

**khresmoi**  
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

# Tutorial on searching text and images in the medical domain

Henning Müller  
MIE 2012, Pisa, Italy

**Hes-SO VALAIS WALLIS**  
Haute Ecole Spécialisée de Suisse occidentale  
Fachhochschule Westschweiz  
University of Applied Sciences  
Western Switzerland

**Overview**

**khresmoi**  
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

- Improving search in the medical domain
  - Allan and Henning
- Searching for medical images
- Who searches for medical images and how?
- Combining text and visual search
- Challenges for search
  - Allan and Henning

2



# Improving search in the medical domain

3

## Scenario



- A close friend of yours has her 5-year-old daughter diagnosed with Leukemia
- As you are regarded as a medical professional in the largest sense they ask you to help find more information on the disease
- How do you search for this information, what are the strategies?
- What type of information are you targeting?
- How can you assure that truthful information is being transferred?
- How do you explain it to the 5-year-old?
- What are difficulties and disadvantages of this approach? 4

## Search target?



- Is a **document** searched for?
- Or an answer to a specific **question**?
- Maybe an **expert** in the domain?
- Educational material?
  - Maybe **videos**, or nice pictures also for children?
- Description of the disease vs. symptoms, treatments, chances of survival?
  - With the goal of making a choice of treatment, for example, comparing various options for the choice

5

## Trustfulness



- How can we make sure that information is **correct**, or at least not totally wrong?
  - Sometimes differing opinions exist, knowledge changes over time
  - Cross check several sources, but this can also lead to wrong ways
  - **Quality labels** or certificates of trust can help
    - HONcode
  - **Classification** of web pages into several classes
    - What are the characteristics
    - Success needs to be monitored closely (spam filtering)



6

## Level of understanding



- Different target groups have a totally different **level of information** that is required
  - Also physicians between a specialist and a GP
- What about **children** and the **elderly**?
- Level of understanding changes the more we read about a specific topic, we become experts
  - Not a fixed thing ...
- Users can give more information to get **personalized results**
- Cyberchondria

7

## Document and page ranking



- Should **wikis** and **blogs** be rated highest?
  - Google does this but for medical information several studies show that this might not be the best strategy
- Most queries are short (1-2 terms) and thus **not specific** at all, how to deal with this?
- Is the search target known?
- Users want advanced search options but then these are rarely used in practice
- Go beyond ranking, explore the content space

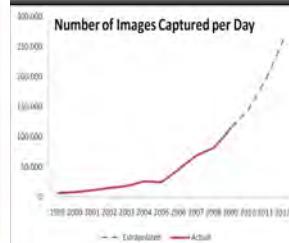
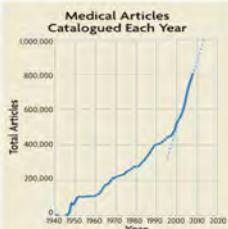
8

# Searching for medical images

9

## Motivation

- Medical imaging is estimated to occupy **30% of world storage capacity** in 2010!
- **Mammography** data in the US in 2009 amounts to **2.5 Petabytes**



### Riding the wave

How Europe can gain from the rising tide of scientific data

Final report of the High Level Expert Group on Scientific Data  
A submission to the European Commission  
October 2010

Riding the wave – how Europe can gain from the rising tide of scientific data, report of the European Commission, 10/2010.

10

## Retrieval of images



- **Text** retrieval of images
  - Is there any text attached to the images?
  - Doing this manually is expensive, subjective, **language** dependent, ...
  - Take text close to the images (such as captions)
  - Semantic **concepts** could help in some cases
- **Visual** retrieval of images
  - Using automatically extracted visual features
  - Content-based image retrieval (**CBIR**)
  - Query by Image Example(s) (QBE)
- **Multimodal** retrieval (**text+images**)

11

## Types of annotation

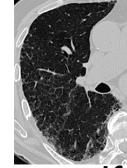


- What is **in** the image?
- What is the image **about**?
- What does the image evoke (**feelings**)?



## Content vs. context

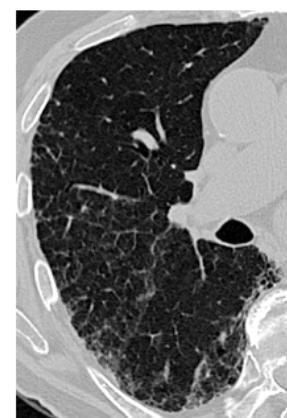
- Most often text around images does not describe the image content itself
  - Unless specifically annotated for retrieval
  - Text often gives the **context** in which the images were taken (private, also medical)
- Image **content** is rarely described precisely and completely with text
  - Visual features describe the pure content
  - Low level of semantics
- Content and context are **complementary** for search



13

## Age matters

  
Khresmoi  
MEDICAL INFORMATION ANALYSIS & RETRIEVAL



14

## Types of information needs/searches



- Known-item search (i.e. telephone number, ...)
- Question answering
- Exploration, exploitation, informational
- Topic search or **open ended** search
- Comparison search (between things)
- **Expert** search, person search, entity search
- Geographical search
- Literature search
- **Multimedia** search (increasingly given as results)
- Browsing, no specific goal



15

## Types of image searches

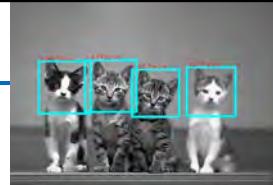


- **Recent** (time) pictures (journalists)
  - Date, given with camera
- Pictures of specific **places**, monuments
  - GPS in many cameras
- Pictures with particular **persons** (private search)
  - Face detection, recognition (Facebook, Picasa)
- Pictures with particular **objects**, types of images
- Pictures evoking specific **feelings**
  - Fear, joy, happiness, ...
- **Similarity** search/browsing (Medicine, journalism)

16

## Visual information for retrieval

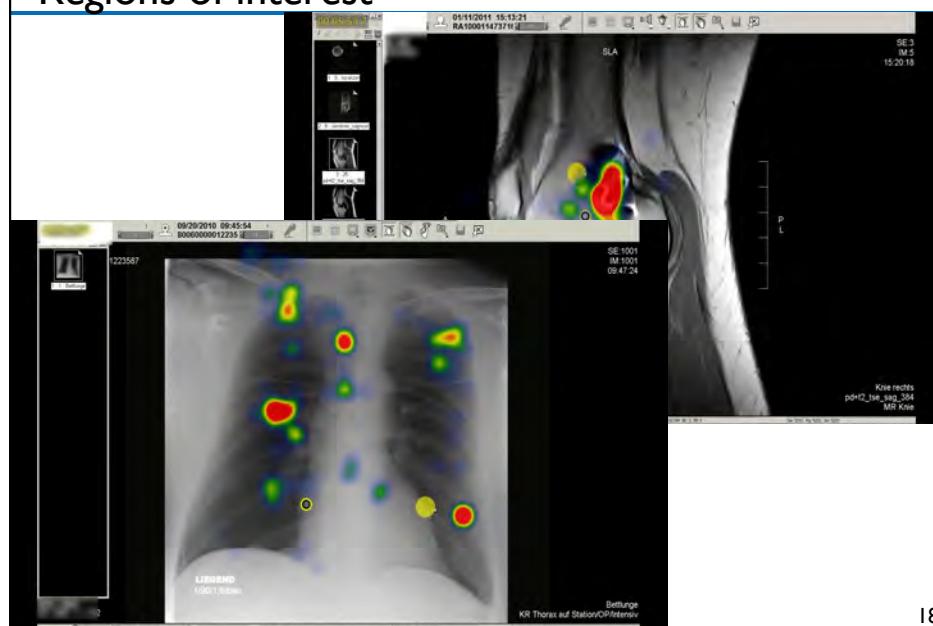
- Object **detection**
  - Then potentially mapping the objects to an ontology
  - Usually works well for a small number of objects
- Image **classification**
  - Training data, limited set of classes
  - **Global** classification of images vs. **local** classification of pixels, regions
- Similarity **retrieval** of images
  - Global image information, regions of interest (ROIs), small
  - No training data, relevance as criterion for quality



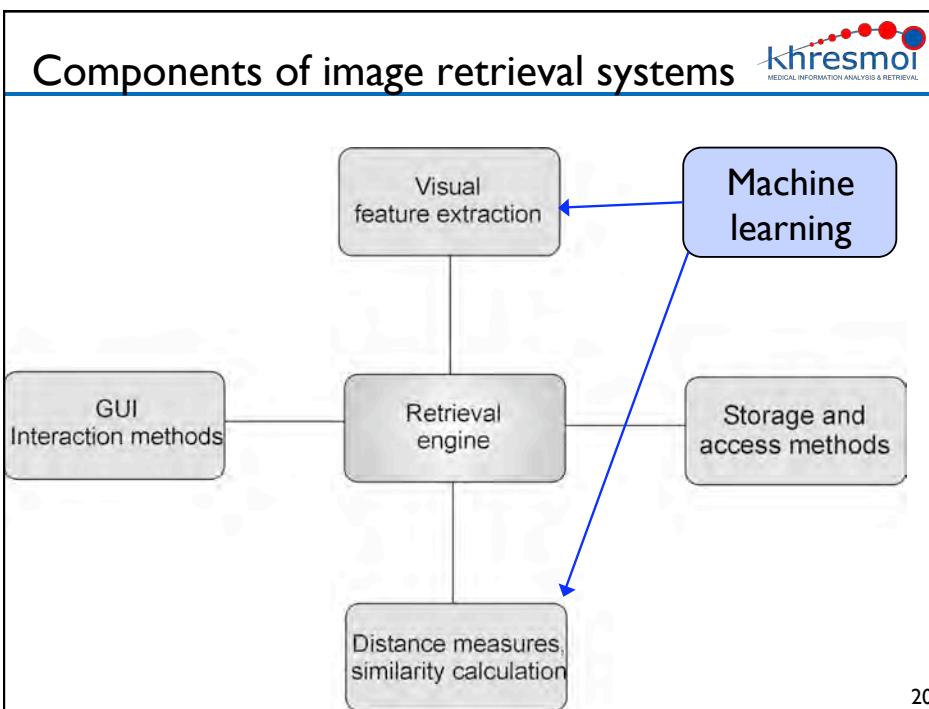
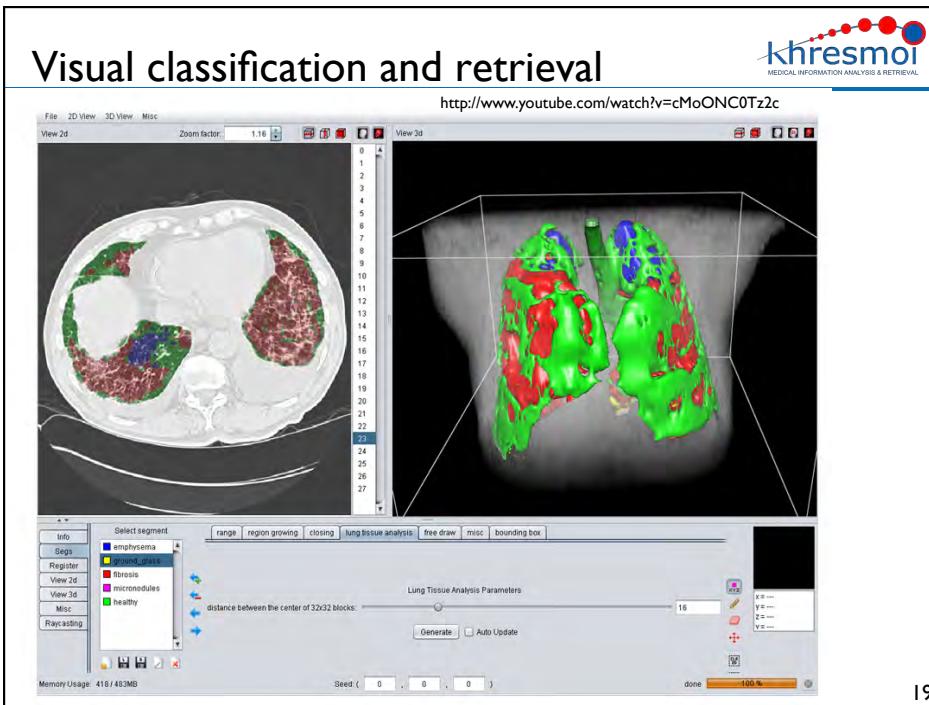
17

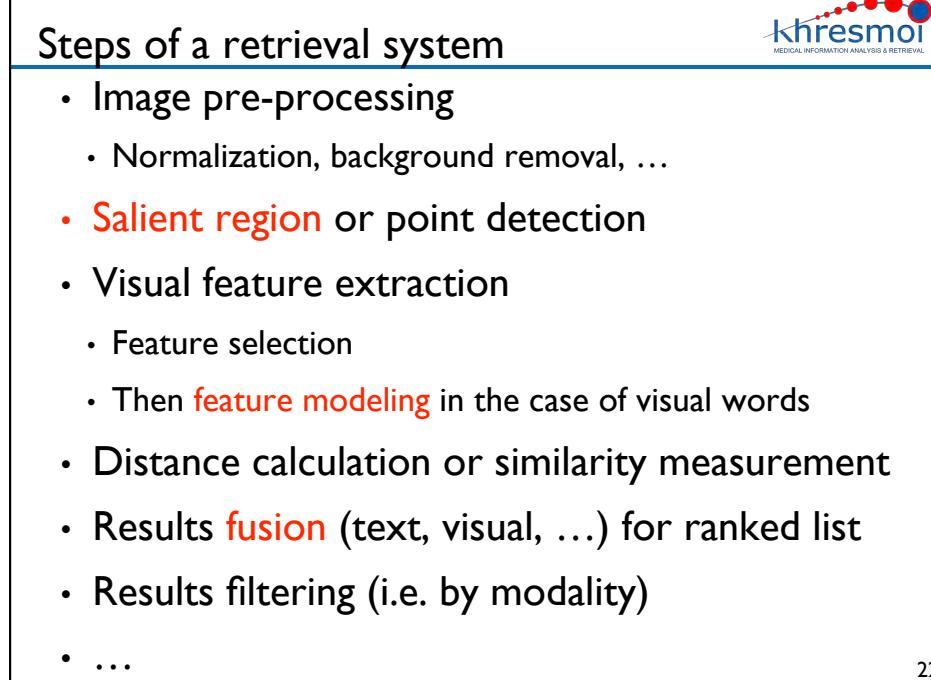
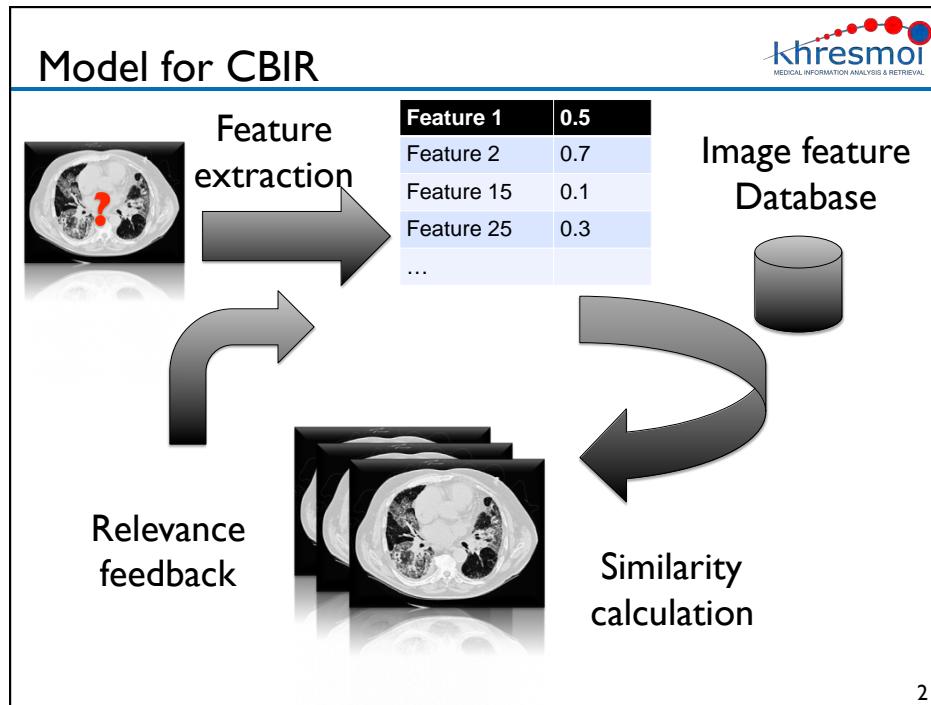
## Regions of interest

**khresmoi**  
MEDICAL INFORMATION ANALYSIS & RETRIEVAL



18





**A typical (old) interface**

The screenshot shows a web-based medical image retrieval system. At the top, there's a logo for "khresmoi MEDICAL INFORMATION ANALYSIS & RETRIEVAL". Below the logo, the title "A typical (old) interface" is displayed. The main area is a Firefox browser window titled "medGIFT online demo - Mozilla Firefox". Inside the browser, the interface is divided into sections:

- Query image:** A chest CT scan image is shown with a red circle around it.
- Diagnosis:** A text field next to the query image.
- Images result:** A grid of 10 images. The first image in the top row is circled in red. Labels with arrows point to specific features:
  - "Link to image database" points to the circled image.
  - "Relevance feedback" points to the circled image.
  - "Similarity score" points to the text "Similarity: 0.999999" below the circled image.
- Feedback:** A blue bar at the bottom of the results section contains the text "top" and "bottom".

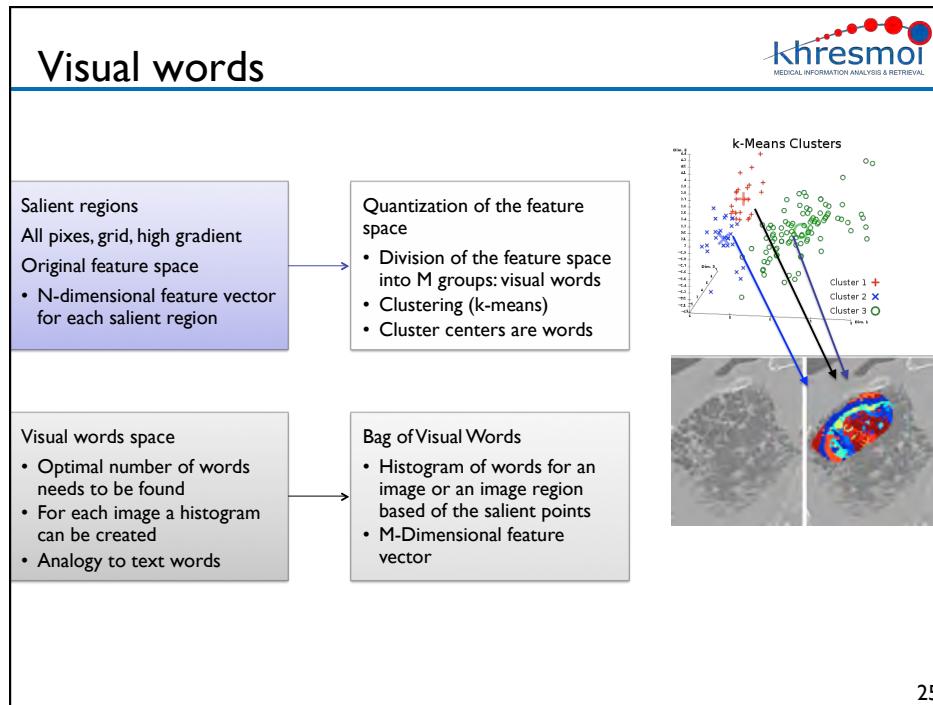
At the bottom right of the browser window, the number "23" is visible.

**Classifying visual features (Eakins)**

The diagram illustrates the classification of visual features into three levels, separated by a blue horizontal bar labeled "Semantic gap".

- Level 1: primitive features**
  - Color, texture, shape, spatial organization
- Level 2: derived features**
  - Individual objects or persons (Eiffel tower, Britney Spears)
  - Objects of a specific type (Volkswagen car)
- Level 3: abstract attributes**
  - High level reasoning about meaning and purpose
  - Emotional or religious significance
  - Find images of “suffering”

At the bottom right of the slide, the number "24" is visible.



**www.springerimages.com (3.3 mio)**

**SpringerImages**

HOME | ABOUT | FOR LIBRARIES | CONTACT US | HELP | REGISTER | LOG IN

lung ct with fibrosis | Search caption | Search | Show advanced options

**REFINE MY SEARCH**  
Search within these results | GO

**IMAGE TYPE**  
 Image    Table  
 Line Art    Video  
 Color    Black & White  
 Only accessible images (free + subscribed)

**SUBJECTS**  
Medicine & Public Health 40 | Imaging / Radiology 22 | Diagnostic Radiology 17 | Neuroradiology 16

**SEARCH RESULTS**  
41 RESULTS >> Save this search  
You searched for: lung ct with fibrosis REMOVE | Hide captions | RSS

Zoom: Sort by: Relevancy | Display: 25 | 50 | 100

	A 70-year-old male who had received radiation therapy.		Clinical and CT features in nine patients with focal		Schema of ROIs for tumour
	radiation fibrosis		A 60-year-old woman with focal interstitial fibrosis		Lung involvement by LOR in the same
	fibrosis		lungs		symptom
	Lung volumes.		Cervical and		CT images of the
	Table		Chest CT scan on		A 70-yearold

**MY IMAGE SETS**  
Never Set | View all images | Log In or register to save your favorite images and download them as high-quality PowerPoint or PDF files.

**SAVED SEARCHES**  
Save current search | Log In or register to save your search criteria.

Springer 2011. Produced by Current Medicine Group Ltd, a part of Springer Science+Business Media.  
Springer Address: 46127-105-58, Berlin, Germany

Home | About | Contact Us | Privacy Policy | Terms of Use | RSB | Help

26

**Goldminer.arrs.org (249,000 images)**

**Khresmoi**  
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

270 images      Filter: Modality - Age - Sex - Reset      1 - 10 · Next »

Recurrent malignancy. Image from a 74-year-old woman with history of middle lobe NSCLC cancer 4 years ... (74y F)  
[View full abstract](#) [Article in PDF-Version](#) [View all images / Visual search](#) [Similar articles](#)  
  
**CT**

CT-scan, mediastinal window. -- The bifurcation of the pulmonary artery appears distended and enlarged probably due to the ... (57y M)  
[View full abstract](#) [Article in PDF-Version](#) [View all images / Visual search](#) [Similar articles](#)  
  
**CT**

CT-scan, lung window. -- The subpleural and parabronchial parenchyma exhibits emphysema with multiple bullae. (57y M)  
[View full abstract](#) [Article in PDF-Version](#) [View all images / Visual search](#) [Similar articles](#)  
  
**CT**

CT-scan, lung window. -- Multiple central and peripheral traction bronchiectases are clearly demonstrated in both lungs.  
[View full abstract](#) [Article in PDF-Version](#) [View all images / Visual search](#) [Similar articles](#)  
  
**CT**

CT-scan, lung window. -- Lung window shows diffuse interstitial thickening of pulmonary apex in both lungs. (57y M)  
[View full abstract](#) [Article in PDF-Version](#) [View all images / Visual search](#) [Similar articles](#)  
  
**CT**

Axial non-contrast HR/CT images (right side) -- HR/CT image showing extensive fibrosis at the base of the right lung  
[View full abstract](#) [Article in PDF-Version](#) [View all images / Visual search](#) [Similar articles](#)  
  
**CT**

© 2010 American Roentgen Ray Society      AJR® Goldmine® is a registered trademark of the American Roentgen Ray Society

27

**medgift.hevs.ch/, demos (300,000 images)**

**Khresmoi**  
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

[Detection of emphysema progression in alpha 1-antitrypsin deficiency using CT densitometry](#)  
[Methodological advances](#)  
 2008-2-13 [Respiratory Research](#)  
 Computer tomography (CT) densitometry is a potential tool for detecting the progression of emphysema but the optimum methodology is uncertain. The level of inspiration affects reproducibility but the ability to adjust for this variable is facilitated by whole lung scanning methods. However, emphysema is frequently localised to sub-regions of the lung and targeted densitometric sampling may be more informative than whole lung assessment.

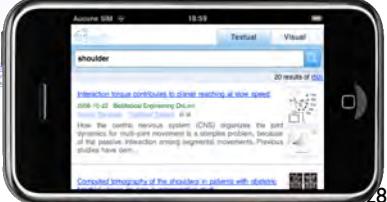


[View full abstract](#) [Article in PDF-Version](#) [View all images / Visual search](#) [Similar articles](#)  
 Authors: [David G Par](#) [Martini Seivoaka](#) [ChunQiri Deng](#) [Berend C Stiel](#) [Robert A Stockley](#)  
<http://respiratory-research.com>

[Bronchiolitis obliterans organizing pneumonia \(BOOP\) after thoracic radiotherapy for breast carcinoma](#)  
 2007-1-03 [Radiation Oncology](#)  
 Common complications of thoracic radiotherapy include esophagitis and radiation pneumonitis. However, it is important to be aware of uncommon post-radiotherapy complications such as bronchiolitis obliterans organizing pneumonia (BOOP). We report on two patients with carcinoma of the breast who developed an interstitial lung disease consistent with BOOP. BOOP responds to treatment with corticosteroids and the prognosis is generally good despite of the need for long-term administration of corticos...



[View full abstract](#) [Article in PDF-Version](#) [View all images / Visual search](#) [Similar articles](#)  
 Authors: [Robin Comellissen](#) [Suresh Saran](#) [Imogen E Antonisse](#) [Hauw Lier](#) [Youke K Aerts](#)  
<http://www.ro-journal.com>



28

[www.yottalook.com \(70,000 images\)](http://www.yottalook.com)

**Yottalook™** MEDICAL INFORMATION ANALYSIS & RETRIEVAL

[Yottalook Home](#)  
**BREAKING NEWS:** Yottalook to be acquired by [Montage Healthcare Solutions, Inc.](#)

Web Images Journals Books Anatomy  
   
 Detailed Thumbnails Only  Start Slide Show

Related terms: [pulmonary](#) - (powered by Yottalook)  
 Radiograph: 0 CT: [554](#) MR: 0 NM: 0 US: 0 [Show All](#)

No Subscription Mode: [ON](#) | [OFF](#)

1 to 10 results out of [554](#) for **lung CT with fibrosis**

Fig. 68 51-year-old man after lung transplant for cystic fibrosis. Patient had free air on routine chest radiograph and no abdominal symptoms and normal laboratory results-benign cause of pneumatois intestinalis (PI). Digital abdominal radiograph (A) and abdominal CT images (B and C) show free air (arrows, A and B) and...  
[View Article](#) | [Download image](#) | [Download article](#) | [What's This?](#)

Fig. 66 51-year-old man after lung transplant for cystic fibrosis. Patient had free air on routine chest radiograph and no abdominal symptoms and normal laboratory results-benign cause of pneumatois intestinalis (PI). Digital abdominal radiograph (A) and abdominal CT images (B and C) show free air (arrows, A and B) and...  
[View Article](#) | [Download image](#) | [Download article](#) | [What's This?](#)

Figure 8a. Focal intestinal fibrosis in a 40-year-old woman. (a) Thin-section CT image at the level of the superior segmental bronchus shows a 25-mm well-defined nodular ground-glass opacity with no solid component in the lower lobe of the left lung.  
[From Radiographics: Nodular Ground-Glass Opacity at Thin-Slice CT: Histologic Correlation and Evaluation of Change at Follow-up](#) | [View all images from this article](#)  
[Download image](#) | [Download article](#) | [What's This?](#)

Figure 14a. Pseudonodule in a 56-year-old woman who underwent a previous percutaneous lung biopsy. (a) Thin-section CT image obtained at the level of the aortic arch shows a 9-mm well-defined nodular ground-glass opacity (arrow) in the right upper lobe.  
[Feedback](#)

29

**khresmōi** MEDICAL INFORMATION ANALYSIS & RETRIEVAL

# Who searches for medical images and how?

30

## Background



- Health professionals' **image search behavior** has been subject to several surveys and log file analyses (MedLine, HONmedia, Goldminer)
- **Goldminer** log files of over 200'000 searches most comprehensive so far
- Khresmoi project performed a **survey among radiologists** to develop first prototypes
- Results change over time as user requirements change with the generation of physicians
  - Current medical students grow up with Google, Facebook and iPhones

31

## Methods of Khresmoi survey



- Survey among radiologists, mainly in Geneva and Vienna University hospitals
- Paper and electronic version
- Survey filled in with a person explaining the goals
- Survey took approximately 1 hour, research/clinical /teaching work separated with same questions
- **26 radiologists answered, 13 from Austria, 9 from Switzerland**, form only sent on invitation
- 17 males, 9 females, mainly 26-35 years, few years of experience

32

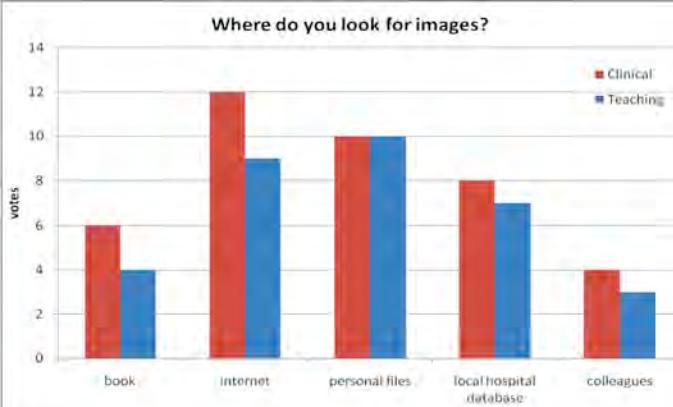
## Reasons for image search



- Clinicians: Information on unclear cases, illustrate presentations, **differential diagnosis** were most frequent
- Teaching: Find **similar cases**, for example an easy, a medium and a tricky case for the same disease, problem-based learning
- Problem-based learning requires increased search skills for the students

33

## Where people search



- The **Internet** replaces books and colleagues
- Personal files are not always optimal

34

**Determination of relevance**

Category	Clinical (red)	Teaching (blue)
image properties	6	4
colleagues	2	0
description	3	0
experience	11	7
quality	4	4
other	4	1

- Experience determines relevance
- Often additional **proof** such as biopsies or other clinical tests are requested

35

**Success rates**

Success Percentage Range	Clinical (red)	Teaching (blue)
>81%	8	2
61-80%	7	2
41-60%	5	3
21-40%	3	0
<20%	1	0

Reason for unsuccessful search	Clinical (red)	Teaching (blue)
no search possible	1	1
too general	1	1
too many / race	5	5
no time	1	1
other	1	2

- For teaching success rates are higher
- Clinical work might have less well defined tasks, average of success at 60%

36

**Search time**

**khresmoi**  
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

The figure consists of two side-by-side bar charts. The left chart is titled 'Image search success time' and the right is titled 'Image search failure time'. Both charts have 'percentage' on the y-axis (0 to 60) and 'time' on the x-axis with categories: <3min, 3min, 5min, 10min, 15min, >15min. The legend indicates three categories: Clinical (blue), Teaching (red), and Research (green).

Time Category	Clinical (%)	Teaching (%)	Research (%)
<3min	10	15	20
3min	35	25	20
5min	30	45	20
10min	25	10	15
15min	10	10	20
>15min	15	15	40

Time Category	Clinical (%)	Teaching (%)	Research (%)
<3min	10	10	15
3min	15	10	15
5min	25	25	20
10min	30	30	25
15min	15	15	10
>15min	30	25	55

- 70% of successful searches less than ten minutes
- Failure often after over 15 minutes
- Less time available for clinical search than research/teaching
- Could a **faster and more targeted** image search system help?

37

**Useful additions**

**khresmoi**  
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

- Search by
  - pathology (13 times)
  - modality (10 times)
  - patient demography (6 times)
  - **similar images** (8 times)
- Other comments:
  - Multilingual retrieval
  - Pathology classification (using **ontologies**)
  - Search in 3D data
  - Confidence in the diagnosis

38

**A perfect search system**

**desired search system input**

Category	votes
image	10
keywords	10
extra information	1
feedback loops	2

**desired search system output**

Category	votes
image	10
differential diagnosis	2
text	4
references	2

- Many free comments
  - Like Google but DICOM and text
  - Structure information and confidence in diagnosis
  - Search by **regions of interest**
  - Social networking, comments of other physicians
  - Search for **similar cases**
  - **Quantification** of structures (size, volume)

39

**Goldminer log files**

- Monday 12h15, talk on the analysis
- 25'000 consecutive queries

number of query pairs

number of results retrieved	add term(s)	remove term(s)	replace term(s)
0	~100	~1300	~1800
1-10	~100	~200	~500
1-40	~200	~250	~750
>40	~850	~100	~400

query modification type

query modification type	1 term	2 terms	3 terms	4 terms	≥5 terms
term(s) added	~900	~150	~50	~10	~5
term(s) removed	~1200	~300	~100	~20	~10
term(s) replaced	~2400	~450	~150	~30	~15

40



# Combining text and visual search

41

## Background



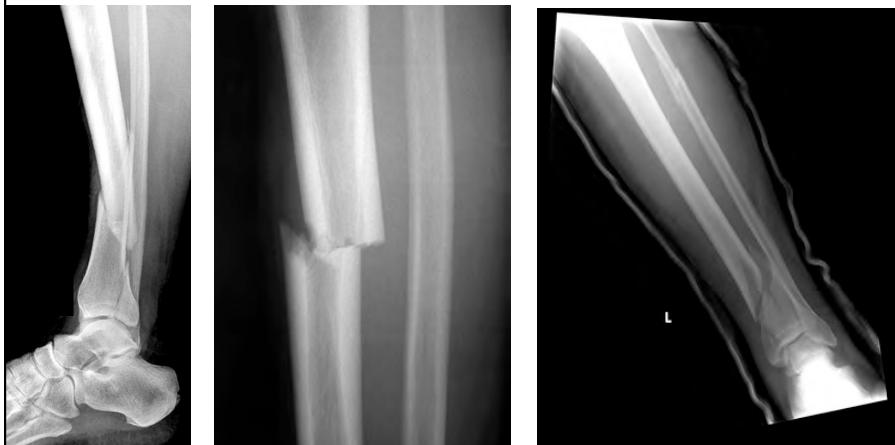
- ImageCLEF image retrieval benchmark has been run each year since 2004
- 12-20 research groups compare their tools and approaches on the same tasks and DBs
  - Visual, textual and combined approaches are used
  - Multilingual approaches are also possible
- Sometimes visual, sometimes textual and sometimes mixed approaches perform best
  - No clear outcome
  - Combination of results can be delicate, instable

42

## Example search topic



- Show me x-ray images of a tibia with a fracture.
- Zeige mir Röntgenbilder einer gebrochenen Tibia.
- Montre-moi des radiographies du tibia avec fracture.



43

## Visual vs. semantic vs. mixed searches



- Experts can **predict** what **type of technique** will most likely perform best
  - Can this prediction be modeled automatically?
  - If results were visually relatively homogeneous **visual** search can work
    - Same anatomic region, same view, same modality
  - If results are expected to be very different (no modality given), **text** would work best
  - **Combinations** often work best when some common aspects but some variability as same modality
- Visual and text retrieval are complementary

44

## Early vs. late fusion



- **Early fusion**
  - **Feature** spaces are directly combined, so visual features and textual words treated in the same way
  - Number of features needs to be similar to avoid bias
- **Late fusion**
  - **Results** of systems are combined, not features
  - Each system can have a varying number of features
- For text/visual combinations late fusion is often simpler to employ and works better
  - When using visual words both could be used

45

## Score based vs. rank-based fusion



- **Score-based fusion**
  - Score of the single systems is used for the combination of the results sets
  - Score needs to be normalized, potentially to have similar characteristics
- **Rank-based fusion**
  - The rank of an element is used to calculate fusion
    - Can be linear or logarithmic or in another form
  - Avoids the bias that very differing results sets of system can have
  - Often have better results when visual and textual systems are combined

46

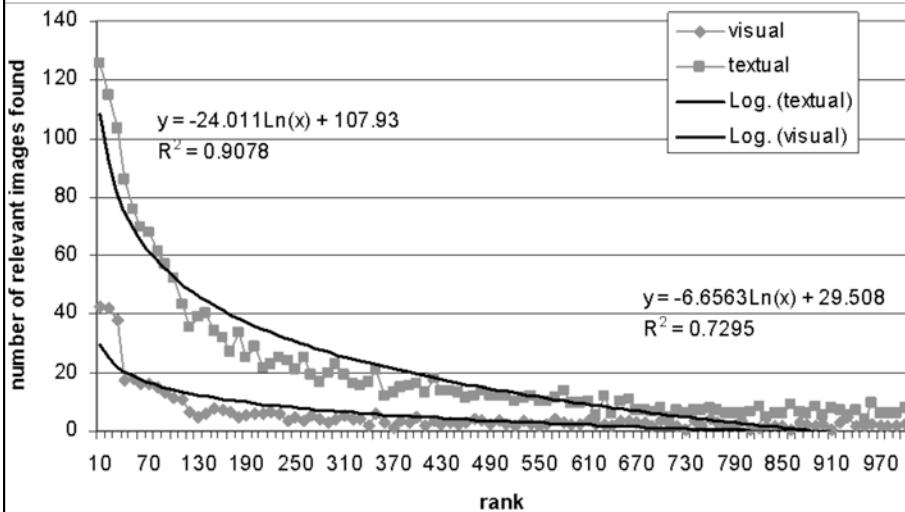
## Types of fusion techniques



- Many types of fusion techniques exist
  - combSUM :  $V_{\text{combSUM}}(i) = \sum_k V_k(i)$
  - combMAX :  $V_{\text{combMAX}}(i) = \text{argmax}(V_k(i))$
  - combMNZ :  $V_{\text{combMNZ}}(i) = f(i) \cdot V_{\text{combSUM}}(i)$   
Where  $f(i)$  is the frequency of image  $i$  in the results
- At the ICPR 2010 conference a competition on fusion techniques was organized using the best ImageCLEF runs
  - Rank-based techniques using logarithmic decrease performed best in a variety of different approaches

47

## Distribution of relevant documents



48

## Other combinations



- Modality detection (using visual techniques or text of the captions) can work very well (80+%)
- Allowing to select the **target modality** can improve image search
  - Tests with all runs of ImageCLEF 2009
  - Many search engines allow for this such as Goldminer
    - This can be used for tabbed browsing as well
- **Exclusion criteria** for images can be chosen based on the text
  - Age group, gender, ...

49

## Combinations for **case-based** retrieval



- **Mix** of free text, structured data, images, and many other forms
- **Interactions** of the data can vary strongly between patients and diseases, also over time
- More **complex combinations** for images need to be found
  - Match images between case to match for similarity
  - Currently text is better than most fusions
- Case description including images without diagnosis, find images for differential diagnosis

50

## Some ImageCLEF lessons learned



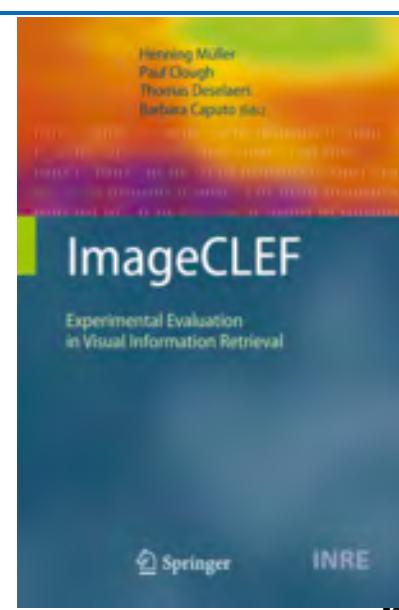
- Text retrieval techniques are stable and deliver good results (i.e. Lucene is above average)
- Visual has had less evolution than text retrieval
  - GIFT (old!) has still relatively good results
    - Semantic gap is very present
  - Visual words-based approaches can be much better when using high quality training data
- Interactive retrieval can improve visual retrieval
- Many features combined deliver best results
- Mapping of images and text to ontologies can help
  - Improve semantic retrieval

51

## Attention: advertisement



- ImageCLEF book
- All on image retrieval
  - Methods of evaluation
  - Task overviews
  - Participant reports
    - The best techniques
    - Industrial requirements
  - Industry perspectives
  - Specific techniques such as fusion



52



## Challenges for search?

53

### BIG data



- How can **scalability** be assured when treating extremely large amounts data?
  - 250'000 images per day in Geneva ...
    - 150 TB of images in Catalonia archive
  - Extremely large scales allow solving many new problems
    - Rare diseases
    - Sufficiently large training data sets
- **Hadoop/Mapreduce** as is also used by Google, Yahoo or Facebook
- Use of **cloud** computing
  - Costs, confidentiality, also bandwidth ...

54

## Confidentiality



- Can patient records be made **available**?
  - Maybe partly, anonymized, only internally?
  - Could the data warehouse be used for this
  - Secondary use of data
- Availability for data mining not a specific very limited scenario (as ethics committees request)
- Can **interoperability** be assured using the same semantic standards
- How to link the literature with specific cases
  - Images not in the same quality, much more than just case descriptions, ...

55

## Search for medical cases not images



- Combine **several data sources**
  - Importance of each source is not fully clear
    - Interaction between content importance is complex
  - Different media from free text, structured data to semantics, signals or images including 3D, 4D
- Some **data sources might be missing**
  - Questions not asked, not documented, errors
  - How to deal with missing data
- **What is a case exactly?**
  - Limited period of time? Or on a patient basis including old data?

56

## Diversity in the results



- Having a list of almost identical texts as result is not useful
  - Google filters out near duplicates
- **Consistency vs. diversity** have limits
- Some search systems cluster results and then present each cluster in a first step
- Diversity can favor **data exploration** and user relevance feedback
- Understand links between documents and content in the results

57

## Retrieval from social networks



- People share data in social networks
  - Blogs, facebook, ...
  - Sometimes more than any physician would share
    - People with rare diseases are sometimes desperate ...
- People can **comment** on data
- Goals of comments are not always clear
  - **Spam** can create problems
- Some metadata is available, so free text, semi-structured metadata, images
  - Can semantics help with this?

58

**A medical blog on movement disorder**

MONDAY, MARCH 7, 2011  
A pleasant genetic appointment!



Bertrand takes his official ring bearer training VERY seriously!

Bertrand's "good news" roll continues! At today's genetic/metabolic follow-up, Bertrand was found to be "improved", much to the puzzlement and pleasure of his geneticist.

WEIGHT CHALLENGE

03/16/11	- 36lbs. 10oz.
04/11/11	- 37lbs. 10oz.
03/14/11	- 38lbs. 10oz.
03/07/11	- 39lbs. 10oz.

STANDER CHALLENGE

Since April 1	- 1-2 hrs. daily!
04/03/11	- 53 min.
03/29/11	- 0 min.
03/28/11	- 23 min.
03/27/11	- 105 min.
03/26/11	- 0 min.
03/25/11	- 36 min.

ABOUT BERTRAND

Born in December 2007, Bertrand is a charming, serious, young man. He lives in Salt Lake City, UT and has global developmental delays (0-6 months-old), leukodystrophy, intractable multifocal epilepsy (Doose Syndrome), peripheral neuropathy, liver fibrosis, gastosmesophagial reflux

59

**Patientslikeme**



patientslikeme®

your profile

60

## Visceral project



**khresmoi**  
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

- VISual Concept Extraction challenge in RAdioLoGy
  - Most likely as a MICCAI workshop in 2013 and 2014
- Two challenges
  - Identify organs in 3D data sets
  - Find similar cases using 3D data and radiology reports
- Very large amounts of data (~10 TB)
  - Data distribution via the cloud
    - Participants will get a virtual machine for free
  - Creation of a gold and silver standard

61

## LinkedIn Khresmoi group

**khresmoi**  
MEDICAL INFORMATION ANALYSIS & RETRIEVAL

### Most Popular Discussions



Why is medicine often not evidence based? (Ben Goldacre)

Many reasons are given, but the following paragraph applies to what KHRESMOI is aiming to do:

"while we do make an effort with ...

posted February 2, 2011

SarathChandra Kambhatla 1 day ago · SarathChandra likes this.

See more \*



30% of world storage is estimated to be medical imaging

According to an EU report by now 30% of worldwide data storage is estimated to be medical imaging, more to be read in <http://cordis.euro...>

posted 27 days ago

All HOSSEINZADEH VAHID 18 days ago · Ali likes this.

See more \*



Big data - and the access to them for scientists

The fact that companies often keep big data sets private but publish with them causes some problems for science.

works on a common policy for data sharing and citing data resources.

Like · Add comment · 6 days ago



Henning Müller started a discussion: WebMill - a tool for collaborative labeling of medical images

Like · Add comment · 6 days ago

See all updates \*

### Ads by LinkedIn Members



Online MBA in Leadership

Top British Online MBA in Leadership and Sustainability - Free catalogue!



Top-Ranked Executive MBA

Fast-track your career! Study in Prague, Pittsburgh & São Paulo. Learn more



Project Management Skills

Become a certified Project Manager now!

### Top Influencers This Week



Allan Hanbury

62

## A few references



- Eakins J, Graham M, Content-based Image Retrieval, JTAP-039, JISC report, 1999.
- Enser P, Visual image retrieval: seeking the alliance of concept-based and content-based paradigms, Journal of Information Science, 2000.
- Müller H, Michoux N, Bandon D, Geissbuhler A. A re-view of content-based image retrieval systems in medical applications-clinical benefits and future directions. Int J Med Inform. 2004 Feb ;73(1):1-23.
- Müller H, Despont-Gros C, Hersh W, Jensen J, Lovis C, Geissbuhler A, Medical professionals' image search and use behaviour, pages 24-32, Medical Informatics Europe, Maastricht, The Netherlands, 2006.
- Müller H, Kalpathy-Cramer J, Kahn CE Jr., Hatt W, Bedrick S, Hersh W. Overview of the ImageCLEFmed 2008 Medical Image Retrieval Task, Springer Lecture Notes in Computer Science 5706, 500-510, 2009.
- Smeulders AVM, Worring, Santini S, Gupta A, Jain R. Content-Based Image Retrieval at the End of the Early Years, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22(12), pages 1349-1380, 2000.
- Vasconcelos N, Lippman A, A Probabilistic Architecture for Content-based Image Retrieval, IEEE International Conference on computer vision and pattern recognition, volume 1, pages 216-221, 1999.
- ...

63

## Conclusions



- Medical information **search and access** is an important technique in medicine
  - And this includes images!!
- Image information is most often complementary to text
- **Visual information** such as regions of interest can be used to **formulate queries**
  - Radiologists request this increasingly instead of searching in books and discussing with colleagues
- There is still much to be learning for combining visual and textual techniques

64

## Questions?



- More **information** available from
  - <http://www.imageclef.org/>
  - <http://khresmoi.eu/>
  - <http://medgift.hevs.ch/>
  - **Contact** [henning.mueller@hevs.ch](mailto:henning.mueller@hevs.ch)



65