amar al-Kadafi
22) Muammar Ghaddafy
23) Muammar Ghadafi
24) Muammar Ghaddafi
25) Muamar Kaddafi
26) Muammar Quathafi
27) Mohammer Q'udafi
28) Muammar Gheddafi
29) Muamar Al-Kaddafi
30) Moammar Khadafy
31) Moammar Qudhafi
32) Mu'ammar al-Qaddafi

APRP: Why more Accuracy then standard Fuzzy?

BOOT `01000010 01001111 01001111 01010100`

BOAT `01000010 01001111 01000001 01010100`

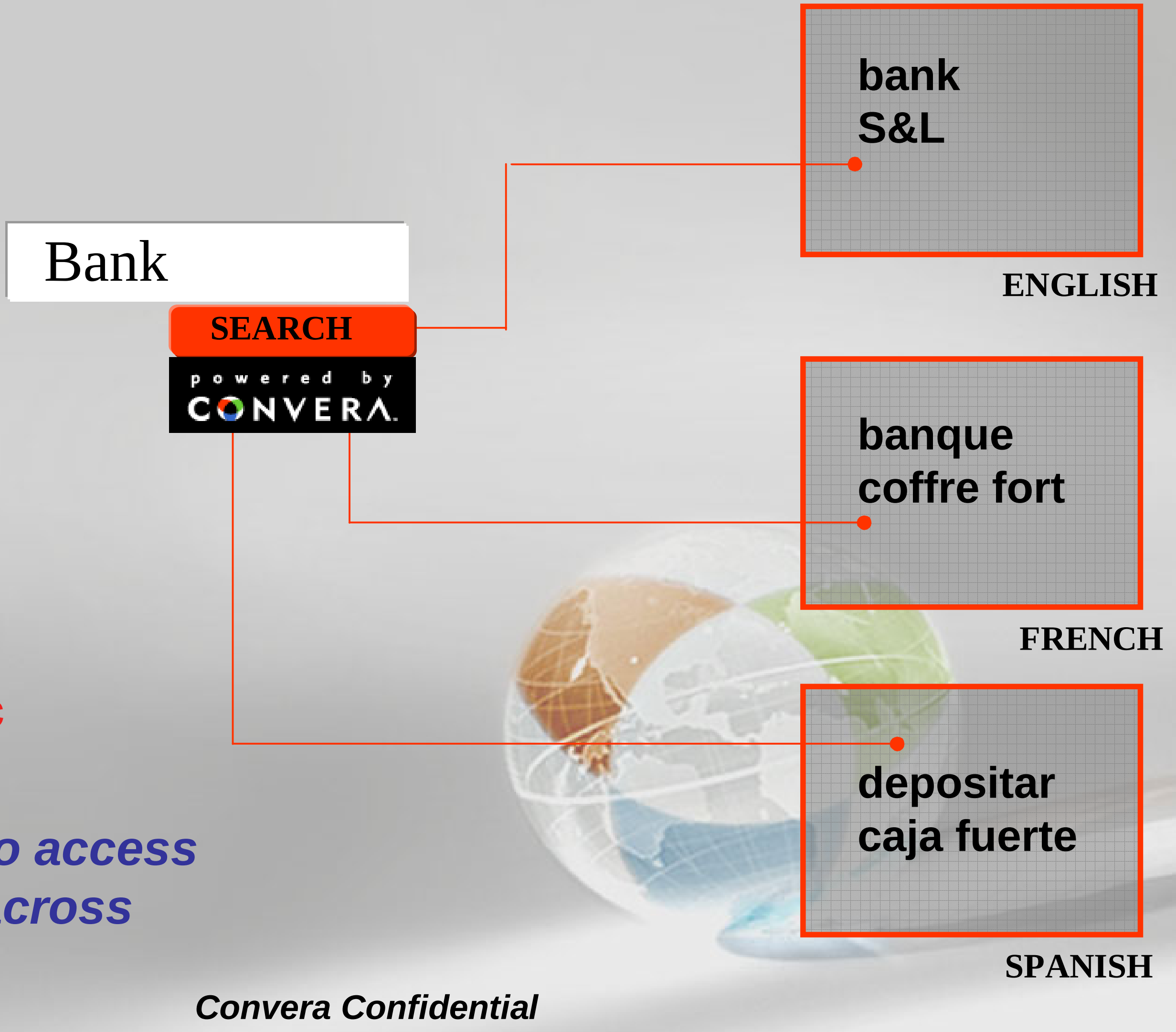


- **Overcomes errors, typos, misspellings**
- **25% inaccuracy in text is only 10% in binary**
- **APRP supports multimedia**

Sample: Names
Rayin al-Abidin Muamar Hussayn
Rayn al-Abidin Muhammar Husayn

Multi-lingual Search

- **Cross-lingual**
  - English
  - German
  - Spanish
  - French
  - Dutch
  - Italian
  - Hebrew and Arabic



Bank

SEARCH

powered by CONVERA

bank
S&L

ENGLISH

banque
coffre fort

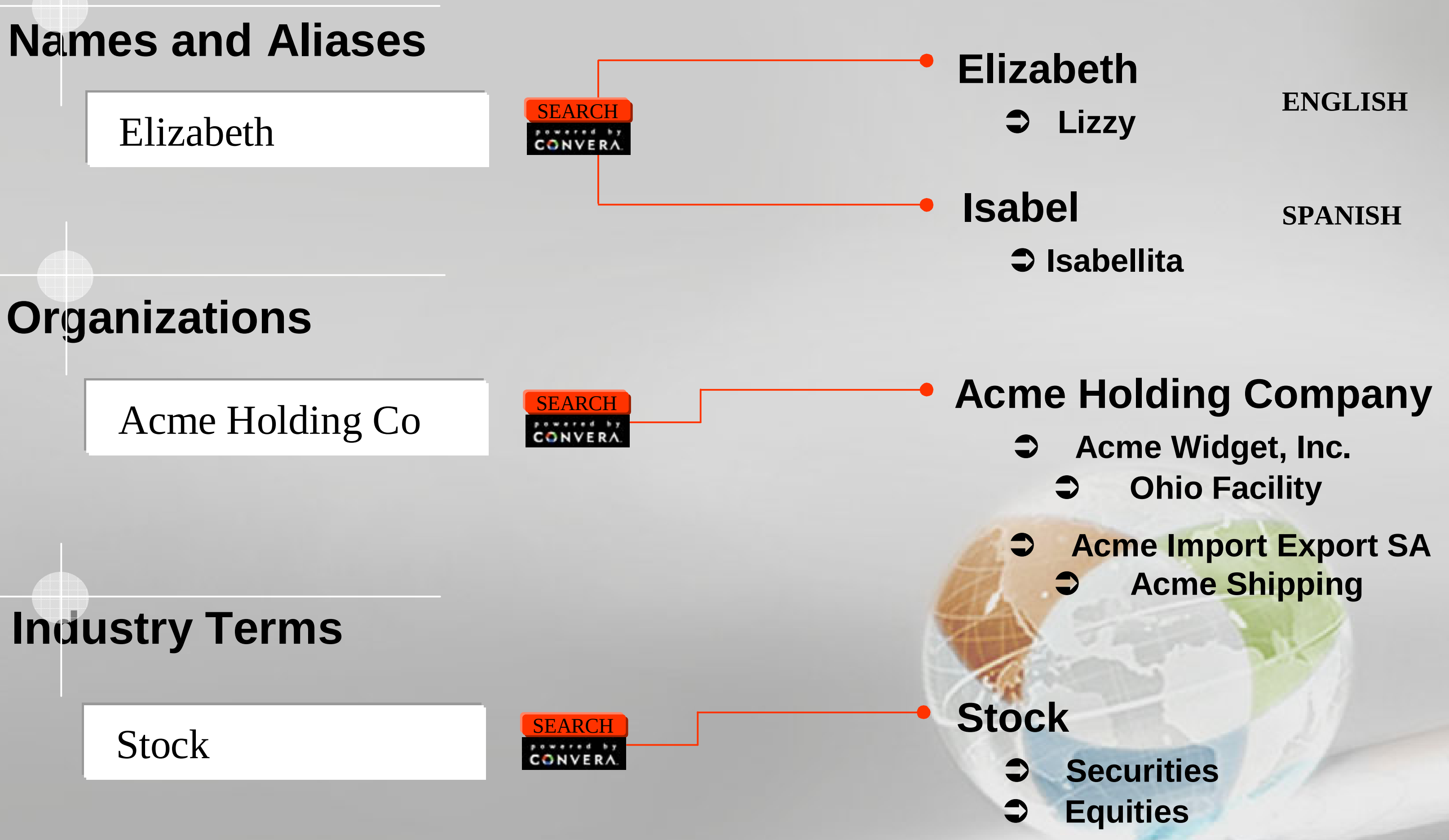
FRENCH

depositar
caja fuerte

SPANISH

Allow stakeholders to access and view results across languages

Applications of Concept Search



Names and Aliases

Elizabeth

SEARCH powered by CONVERA

- Elizabeth — ENGLISH
  - ➲ Lizzy
- Isabel — SPANISH
  - ➲ Isabellita

Organizations

Acme Holding Co

SEARCH powered by CONVERA

- Acme Holding Company
  - ➲ Acme Widget, Inc.
    - ➲ Ohio Facility
  - ➲ Acme Import Export SA
    - ➲ Acme Shipping

Industry Terms

Stock

SEARCH powered by CONVERA

- Stock
  - ➲ Securities
  - ➲ Equities

Provide stakeholders with ‘virtual expertise’ for more accurate search

Search Functionality – In the BODY TEXT and in the FIELDS!

- Query by Example (relevance feedback)
- Idiom (Syntactic) Processing
- Adjustable Stop Words
- Exact Phrases
- Date Ranges
- Fielded Searching
- Search Term Weighting
- Logging functions (user)
- Web crawler

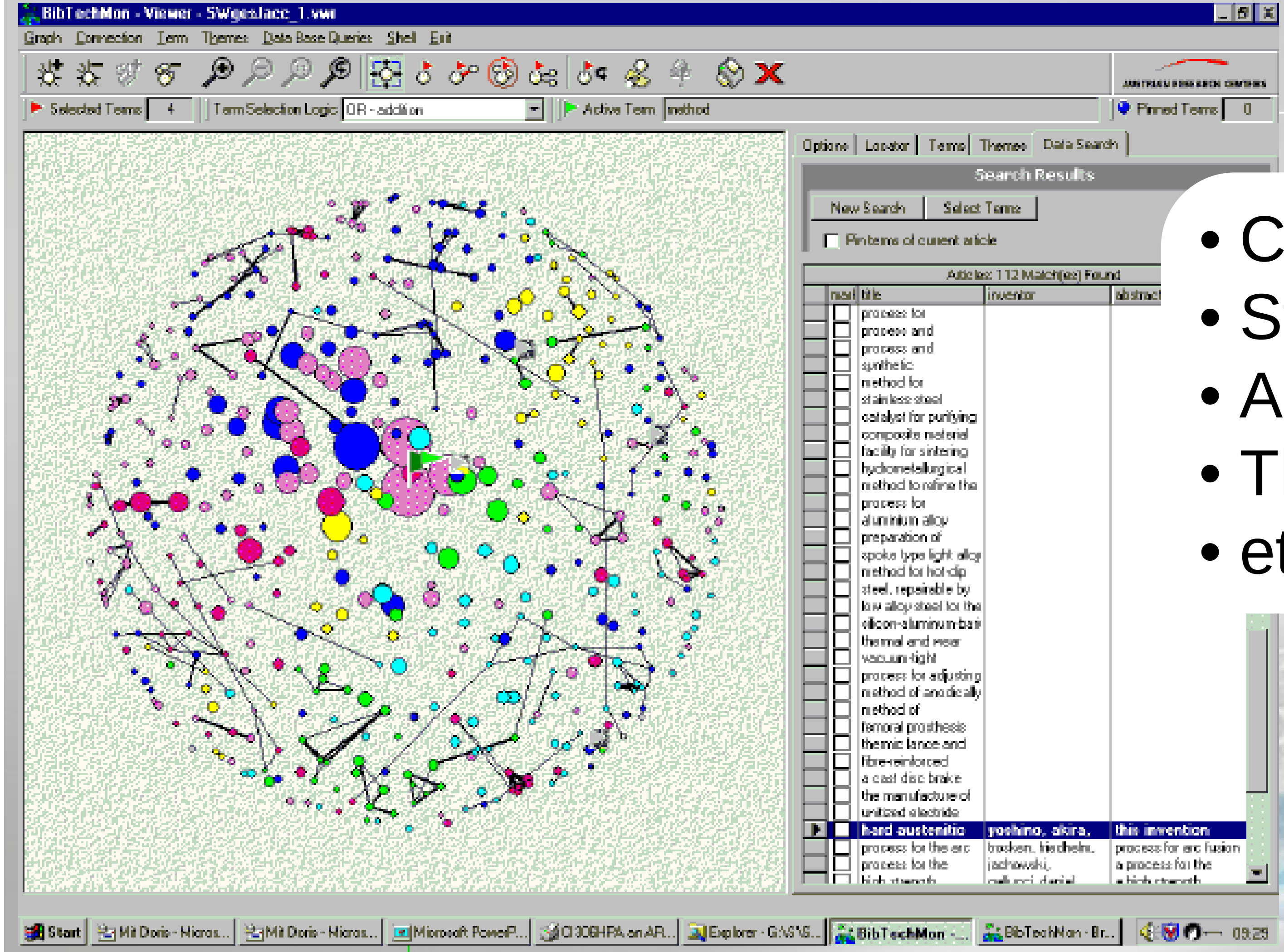
- Numeric Range searching
- Multiple Dictionaries / Thesauri
- Recurrent searching (searching hitlist)
- Multiple options for document display
- Automatic categorization
- Dynamic classification
- User profiling
- Relevance Ranking
- Language / Industry Plug-ins
- Cluster displays
- Visualisation aids

Scalability Accuracy *Flexibility*

Third Patry Knowledge Discovery Components

Third Party	Visualization	Tracking	Pattern Detection
CONVERA	Search	Classification	Mapping
	Indexing	Categorization	Entity Extraction
	Crawling, Filtering		

Postprocessing - Clustering



Selected Terms 4

Term Selection Logic: OR - addition

Active Term: method

Pinned Terms 0

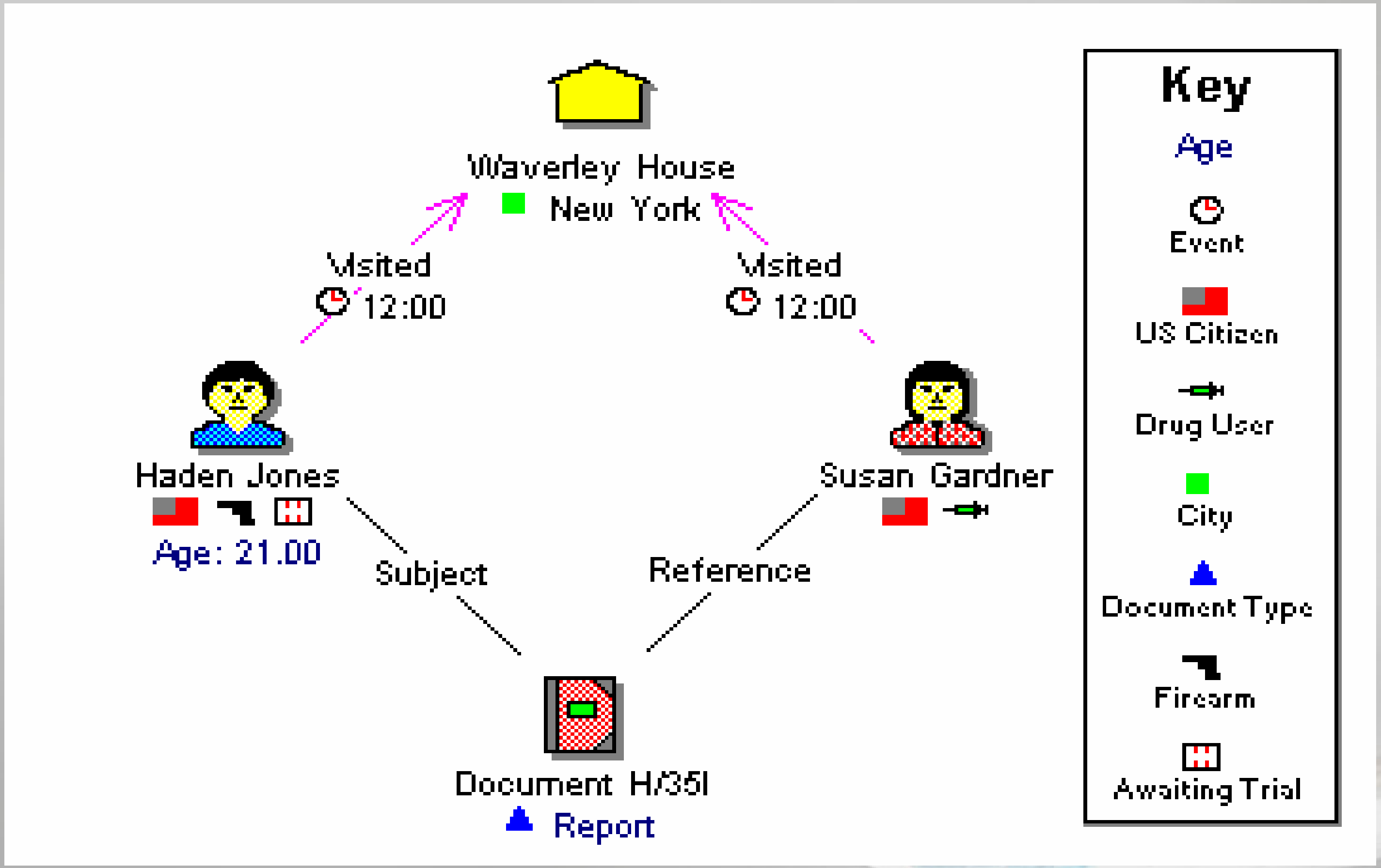
☐ Pin terms of current article

Articles: 112 Match(es) Found

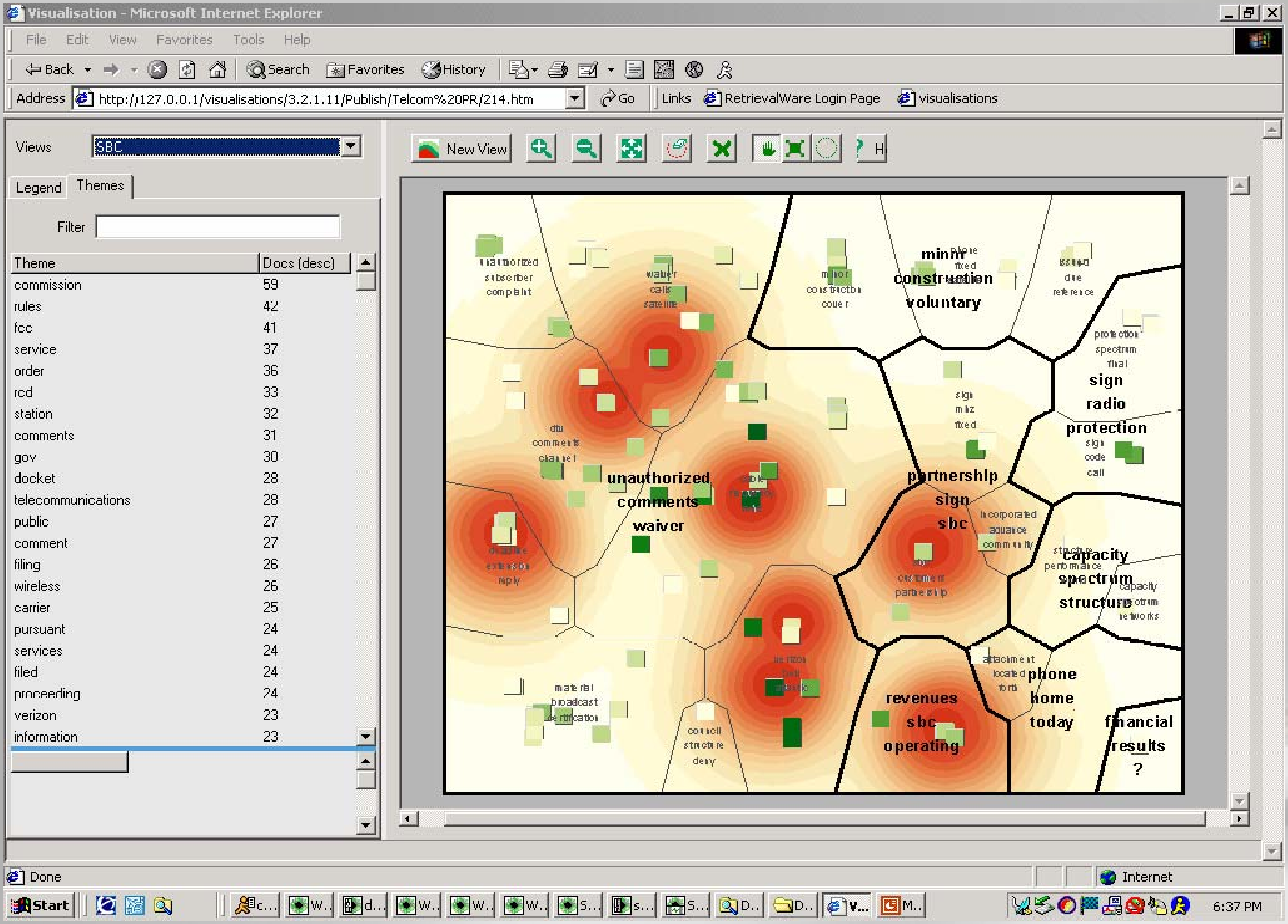
read	title	inventor	abstract
☐	process for		
☐	process and		
☐	process and		
☐	synthetic		
☐	method for		
☐	stainless steel		
☐	catalyst for purifying		
☐	composite material		
☐	facility for sintering		
☐	hydrometallurgical		
☐	method to refine the		
☐	process for		
☐	aluminium alloy		
☐	preparation of		
☐	spoke type light alloy		
☐	method for hot-dip		
☐	steel, repairable by		
☐	low alloy steel for the		
☐	silicon-aluminum-bar		
☐	thermal and wear		
☐	vacuum-tight		
☐	process for adjusting		
☐	method of anodically		
☐	method of		
☐	femoral prosthesis		
☐	thermic lance and		
☐	fibre-reinforced		
☐	a cast disc brake		
☐	the manufacture of		
☐	unitized electrode		
☐	hard austenitic	yoshino, akira,	this invention
☐	process for the arc	brocken, hiedhelm,	process for arc fusion
☐	process for the	jachowski,	a process for the
☐	high strength	pellucci, daniel	a high strength

- Clustering
- Statistical Analysis
- Absolute Linking
- Theme Grouping
- etc …

www.i2.co.uk - Tracking



www.cedar.com – Pattern detection



Visualisation - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://127.0.0.1/visualisations/3.2.1.11/Publish/Telcom%20PR/214.htm
Go Links RetrievalWare Login Page visualisations
Views SBC
Legend Themes
Filter
New View
unauthorized subscriber complaint
waiver calls satellite
minor construction cover
minor construction voluntary
issued due reference
protection spectrum final
sign radio protection
sign code call
sign mhz fixed
dtv comments channel
unauthorized comments waiver
partnership sign sbc
incorporated advance community
capacity spectrum structure
disable extension reply
material broadcast certification
council structure deny
revenues sbc operating
attachment located form
phone home today
financial results ?
Done
Internet
Start
6:37 PM

Theme	Docs (desc)
commission	59
rules	42
fcc	41
service	37
order	36
rcd	33
station	32
comments	31
gov	30
docket	28
telecommunications	28
public	27
comment	27
filing	26
wireless	26
carrier	25
pursuant	24
services	24
filed	24
proceeding	24
verizon	23
information	23

Agenda

- About CONVERA
- Introduction
- CONVERA Platform Infrastructure
- CONVERA‘s Retrieval Technologies
- Visualisation
- Automatic Categorization Technologies
- Dynamic Classification Technologies
- What we do / What‘s Third Party
- Product demonstration

The Problem

- Manual indexing is costly and slow
- Traditional Classification is precoordinated
- Hit lists are OK, but somehow inefficient
- Most information is unstructured
- Information structure is irrelevant

Ontology

- An **ontology** is a foundation of categories representing a view of the world. An ontology reflects the commonly used and trusted breakdown of categories. For example, the breakdown of news items into categories of ‘World’, ‘Sports’, ‘Politics’, etc. is ontological.

Taxonomy

- A **taxonomy** is a hierarchical system describing genera and species. Species derive from a common genus and are hierarchically represented according to their essential characteristics and differences. For example, animals are categorized with the "Taxonomy of Life" which separates mammals from birds and spiders from insects, based on proper features and relative differences. This genus to species nomenclature is highlighted by terminology which moves from generic terms to binomial terms through lexical derivation and compounding.

- A taxonomy doesn’t deal with things, but with the essence of things: a taxonomy is based on an ontology.

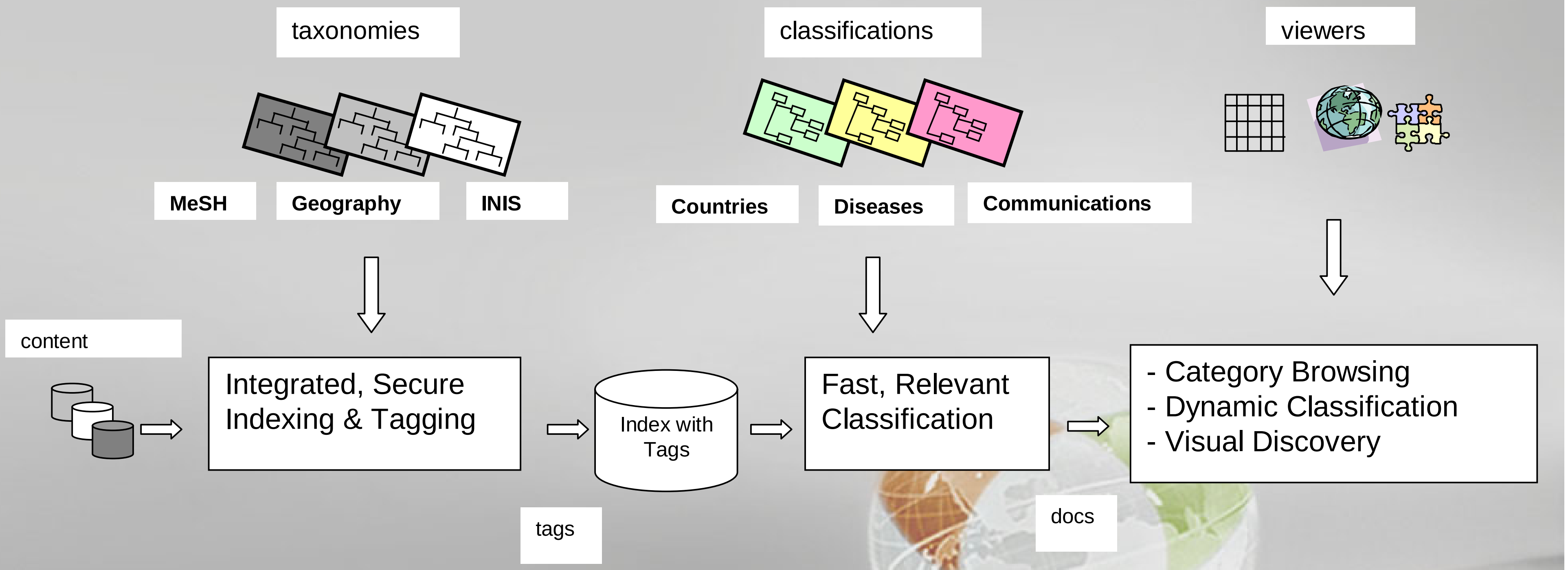
Categorization vs. Classification

- Categorization
  - **Logical**
  - **Taxonomy based**
  - **Consistent (Based on cultural fundamentals)**
  - **Stable**

- Classification
  - **Pragmatic**
  - **Precordinated**
  - **Common sense**
  - **Chaotic (Based on best practices)**
  - **Individual**

Consistent, Scalable, and Flexible Knowledge Architecture

We don‘t know the classification from tomorrows needs!



taxonomies

MeSH

Geography

INIS

classifications

Countries

Diseases

Communications

viewers

content

Integrated, Secure Indexing & Tagging

Index with Tags

tags

Fast, Relevant Classification

docs

- Category Browsing
- Dynamic Classification
- Visual Discovery

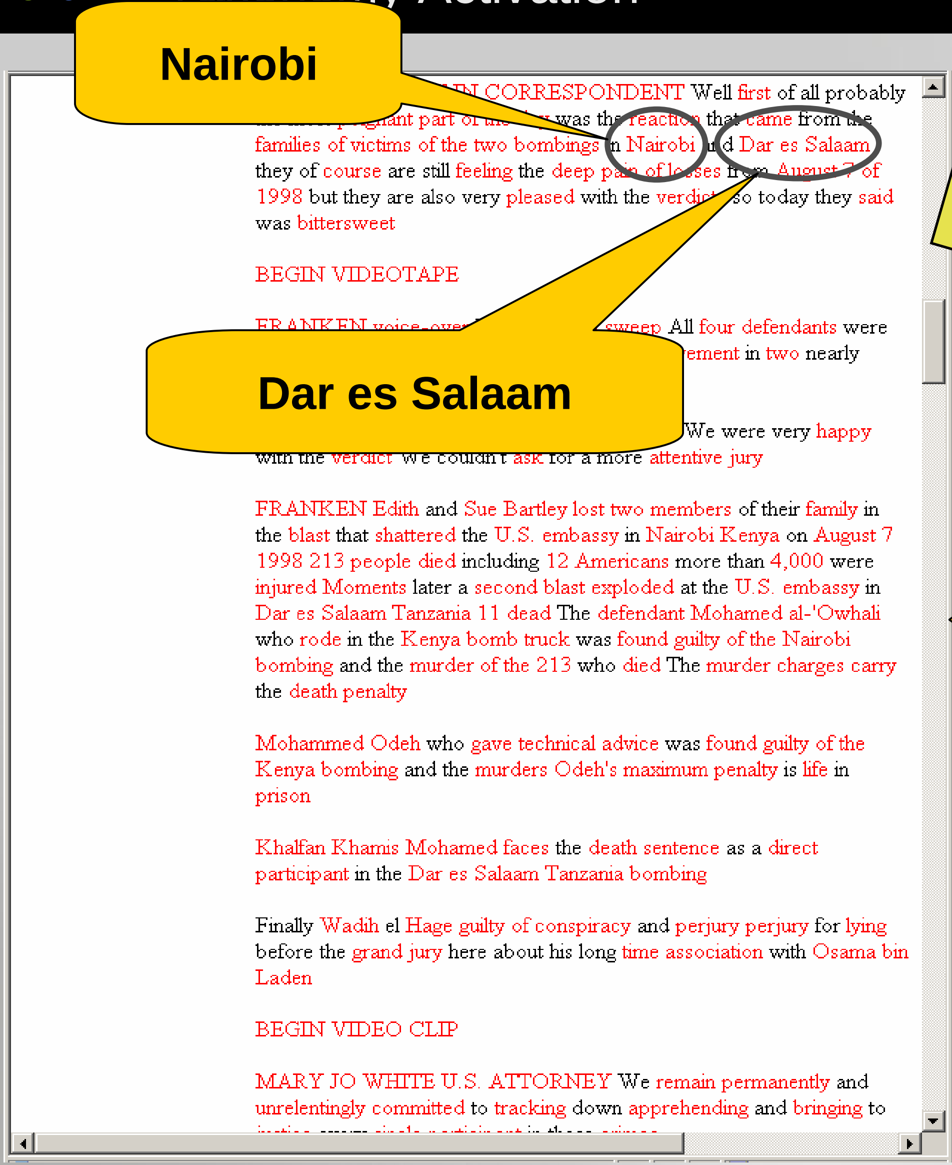
Example 1: Geography

- Africa
  - Algeria
  - Angola
- Asia
  - Afghanistan
  - Armenia
- Europe
  - Albania
  - Andorra
- Middle East
  - Bahrain
  - Iran
- North and Central America
  - Antigua and Barbuda
  - Bahamas
- Pacific
  - Australia
  - Fiji
- South America
  - Argentina
  - Bolivia
- U.S.
  - Alabama
  - Alaska

Example 2 : Defense

- Defense Communications
  - Satellite Communications
  - Tactical Communications
- Defense Systems
  - Air Defense
    - Antiaircraft Defense Systems
      - Gun Air Defense Systems
    - Antimissile Defense Systems
    - Forward Area Air Defense Systems
    - Terminal Defense
  - Aircraft Defense Systems
  - Antisubmarine Defense Systems
  - Antiswimmer Defense Systems
  - Countermeasures
    - Acoustic Countermeasures

Taxonomy Activation



...N CORRESPONDENT Well first of all probably ...gnant part of the ... was the reaction that came from the families of victims of the two bombings in Nairobi and Dar es Salaam they of course are still feeling the deep pain of losses from August 7 of 1998 but they are also very pleased with the verdict so today they said was bittersweet

BEGIN VIDEOTAPE

FRANKEN voice-over ... sweep All four defendants were ...ement in two nearly

...We were very happy with the verdict We couldn't ask for a more attentive jury

FRANKEN Edith and Sue Bartley lost two members of their family in the blast that shattered the U.S. embassy in Nairobi Kenya on August 7 1998 213 people died including 12 Americans more than 4,000 were injured Moments later a second blast exploded at the U.S. embassy in Dar es Salaam Tanzania 11 dead The defendant Mohamed al-'Owhali who rode in the Kenya bomb truck was found guilty of the Nairobi bombing and the murder of the 213 who died The murder charges carry the death penalty

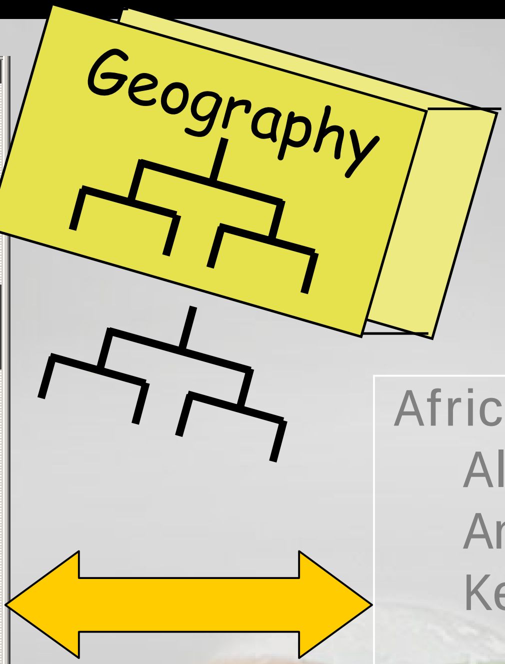
Mohammed Odeh who gave technical advice was found guilty of the Kenya bombing and the murders Odeh's maximum penalty is life in prison

Khalfan Khamis Mohamed faces the death sentence as a direct participant in the Dar es Salaam Tanzania bombing

Finally Wadih el Hage guilty of conspiracy and perjury perjury for lying before the grand jury here about his long time association with Osama bin Laden

BEGIN VIDEO CLIP

MARY JO WHITE U.S. ATTORNEY We remain permanently and unrelentingly committed to tracking down apprehending and bringing to justice every single participant in these crimes



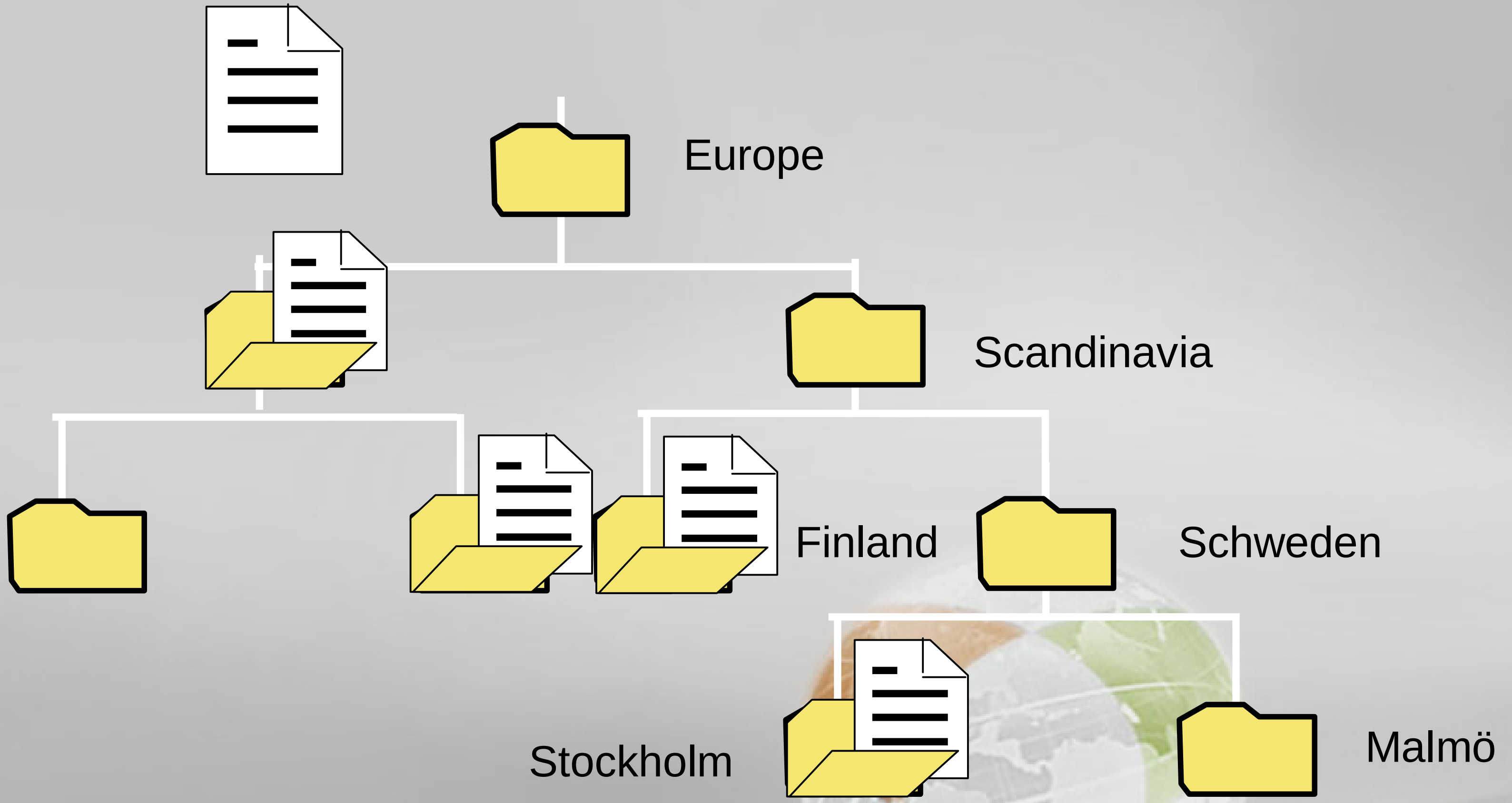
Africa

- Algeria
- Angola
- Kenya
  - **Nairobi**
- Tanzania
  - **Dar es Salaam**

Asia

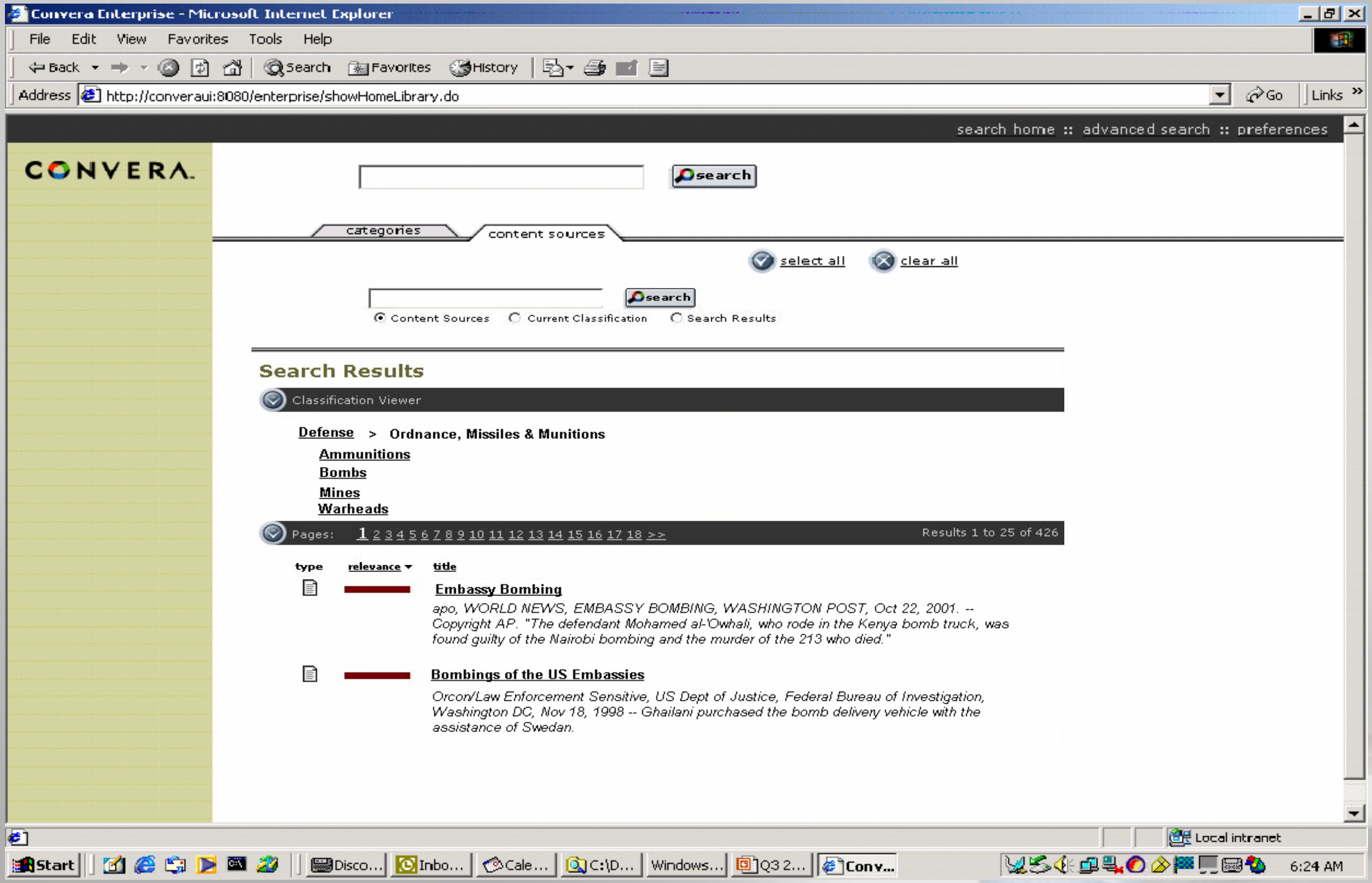
- Afghanistan
- Armenia

Population Mechanism



Convera Confidential

Automatic Categorizing – Personal Dynamic Classification


Convera Enterprise - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://converaui:8080/enterprise/showHomeLibrary.do
Go Links
search home :: advanced search :: preferences
search
categories
content sources
select all
clear all
search
○ Content Sources ○ Current Classification ○ Search Results
Local intranet
Start Disco... Inbo... Cale... C:\D... Windows... Q3 2... Conv...
6:24 AM

Search Results

Classification Viewer

Defense > Ordnance, Missiles & Munitions

- Ammunitions
- Bombs
- Mines
- Warheads

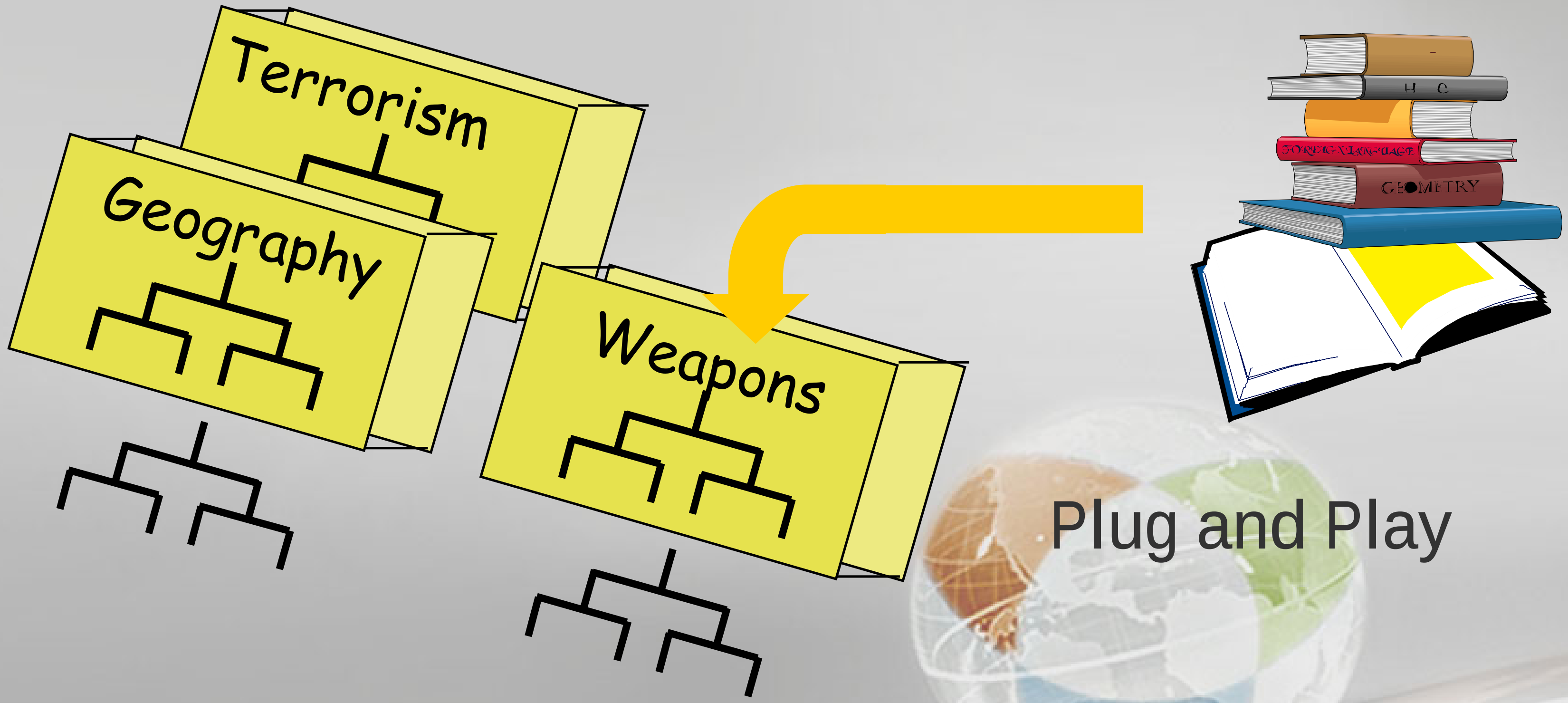
Pages: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 >> Results 1 to 25 of 426

type relevance title

Embassy Bombing
apo, WORLD NEWS, EMBASSY BOMBING, WASHINGTON POST, Oct 22, 2001. -- Copyright AP. "The defendant Mohamed al-'Owhali, who rode in the Kenya bomb truck, was found guilty of the Nairobi bombing and the murder of the 213 who died."

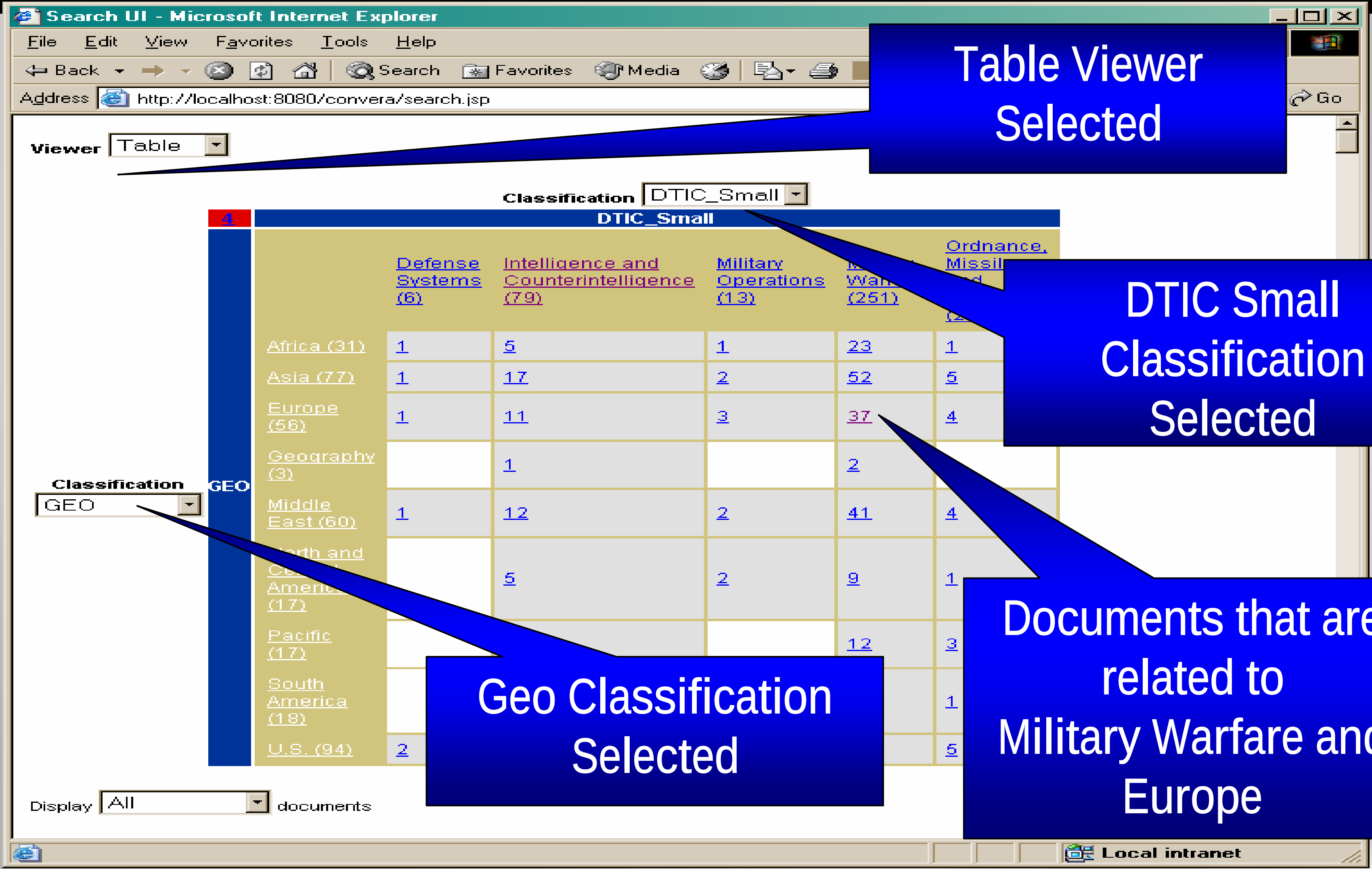
Bombings of the US Embassies
Orcon/Law Enforcement Sensitive, US Dept of Justice, Federal Bureau of Investigation, Washington DC, Nov 18, 1998 -- Ghailani purchased the bomb delivery vehicle with the assistance of Swedan.

Vertical Cartridges



Plug and Play

Visual Discovery



Search UI - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://localhost:8080/convera/search.jsp
Go
Local intranet

Viewer: Table

Classification: DTIC_Small

Classification: GEO

4	DTIC_Small					
GEO		Defense Systems (6)	Intelligence and Counterintelligence (79)	Military Operations (13)	[illegible] Wa[illegible] (251)	Ordnance, Missil[illegible]
	Africa (31)	1	5	1	23	1
	Asia (77)	1	17	2	52	5
	Europe (56)	1	11	3	37	4
	Geography (3)		1		2	
	Middle East (60)	1	12	2	41	4
	[illegible]rth and Ce[illegible] Americ[illegible] (17)		5	2	9	1
	Pacific (17)		[illegible]		12	3
	South America (18)		[illegible]	[illegible]	[illegible]	1
	U.S. (94)	2	[illegible]	[illegible]	[illegible]	5

Display: All documents

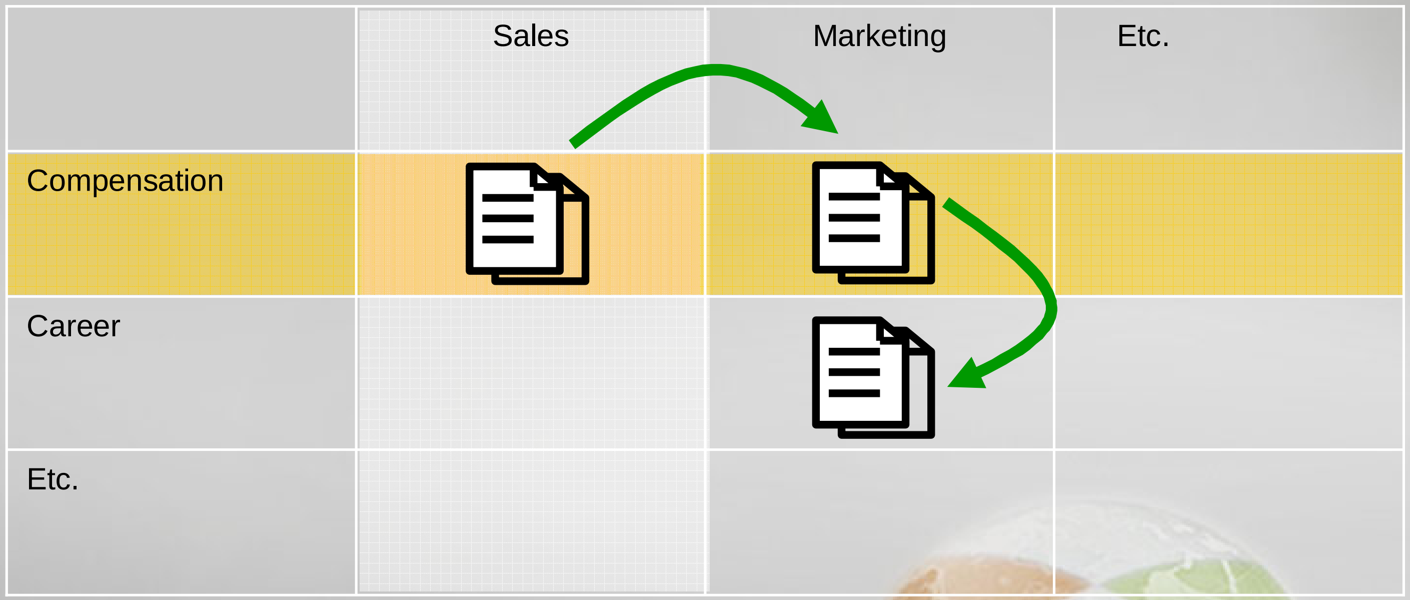
Table Viewer Selected

DTIC Small Classification Selected

Documents that are related to Military Warfare and Europe

Geo Classification Selected

Expand Your Search



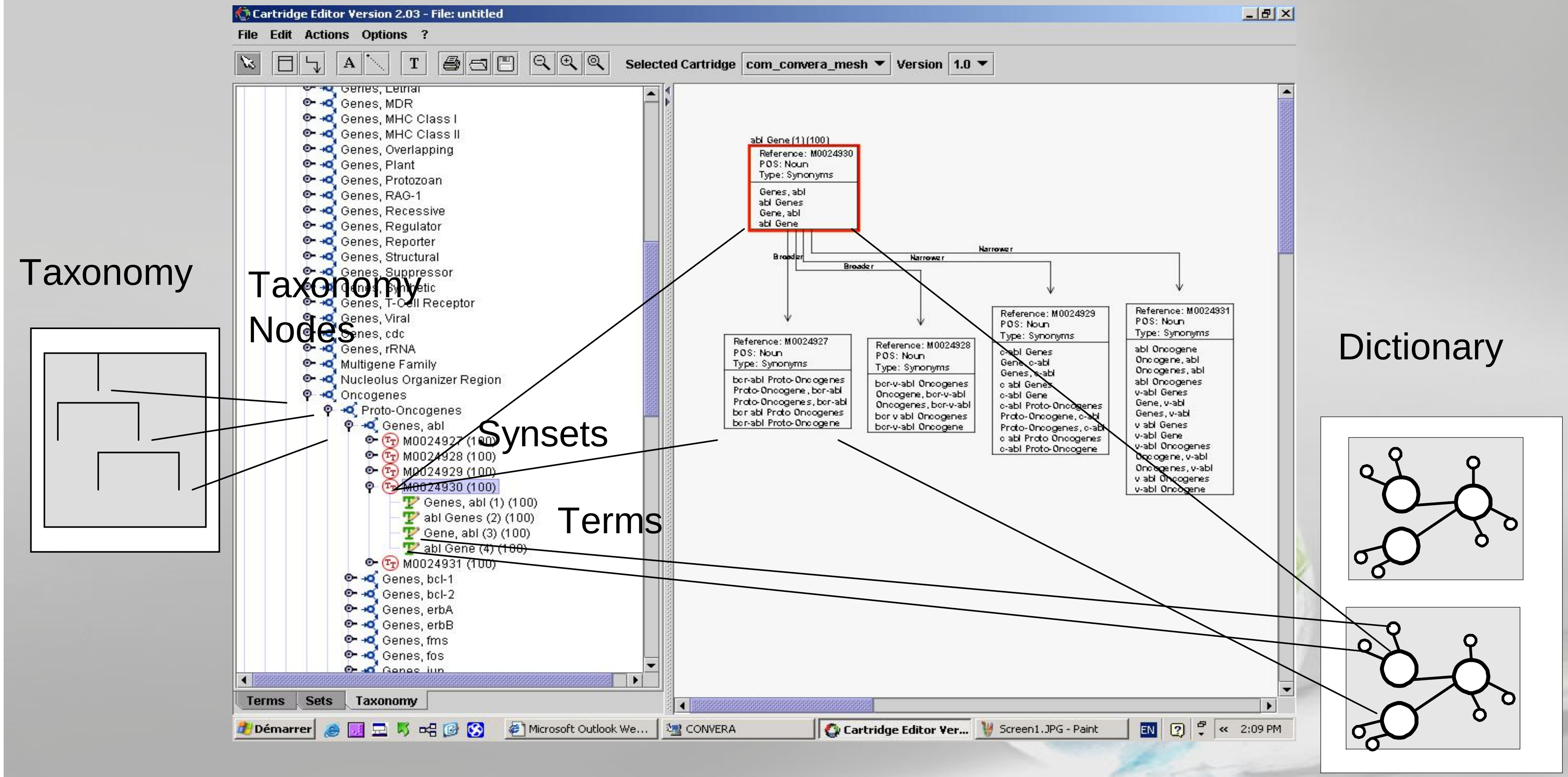
	Sales	Marketing	Etc.
Compensation	📄	📄	
Career		📄	
Etc.			

- Search is a journey through a multi-dimensional grid of topics
- The ability to visualize all possible combinations at once will save time and increase focus

Already Existing Taxonomy Cartridges (Samples)

- Biology
- Chemistry
- Computers
- Electronics
- Finance
- Food Science
- Geography
- Geology
- Health Sciences
- Information Science
- Law
- Mathematics
- MeSH (Medical Subject Headings)
- Military
- Petroleum Natural Gas & Petrochemicals
- Pharmacology
- Physics
- Plastics
- Rubber
- Telecommunications

Cartridge Editor enables editing Taxonomy and Dictionary



Genes, Lethal
Genes, MDR
Genes, MHC Class I
Genes, MHC Class II
Genes, Overlapping
Genes, Plant
Genes, Protozoan
Genes, RAG-1
Genes, Recessive
Genes, Regulator
Genes, Reporter
Genes, Structural
Genes, Suppressor
Genes, Synthetic
Genes, T-Cell Receptor
Genes, Viral
Genes, cdc
Genes, rRNA
Multigene Family
Nucleolus Organizer Region
Oncogenes
Proto-Oncogenes
Genes, abl
M0024927 (100)
M0024928 (100)
M0024929 (100)
M0024930 (100)
Genes, abl (1) (100)
abl Genes (2) (100)
Gene, abl (3) (100)
abl Gene (4) (100)
M0024931 (100)
Genes, bcl-1
Genes, bcl-2
Genes, erbA
Genes, erbB
Genes, fms
Genes, fos
Genes, jun

abl Gene (1) (100)

Reference: M0024930
POS: Noun
Type: Synonyms
Genes, abl
abl Genes
Gene, abl
abl Gene

Broader
Broader
Narrower
Narrower

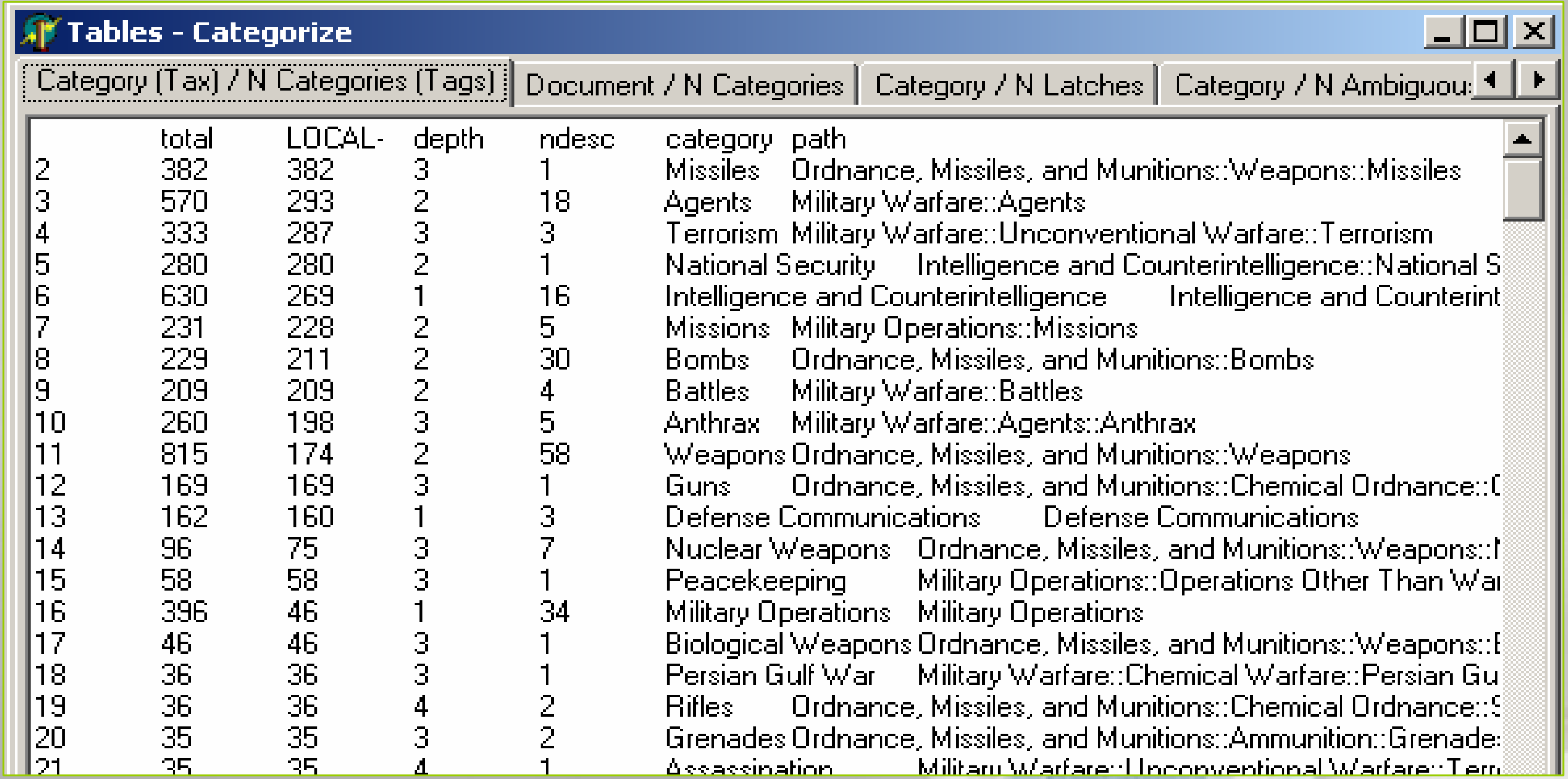
Reference: M0024927
POS: Noun
Type: Synonyms
bcr-abl Proto-Oncogenes
Proto-Oncogene, bcr-abl
Proto-Oncogenes, bcr-abl
bcr abl Proto Oncogenes
bcr-abl Proto-Oncogene

Reference: M0024928
POS: Noun
Type: Synonyms
bcr-v-abl Oncogenes
Oncogene, bcr-v-abl
Oncogenes, bcr-v-abl
bcr v abl Oncogenes
bcr-v-abl Oncogene

Reference: M0024929
POS: Noun
Type: Synonyms
c-abl Genes
Gene, c-abl
Genes, c-abl
c abl Genes
c-abl Gene
c-abl Proto-Oncogenes
Proto-Oncogene, c-abl
Proto-Oncogenes, c-abl
c abl Proto Oncogenes
c-abl Proto-Oncogene

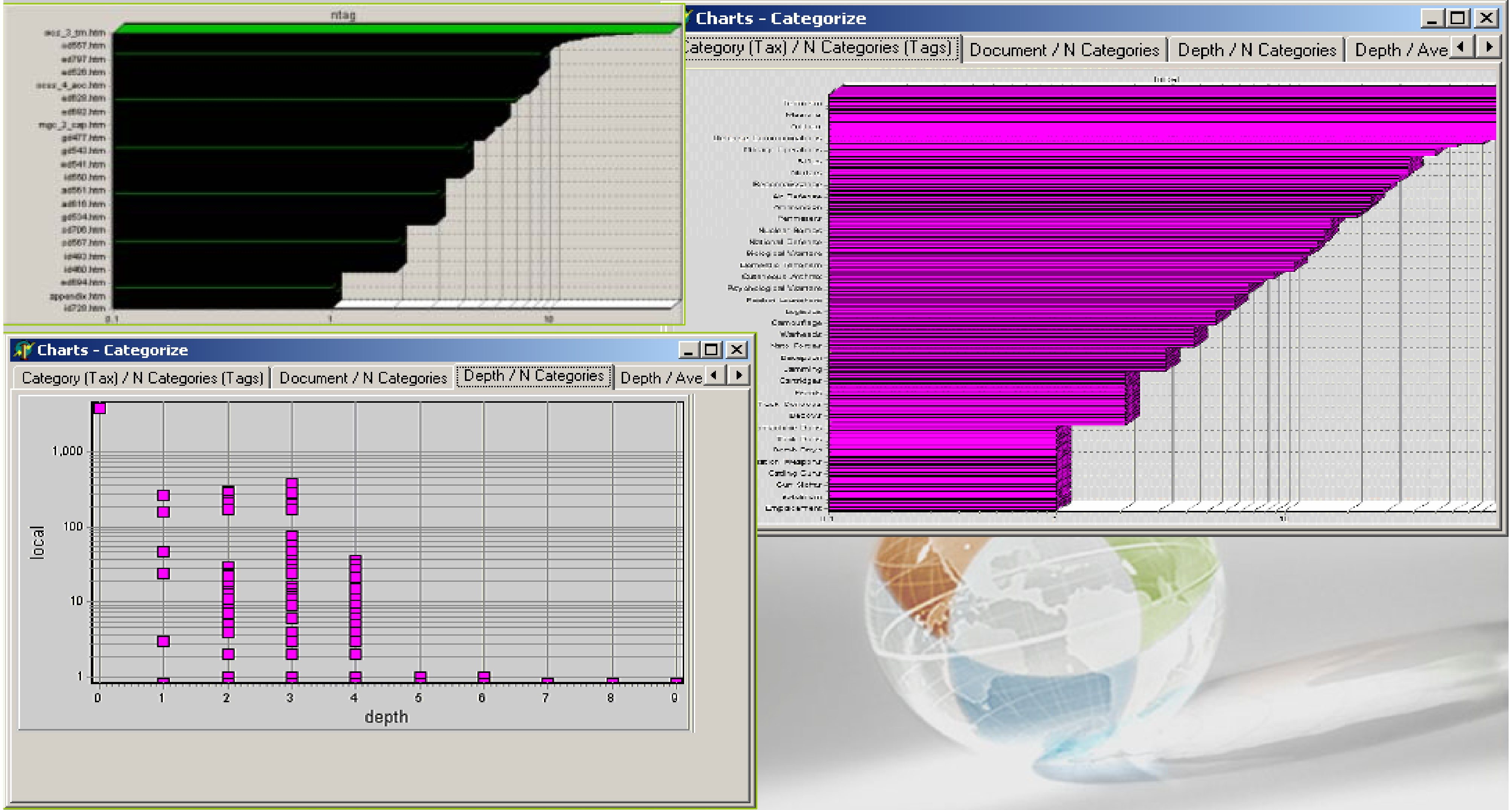
Reference: M0024931
POS: Noun
Type: Synonyms
abl Oncogene
Oncogene, abl
Oncogenes, abl
abl Oncogenes
v-abl Genes
Gene, v-abl
Genes, v-abl
v abl Genes
v-abl Gene
v-abl Oncogenes
Oncogene, v-abl
Oncogenes, v-abl
v abl Oncogenes
v-abl Oncogene

Tables to improve the Cartridge Quality



	total	LOCAL-	depth	ndesc	category	path
2	382	382	3	1	Missiles	Ordnance, Missiles, and Munitions::Weapons::Missiles
3	570	293	2	18	Agents	Military Warfare::Agents
4	333	287	3	3	Terrorism	Military Warfare::Unconventional Warfare::Terrorism
5	280	280	2	1	National Security	Intelligence and Counterintelligence::National S
6	630	269	1	16	Intelligence and Counterintelligence	Intelligence and Counterint
7	231	228	2	5	Missions	Military Operations::Missions
8	229	211	2	30	Bombs	Ordnance, Missiles, and Munitions::Bombs
9	209	209	2	4	Battles	Military Warfare::Battles
10	260	198	3	5	Anthrax	Military Warfare::Agents::Anthrax
11	815	174	2	58	Weapons	Ordnance, Missiles, and Munitions::Weapons
12	169	169	3	1	Guns	Ordnance, Missiles, and Munitions::Chemical Ordnance::
13	162	160	1	3	Defense Communications	Defense Communications
14	96	75	3	7	Nuclear Weapons	Ordnance, Missiles, and Munitions::Weapons::
15	58	58	3	1	Peacekeeping	Military Operations::Operations Other Than Wa
16	396	46	1	34	Military Operations	Military Operations
17	46	46	3	1	Biological Weapons	Ordnance, Missiles, and Munitions::Weapons::
18	36	36	3	1	Persian Gulf War	Military Warfare::Chemical Warfare::Persian Gu
19	36	36	4	2	Rifles	Ordnance, Missiles, and Munitions::Chemical Ordnance::
20	35	35	3	2	Grenades	Ordnance, Missiles, and Munitions::Ammunition::Grenade
21	35	35	4	1	Assassination	Military Warfare::Unconventional Warfare::Terr

Charts to improve the Cartridge Quality



Convera Confidential

Conclusion

Relate available information to

YOUR

decision-making processes

- Categorize with consistency
- Classify in context

Key Differentiators

- **The directory is a result, not a starting point.**
- **Ontologies are real ontologies: conceptual and explicit.**

Typical Structures

- Geography / Topic
  - **Terrorism in Philippines**
  - **Criminal Law in Texas**
  - **Domestic Sales**
  - **Security in Building C**
- Horizontal / Vertical
  - **Petroleum Business**
  - **AML Regulations**
- Vertical / Vertical
  - **Chemical Compounds for Alzheimer**

Over Defined Context

- Very large computational space
  - **“Chemical Compounds in Alzheimer Genomics”**
  - **-> 8500 diseases**
  - **-> 1,000 genes**
  - **-> 30,000 compounds**
  - **= 255 billion folders**
- Reversely proportional number of successfully populated folders
- Can’t be done by automatic CLASSIFICATION!

What We Do / What‘s Third Party

What We Do (Some Samples)

- List Search (Batch Mode)
- Information Profiling
- Automatic Categorizing
- Dynamic Classification
- Multi- and Cross Language
- Content Management
- Voice To Text and Automatic Meta Data Generation

Third Party (Some Samples)

- Pre Processing (nCase, etc.)
- Post Processing (Statistics, Facerec, …)

RetrievalWare 8
High End Categorization/Classification

Heinz Bachmann

Convera Confidential Information -- The contents of this Convera product presentation are confidential and governed by the NDA between your company and Convera. Such contents are subject to change by Convera at its sole discretion and Convera assumes no obligation to update such contents. Any binding representations, warranties and covenants by Convera shall be exclusively set forth in writing in a contract mutually agreed to and signed by your company and Convera, and such contract shall exclude all other written and oral communications including this presentation.