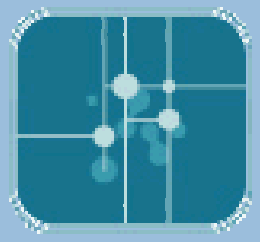
**Transforming Information into Value**



Emnekart Norge
2002

Oct. 18, 2002

**Dr. H. Holger Rath**
Head of Consulting

**empolis GmbH**
**www.empolis.com**

# Managing and Accessing a TV Media Archive with Topic Maps

**The RTV Dubai Project**

empolis

# Roadmap

- The Challenge
- The Idea
- The Solution
- The Demos
- The Conclusions

empolis

# The Challenge

If RTV Dubai knows what RTV Dubai knows

# RTV Dubai Facts at Project Start

- 4 TV stations and 3 radio stations
  - News, sports, entertainment, culture
- Analogue archive of media objects
  - 17,000 hours audio material
  - 21,000 hours video material
  - Material in Arabic and English
  - Plan to grow
- More than a dozen different databases
- Manual management of archive
  - Costly archiving process
  - Slow manual access
- Unsatisfied journalists
- Value of archive not used
- Lots of money spend for archive without getting ROI

empolis

# What to do?

- Introduce digital media archive
  - 1 terabyte online, 50 terabyte nearline
  - later on: 200 terabyte nearline
- Introduce “Portal” as single point of access
  - librarians, journalists, editors, producers
- Improve archiving process
  - automatic meta data extraction
  - automatic categorization
  - automatic cataloguing
- Simplify access to material
  - online access
  - intelligent search and explorative navigation

empolis

# What else to do?

- Improve operational efficiency
  - (dead) archive will become (lively) library
  - satisfied journalists will use library
  - reuse of content reduces production costs
- Get ROI
  - Spend money in new technology instead of wasting it for in-efficient archive
  - little risk when project fails, but big ROI when it succeeds

empolis

# Project Partners

**Customer:**

- Dubai Department of Information in United Arab Emirates

**Prime contractor:**

- BAMES – Bertelsmann Arvato Middle East Sales – an empolis company

**Sub contractor:**

- blue order – a tecmath company
  - media archive®
- empolis GmbH
  - empolis knowledge manager – formerly known as orenge and k42®

**Total project order value:**

- US$ 6.4 million over 4 years
- mainly hardware costs

empolis

# Project Challenges

- Legacy data migration
  - analogue tapes
  - other digital legacy systems
- Bi-lingual (Arabic & English)
  - material
  - search
  - ontologies (topic maps)
- Integrate blue order's Media Asset Management system with empolis' Information Access Management system
  - close integration necessary
  - interface level and data level
  - empolis built business connector
  - strategic partnership

# The Idea

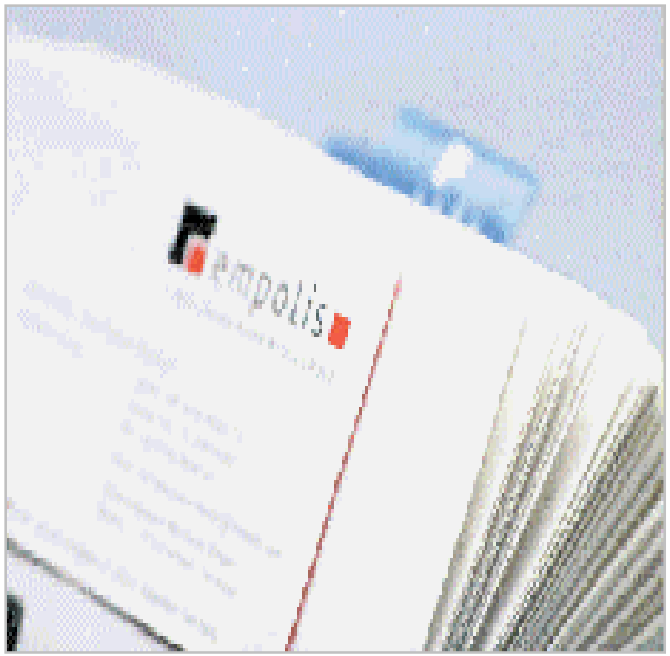
Applying knowledge management technologies

# Knowledge Management

Knowledge management transforms structured and unstructured information, selects and combines the important information for a user in a specific context, so that decisions and the actions of companies are supported.

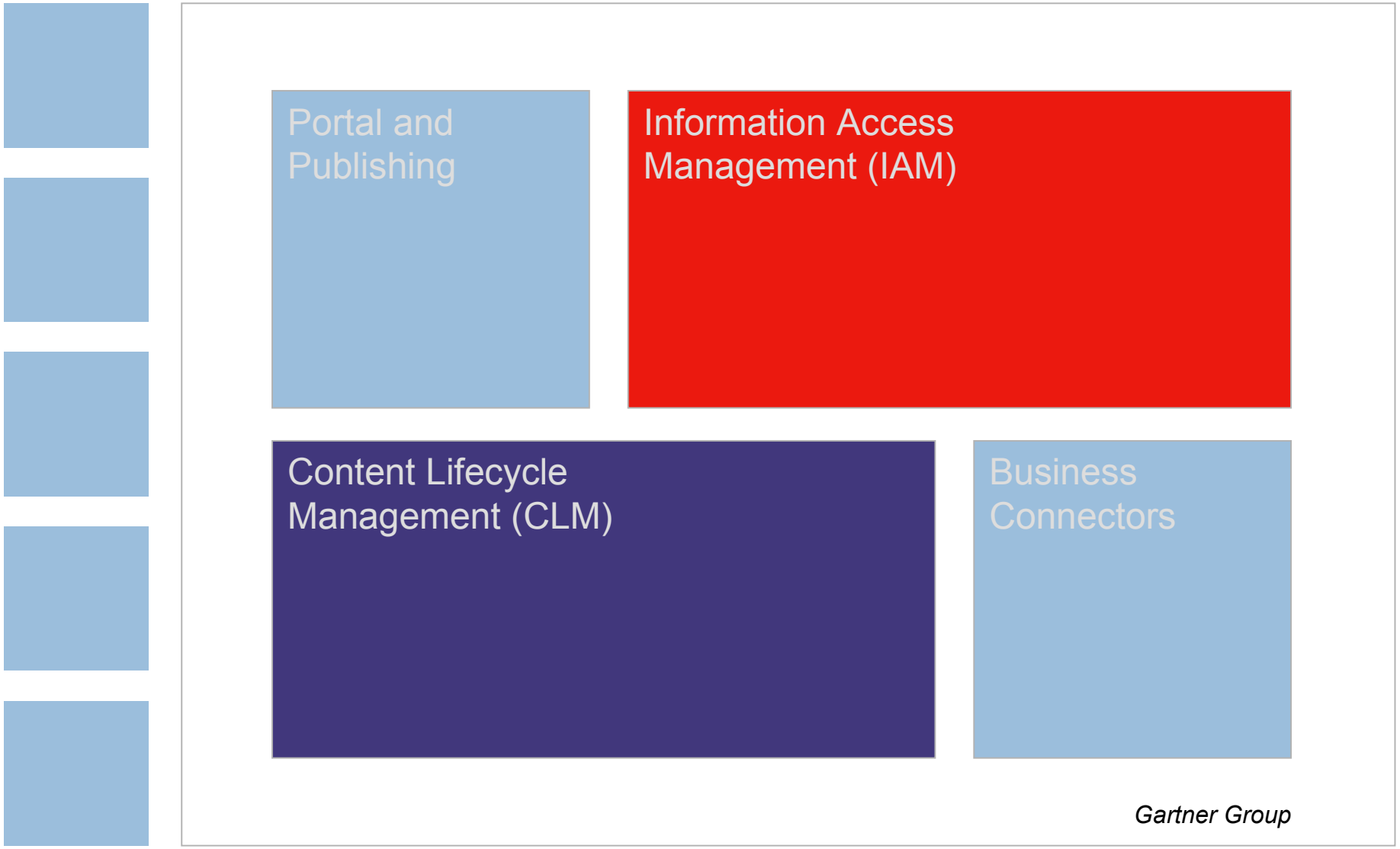
To be able to realize knowledge management in a company, a multitude of individual IT solutions must be taken into consideration and viewed within a holistic approach, e.g. archiving and DMS systems, retrieval and agent technologies, workflow, portal and content management systems.

*Meta Group 2001*

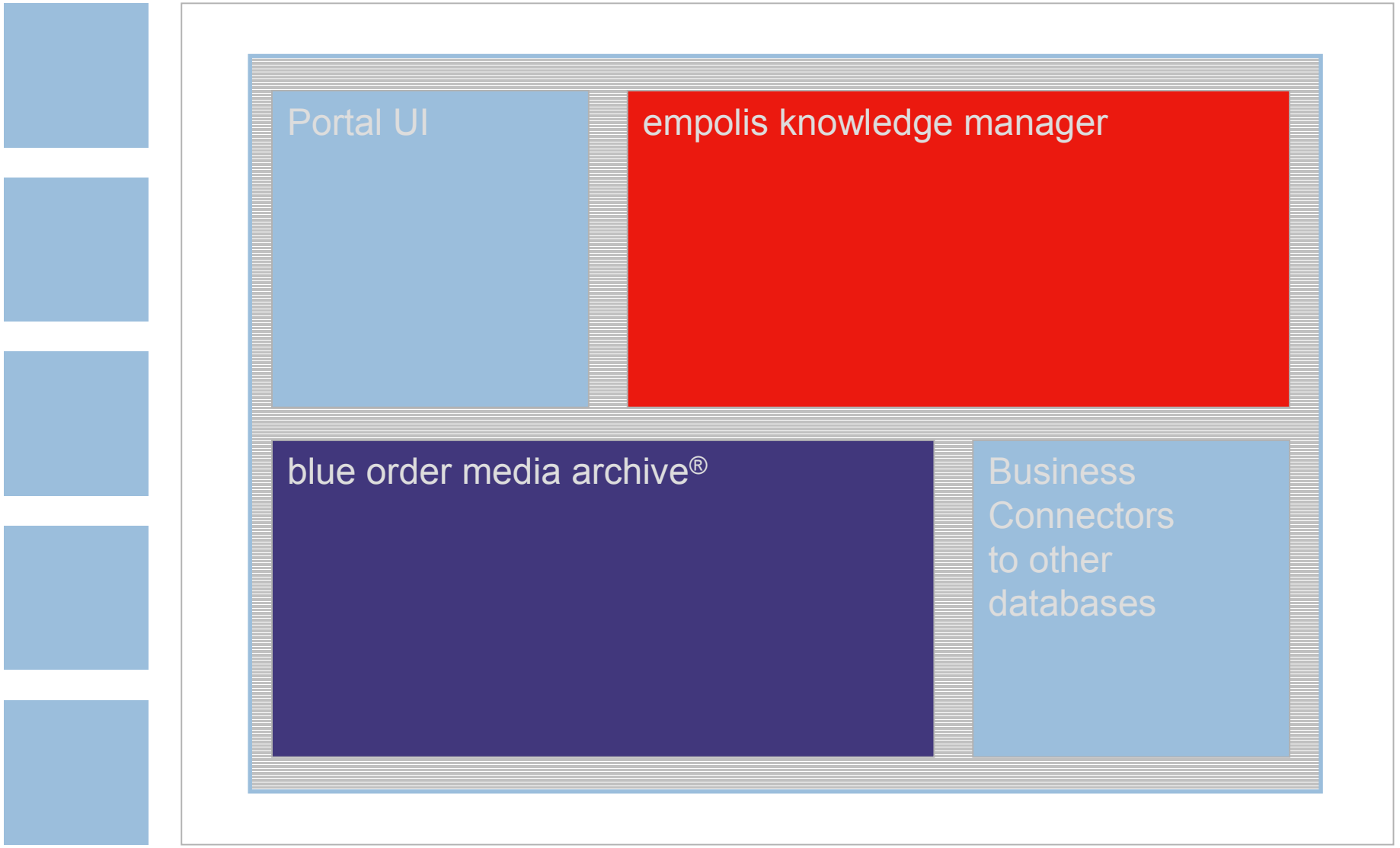



empolis

# Enterprise Knowledge Management = *CLM + IAM*

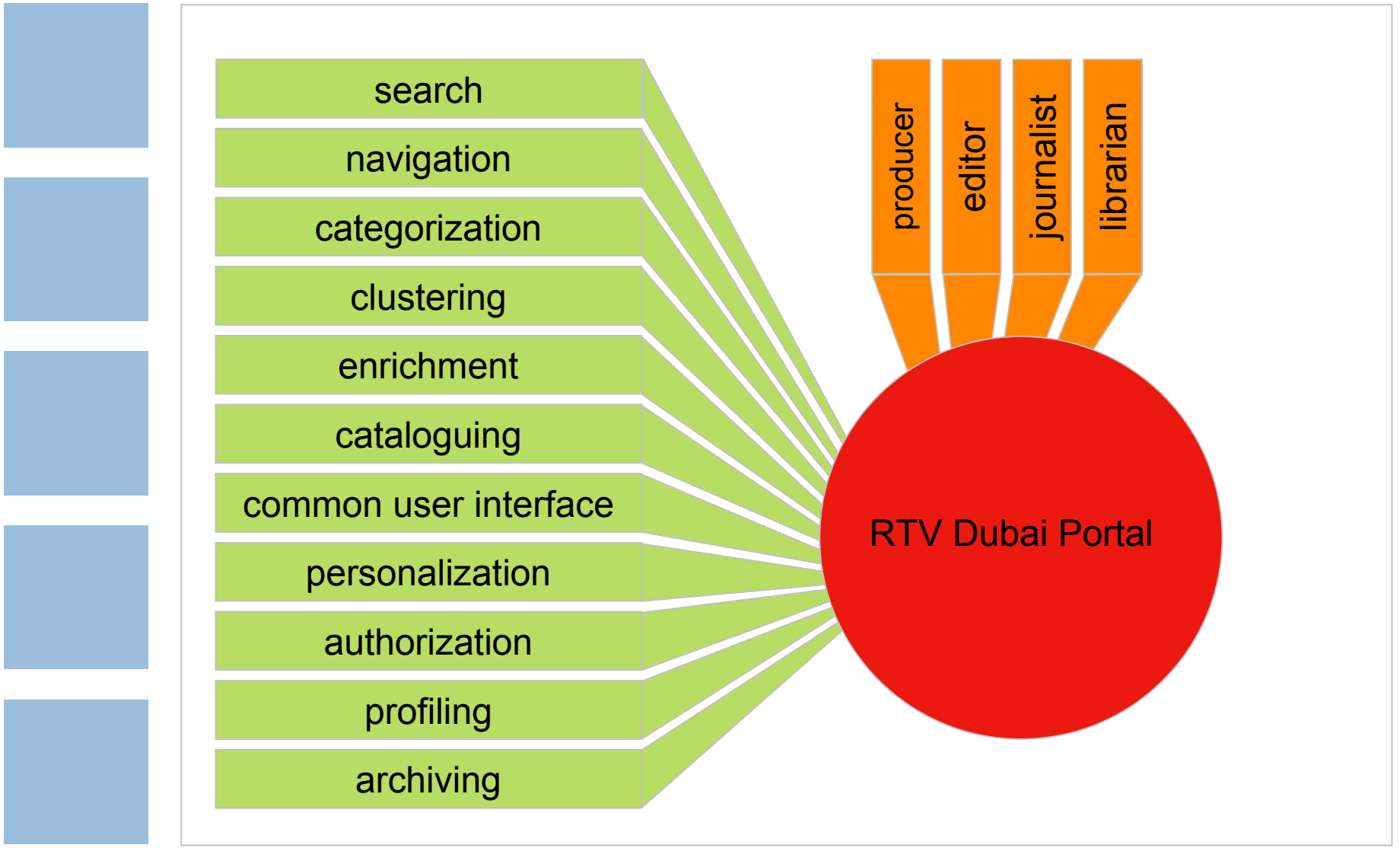


# RTV Dubai Media Library Solution



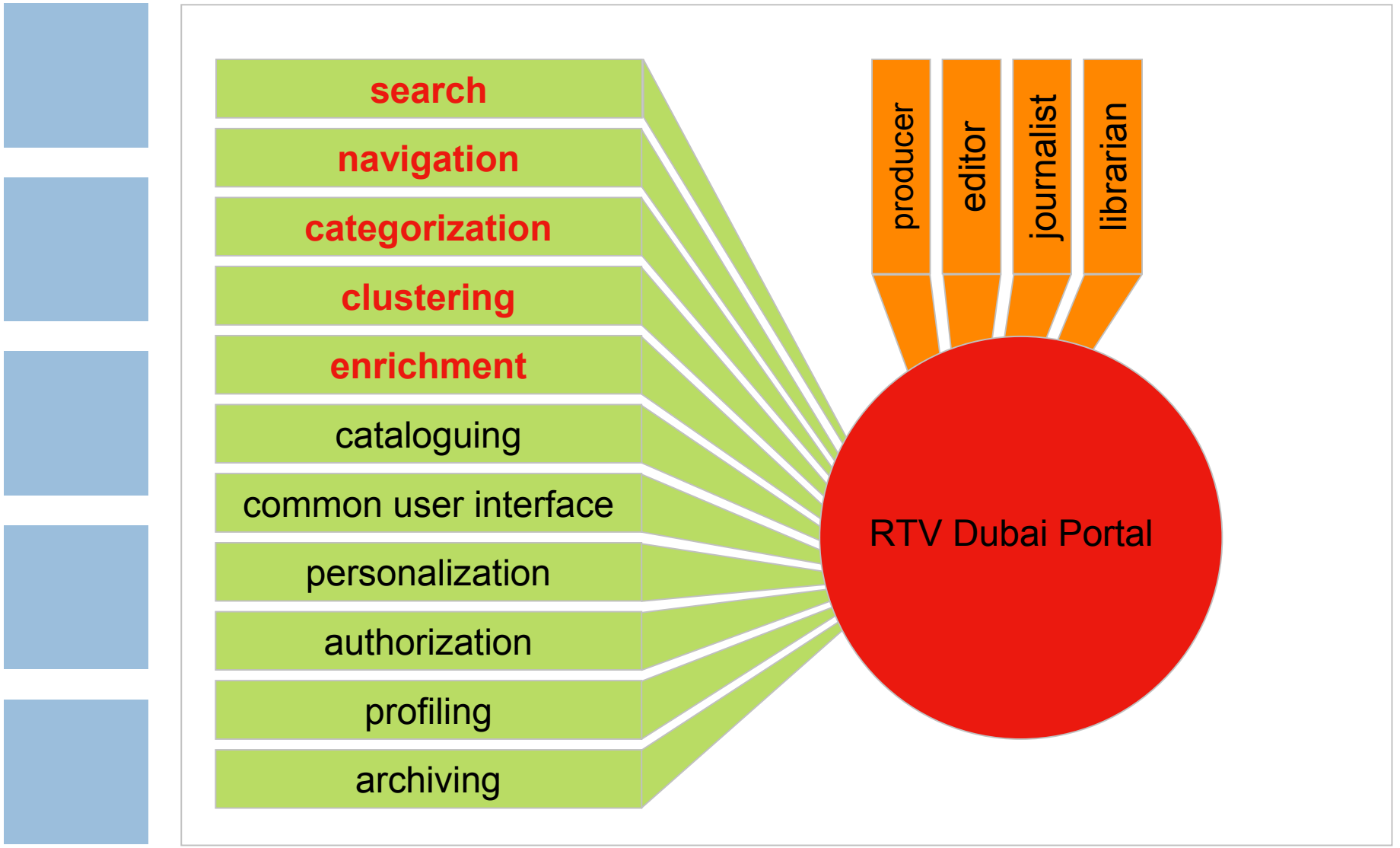

empolis

# RTV Dubai Media Portal




empolis

# RTV Dubai Media Portal



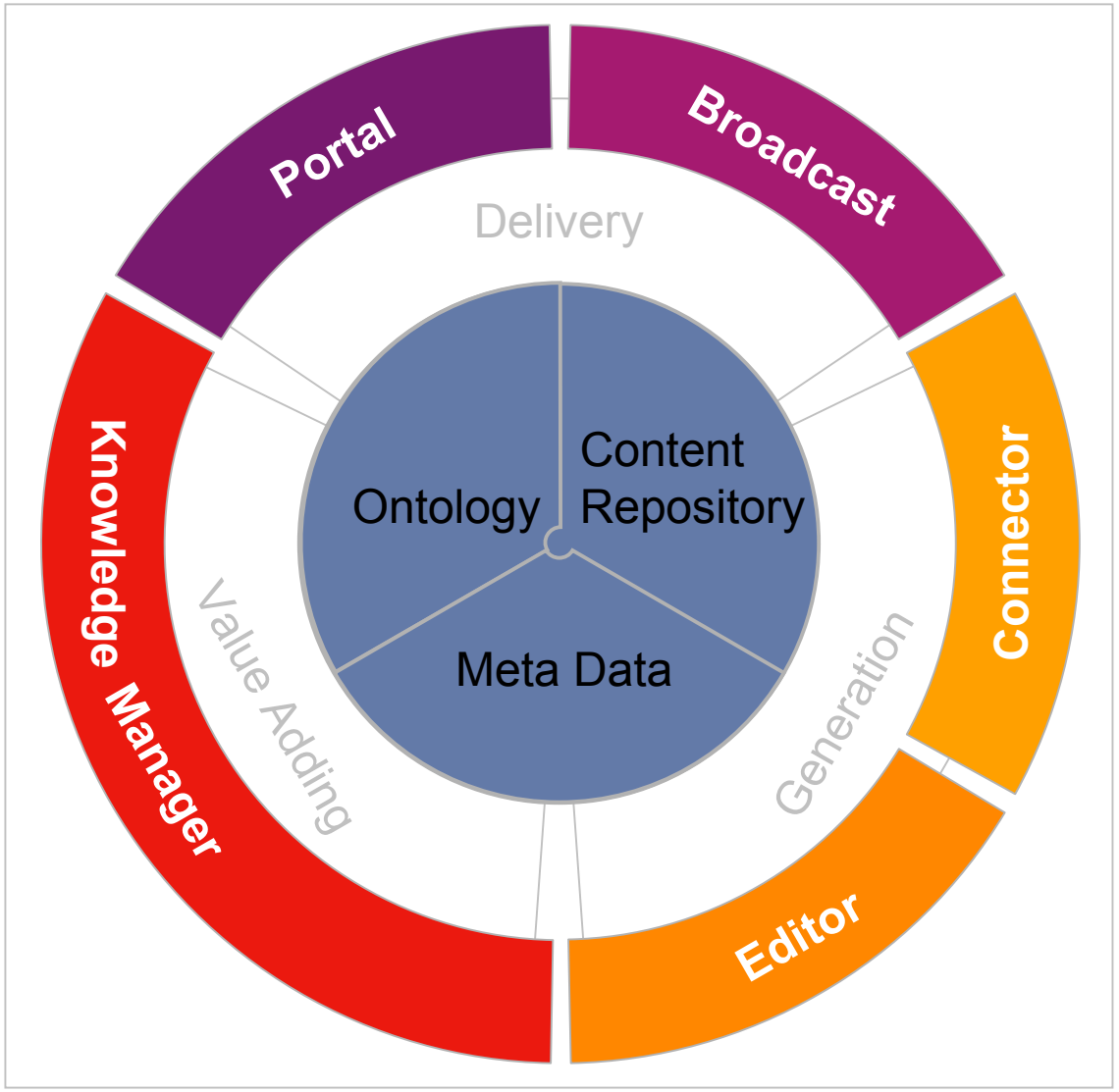

empolis

# Transforming Content into Knowledge

**An explicit knowledge model (= Topic Map) adds significant value to managed content.**

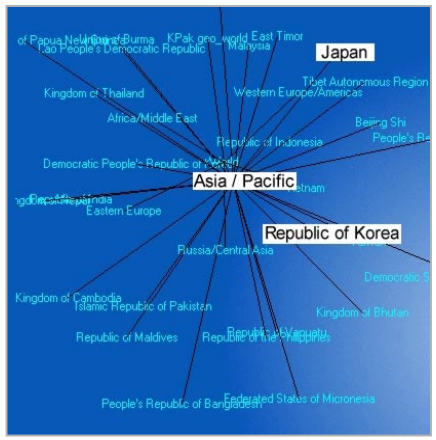
**Topic Map is used for:**

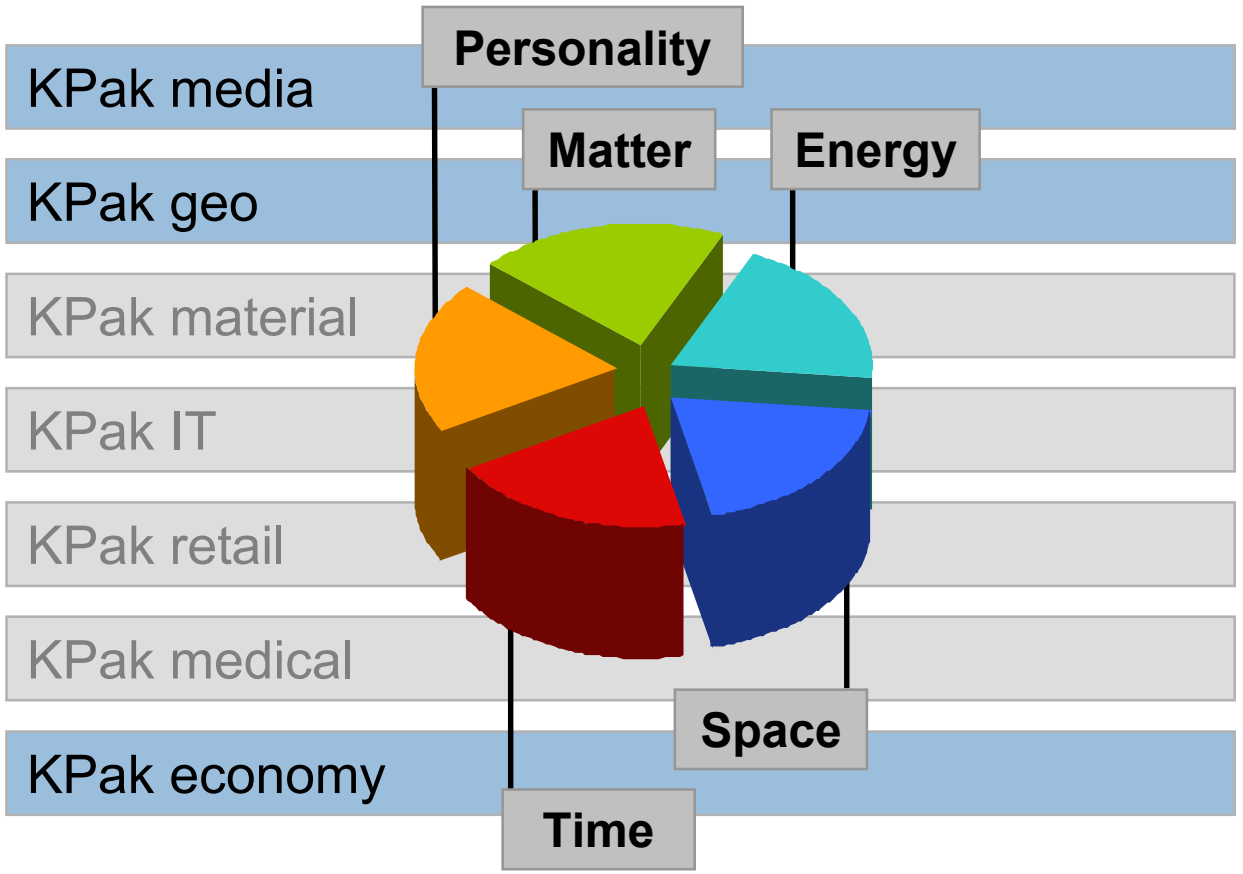
- Semantic Indexing & Intelligent Retrieval
- Navigation
- Categorization
- Classification
- Enrichment




empolis

# Pre-Modelled Topic Maps as Starter Kit




**empolis knowledge packs**

empolis

# The Solution

Searching, filtering, and organizing
the RTV media archive

# Traditional Search in Media Archives

Find media objects (at all)

# Traditional Search in Media Archives

**Full-text retrieval:**

- Query:

  "Formula 1" AND
  "Michael Schumacher" AND
  "year 2000" AND
  "winner"

**Database search:**

- Query:

  Date of event: *May 14, 2001*
  Sport: *Formula 1*
  Athlete: *Michael Schuhmacher*

empolis

# Problems with Traditional Search

**Full-text retrieval:**

- **Too many hits!**
- Query: "winner" and "Michael Schumacher"
- Result:
  - *Formula 1 race in …*
  - *…*
  - *Interview with Michael Schumacher …*
- Reason:
  - Query is too vague

**Database search:**

- **No hits!**
- Query: "Race on May 14, 2001 by Michael Schuhmacher"
- Result: *0 hits found.*
- Reasons:
  - Race was on May 13, 2001; not on May 14
  - Mistyping of athlete

# Intelligent Search

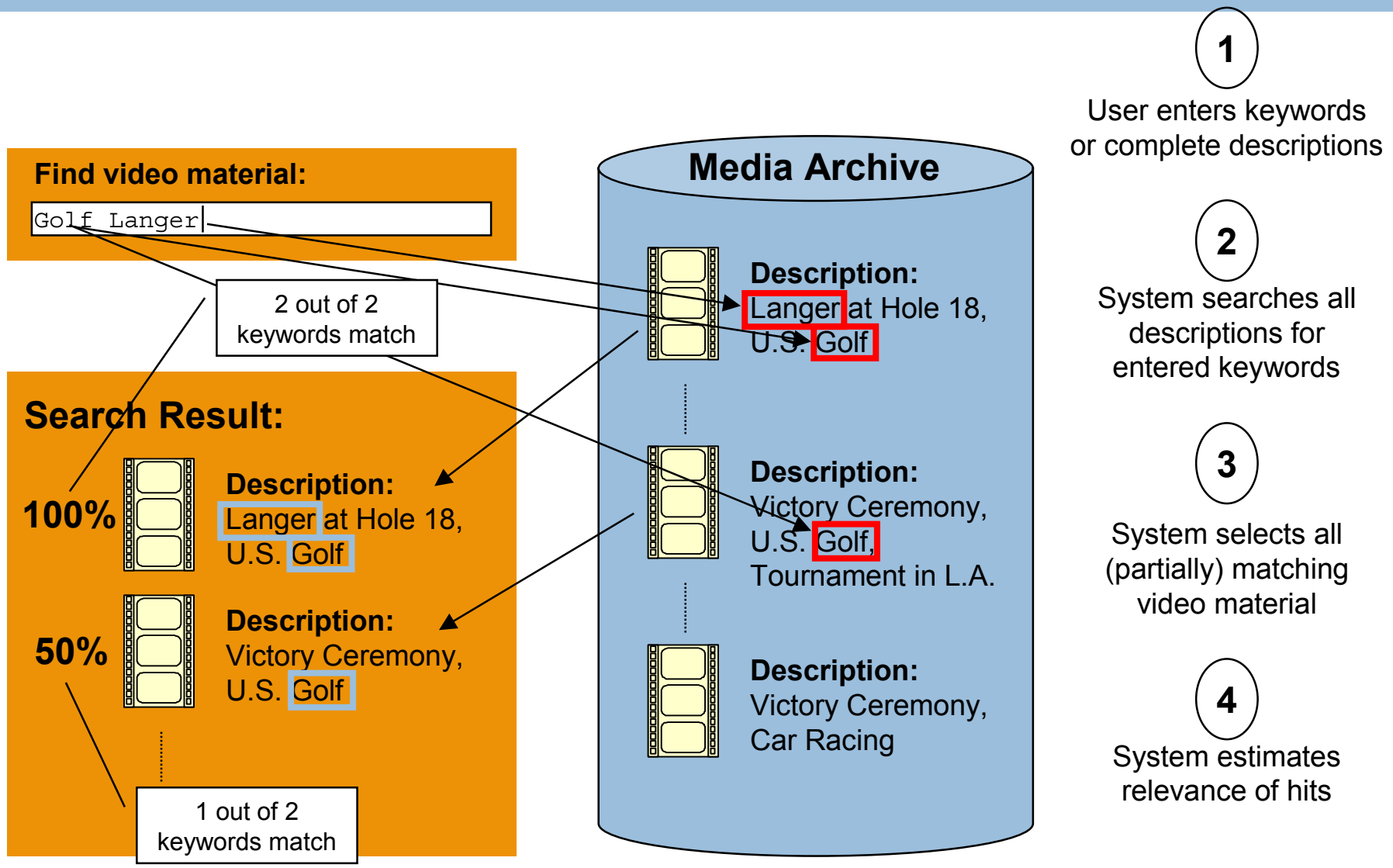
Find relevant media objects

empolis

# Goals for Advanced Search for Media Objects

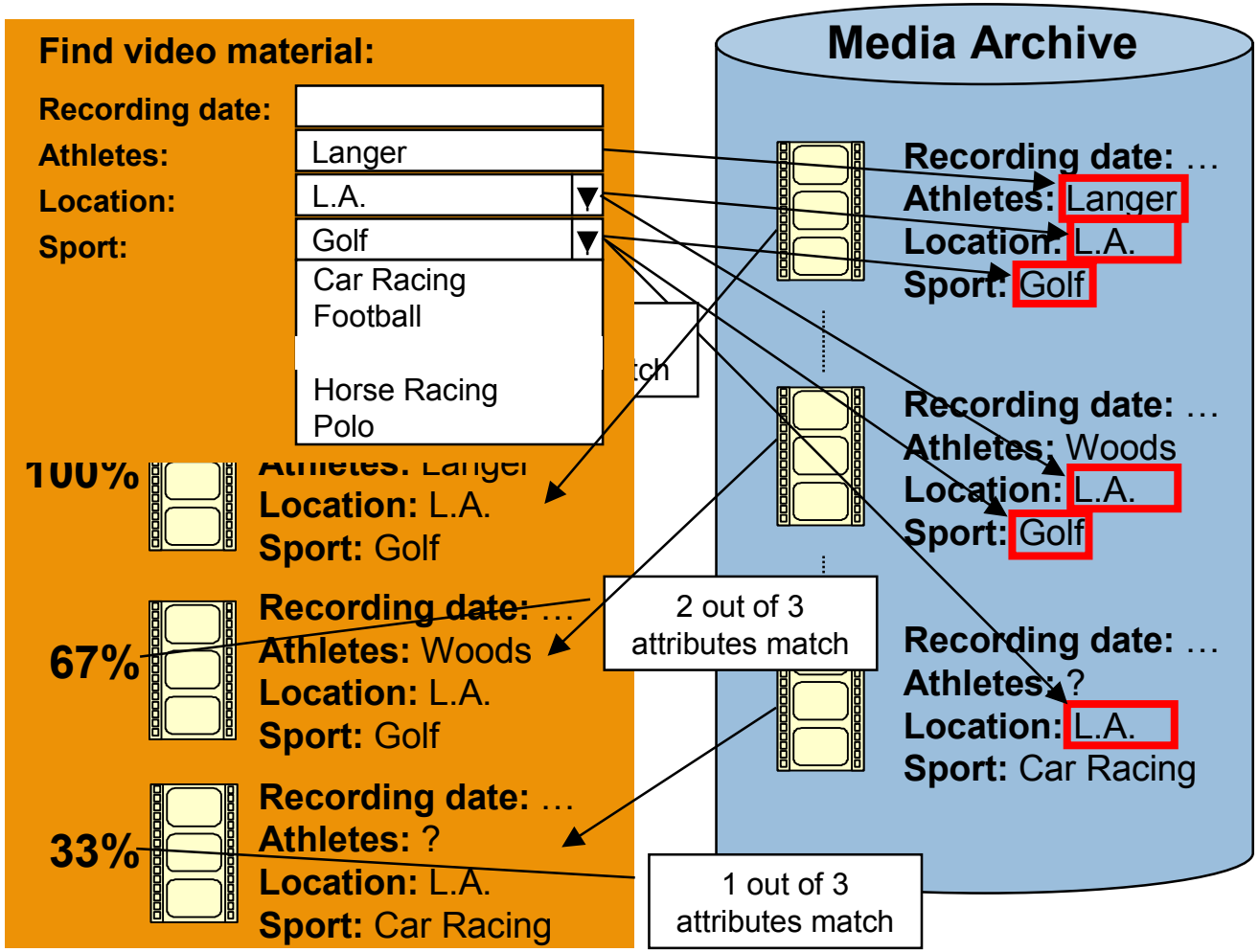
- Find relevant media objects with little effort
  - i.e., with as few search requests and mouse clicks as possible
- Tolerate misspelled input
  - suggested spell correction
  - controlled vocabulary (= topic map)
- Allow cross-lingual search
  - e.g., find English documented media objects with Arabic query and vice versa
- Find related information and media objects
  - content enrichment (= links into topic map)
  - explorative browsing of topic map

# Full-Text Search




empolis

# Search Using Attributes




1

User enters attribute values

2

System searches all characterizations

3

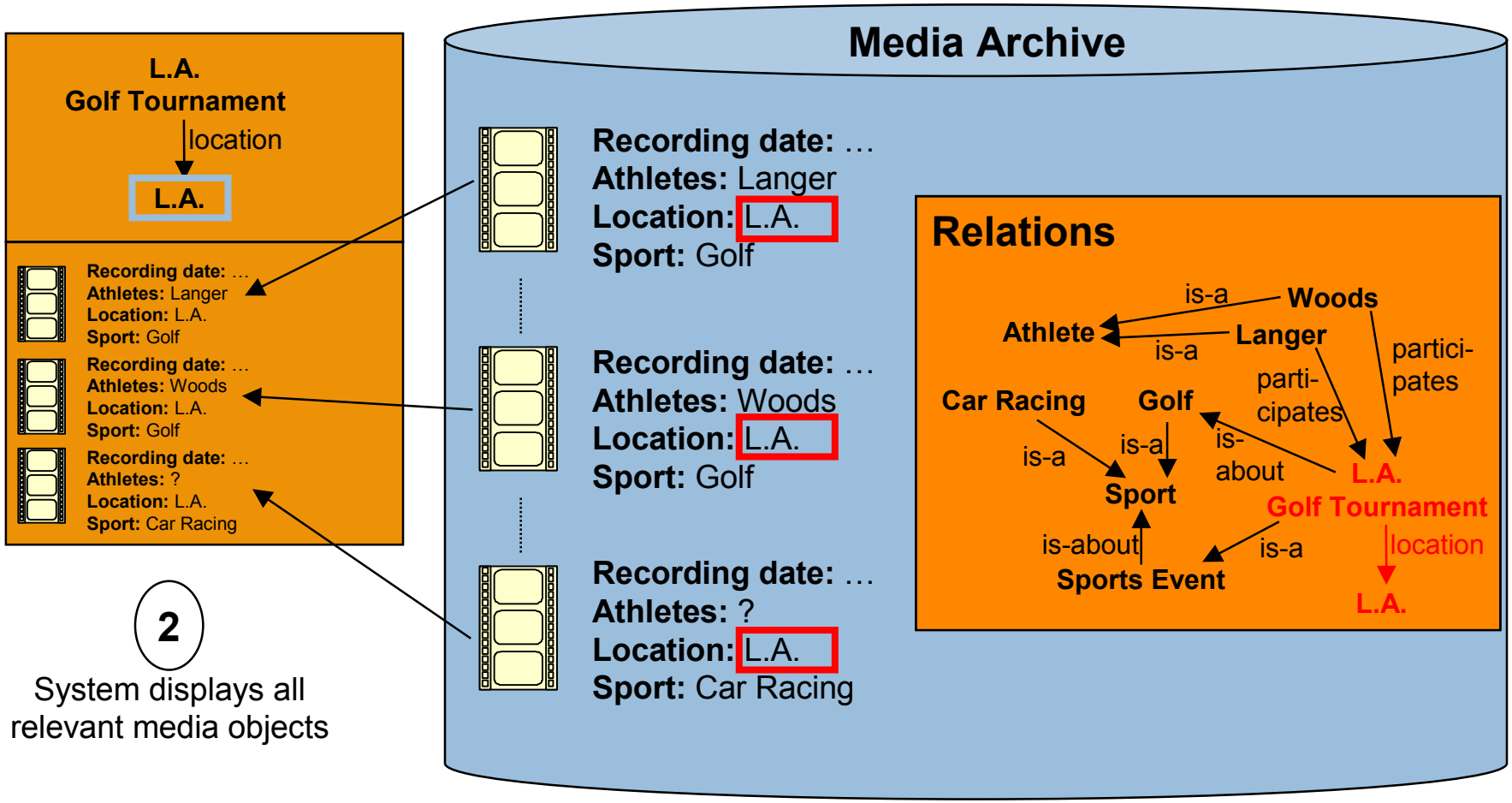
System selects all (partially) matching video material

4

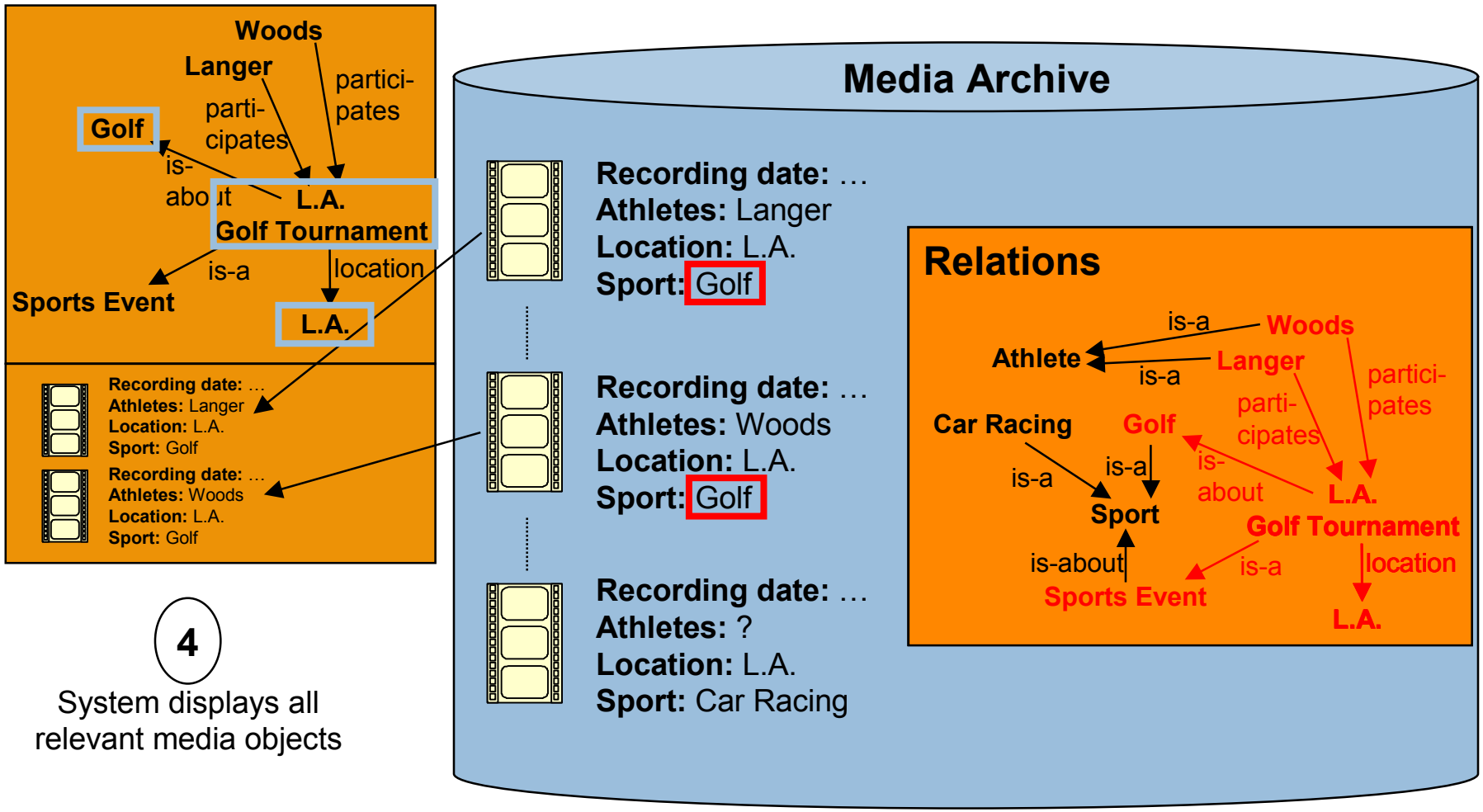
System estimates relevance of hits

empolis

# Search Using Visual Navigation



# Search Using Visual Navigation

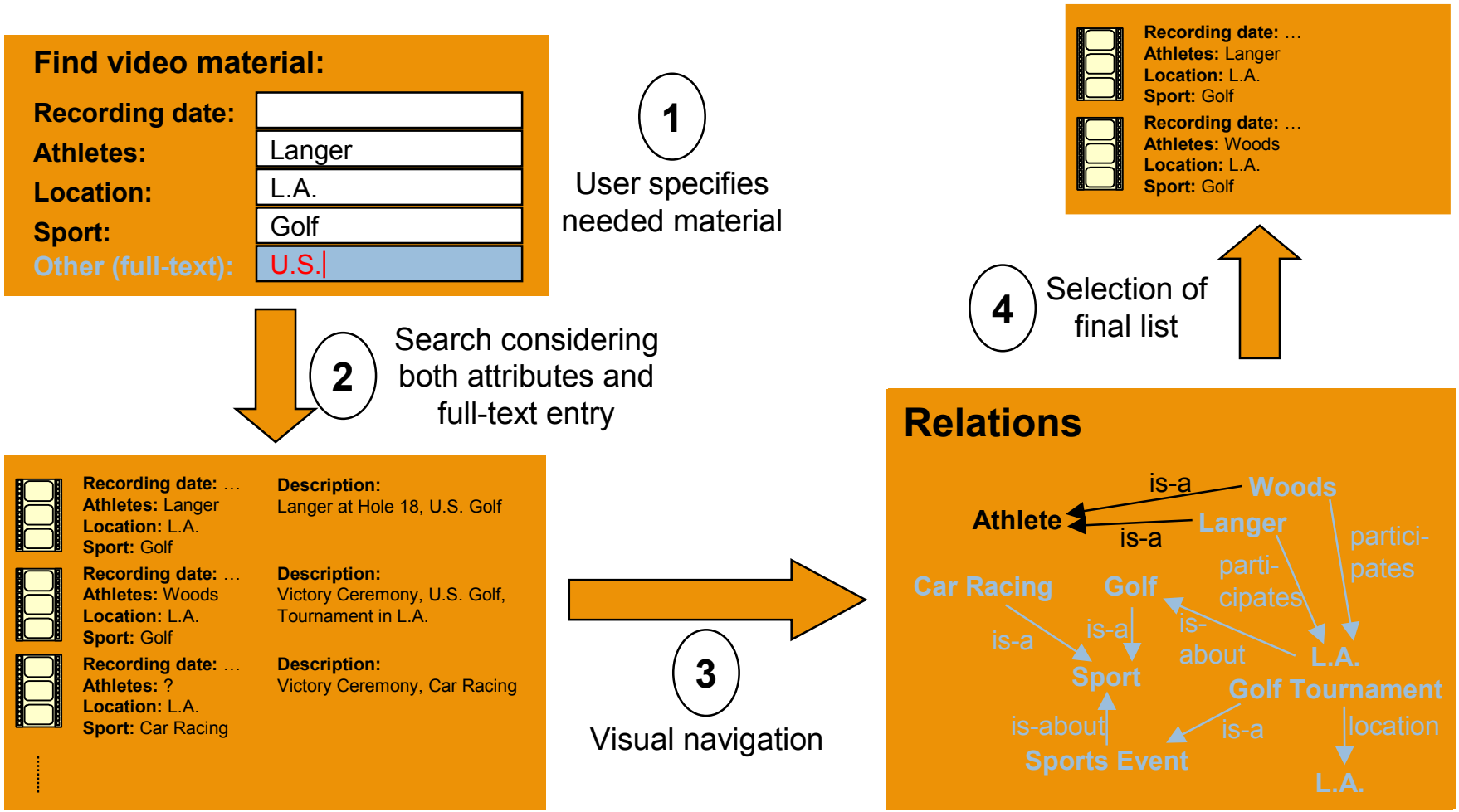



4

System displays all relevant media objects

empolis

# Seamless Integration of Search Methods




empolis

# Advantages of the Proposed Approach

- Spelling correction
- Similarity-based search
- Relevance estimation of hits
- Explanation why hit appears in list
- Intelligent dialog to refine query
- Cross-lingual search
- Visual Navigation for finding related material
- Open Standards (XML, SMEF, Topic Maps, Dewey)

empolis

# Powerful Archiving

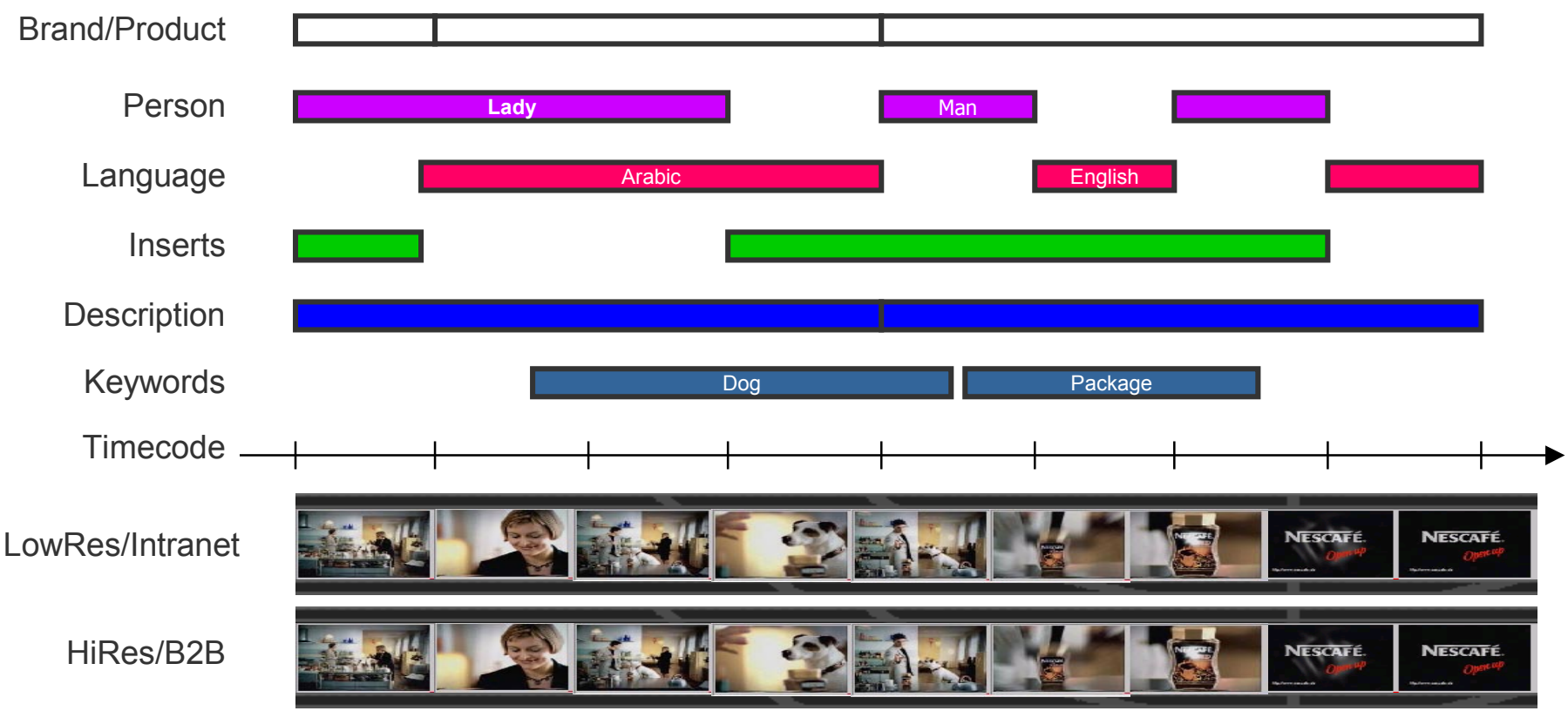
Make media objects available

# Goals for Archiving Media Objects

- Faster archiving
- More accurate archiving
- More complete archiving
- High automation
  - i.e., automatic generation of attribute values and relations

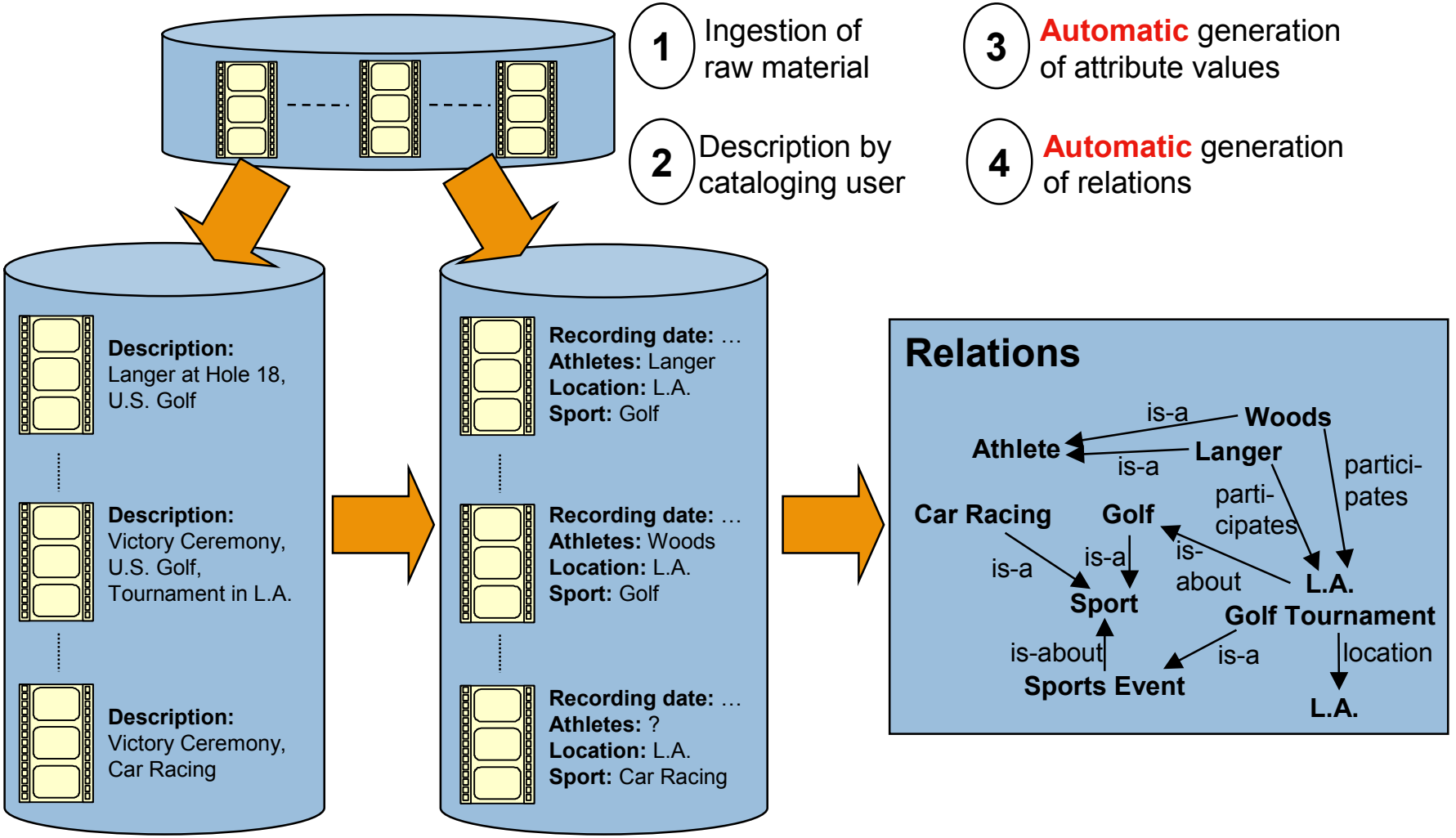
empolis

# Stratified Documentation in media archive®



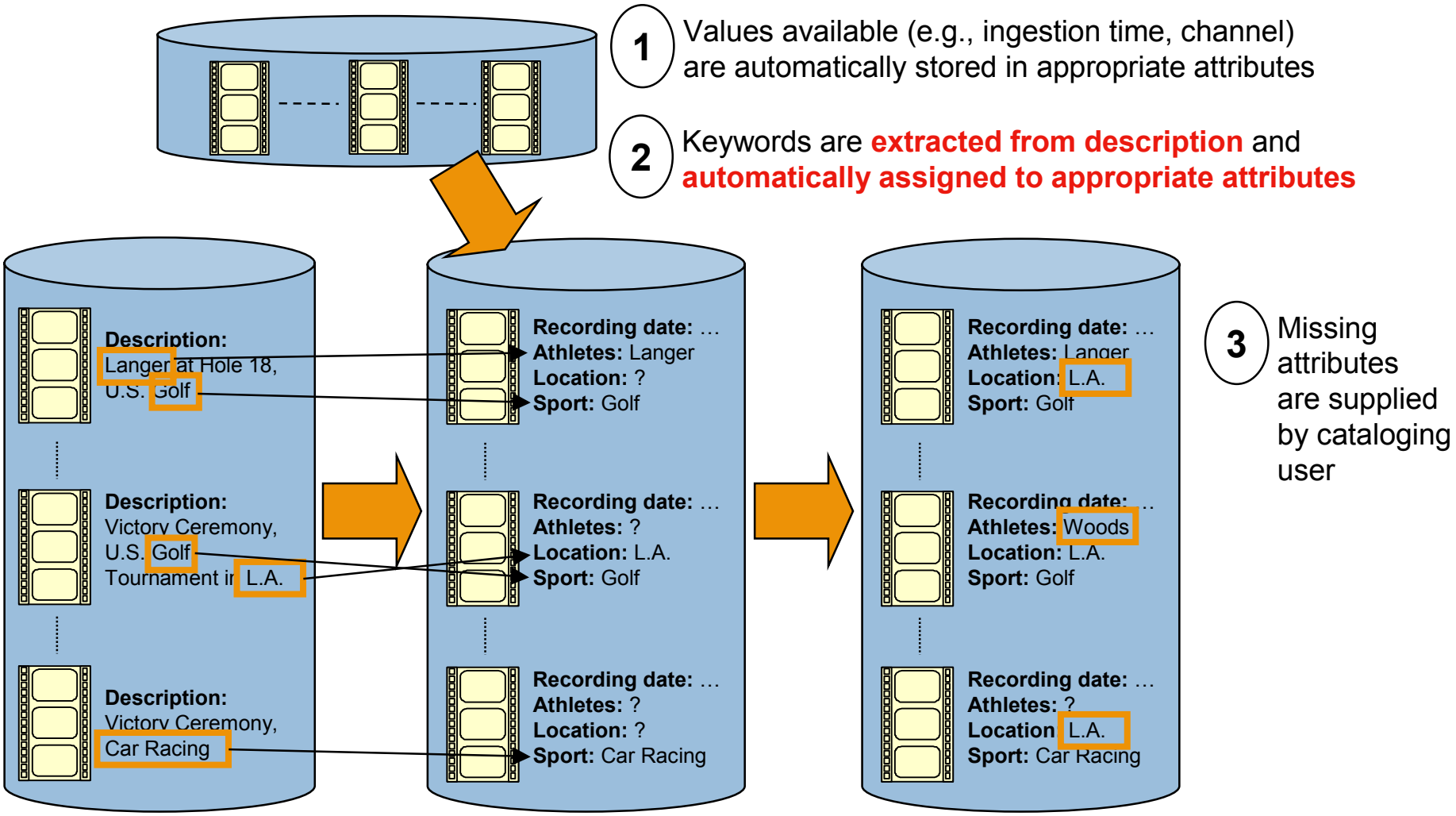

empolis

# Archiving: Workflow




empolis

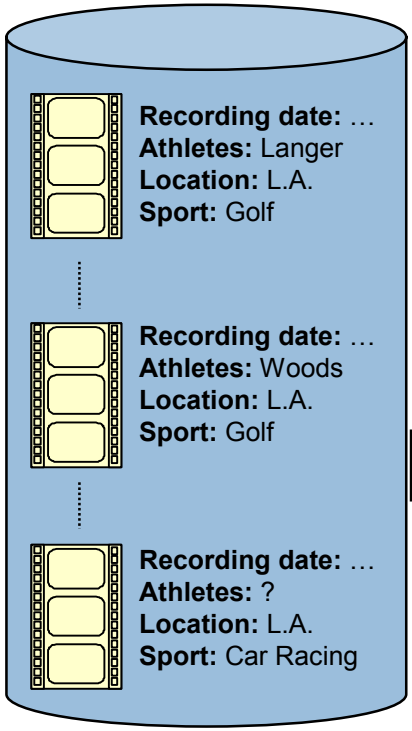
# Generation of Attribute Values

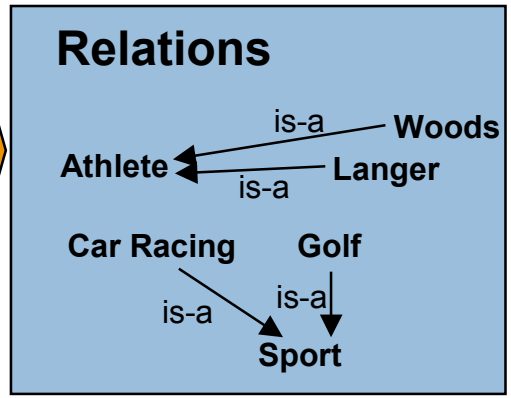


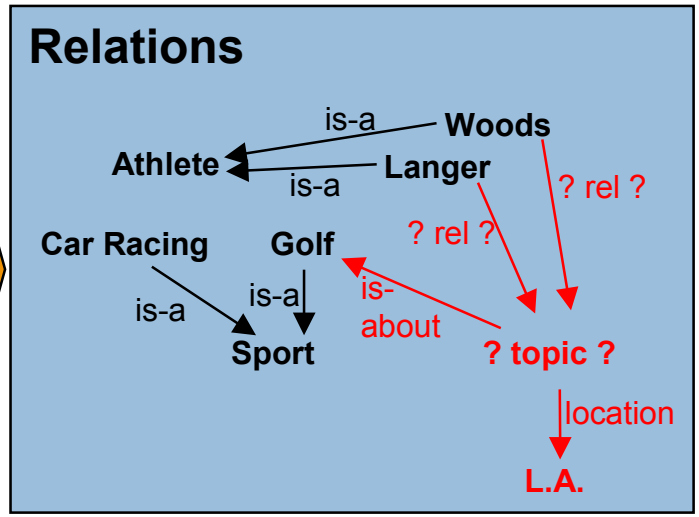

empolis

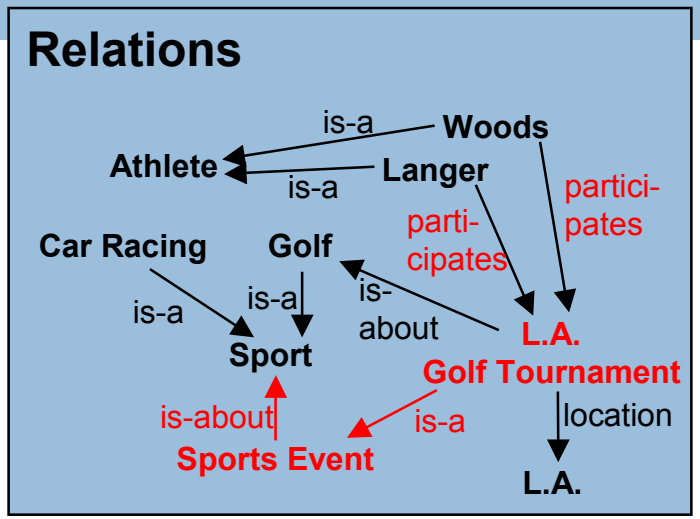
# Generation of Relations

1. Based on attribute names and values, an initial topic map for relations is generated
2. Further relations are **suggested through statistical analysis**
3. Final modifications and extensions are done by cataloging user



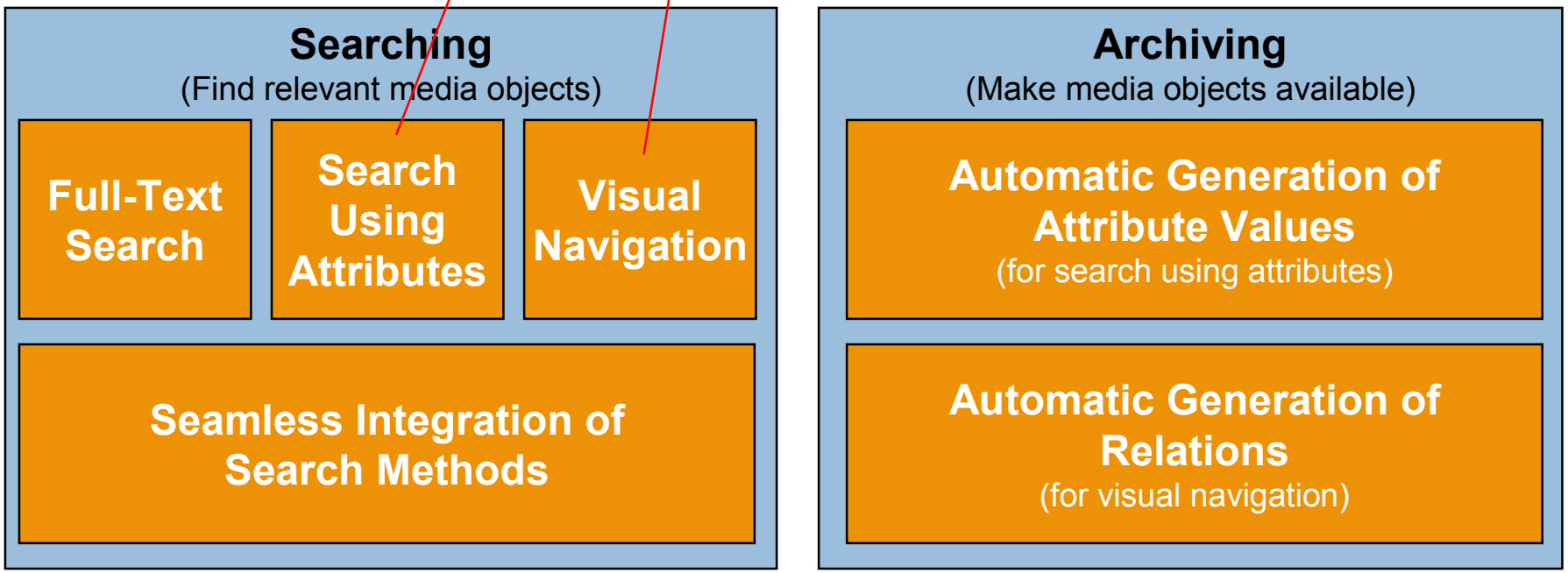




empolis

# Summary

Attribute search supports among others the **Dewey classification** and **SMEF** (Standard Media Exchange Framework – a BBC standard for exchanging video material)

Supports **Topic Map** standard (ISO 13250)




**There are 3 ways to find media objects**

**The search methods can be used in combination**

**Before media objects can be used, they must be stored (archived)**

**The information needed for the sophisticated search methods can be generated automatically**

empolis

# The Demos

“Movie Library Sales“ and “Knowledge in Media”

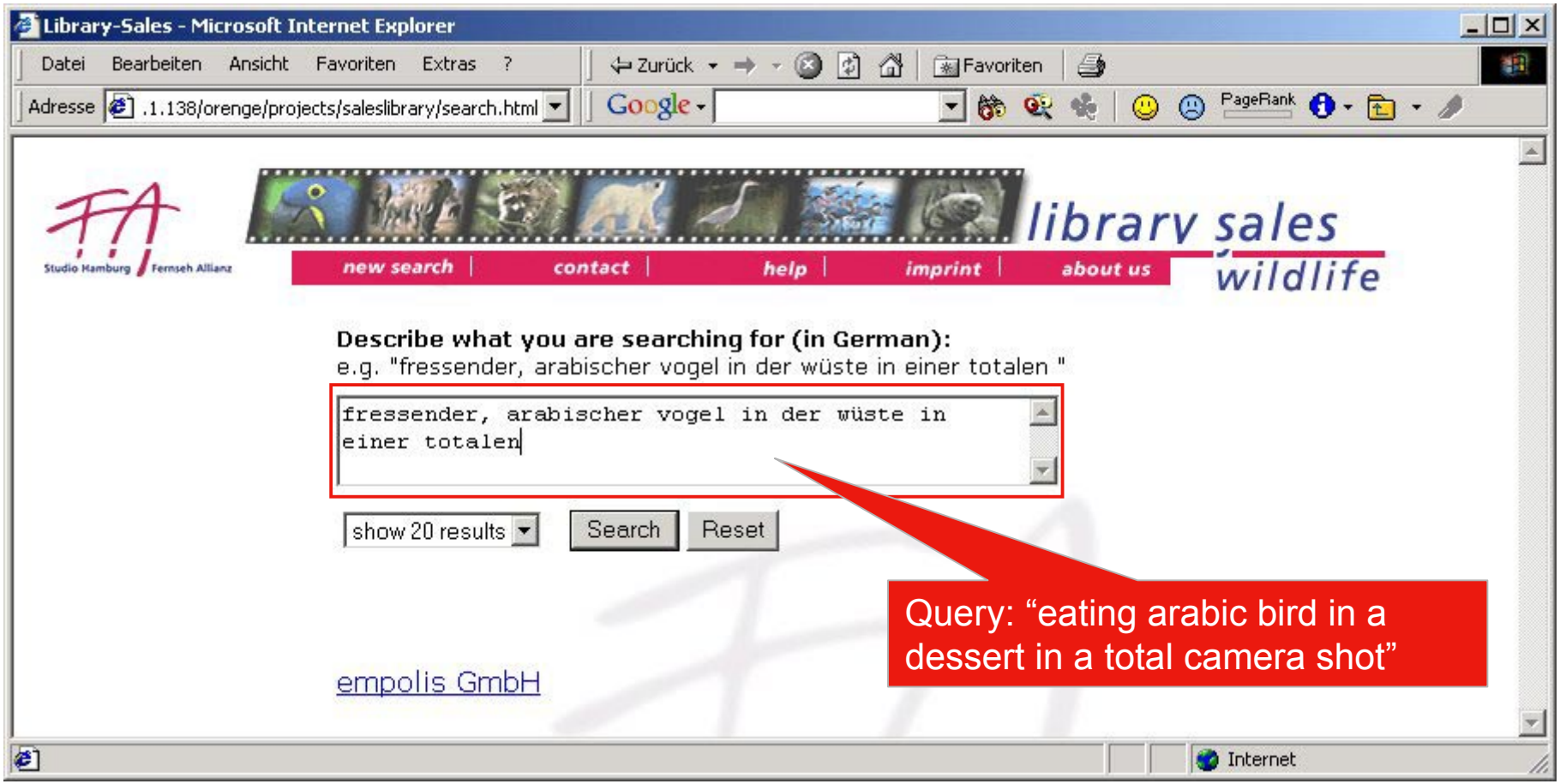
empolis

# Demo – Movie Library Sales

- Finding animal movies
  - intelligent search driven by topic map
  - KPak geo
  - German movie library

empolis

# Demo – Movie Library Sales



# Demo – Movie Library Sales

empolis

# Demo – Knowledge in Media

- Finding news in text archive
  - search & navigation
  - KPak media
  - Reuters news

empolis

# The Conclusions

Topic maps play important role,
but are only one technology among others

empolis

# Conclusions

- RTV Dubai extremely benefits from the combination of
  - digital media archive
  - improved and simplified access
  - automated archiving process
  - increased user satisfaction
  - higher content reuse ratio
- Topic maps are the central ontologies driving
  - categorization & classification
  - search
  - navigation
  - enrichment
- But topic maps have to be accompanied by other technologies
  - Media meta data (= stratified documentation of material)
  - Case Based Reasoning (= intelligence)
  - Portal (= personalization, authorization, integration)

# </End>

Thank you for your attention!

Questions? Answers!

**www.empolis.com**