

# Enterprise Search in Medical Data

Allan Hanbury



TECHNISCHE  
UNIVERSITÄT  
WIEN

Vienna University of Technology

Henning Müller

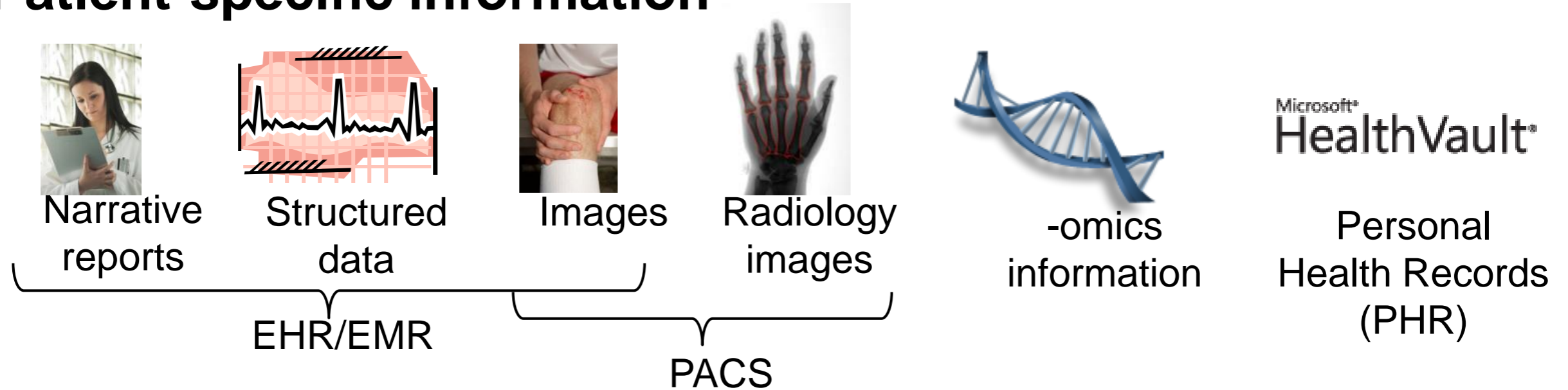
**Hes·SO**

Haute Ecole Spécialisée  
de Suisse occidentale

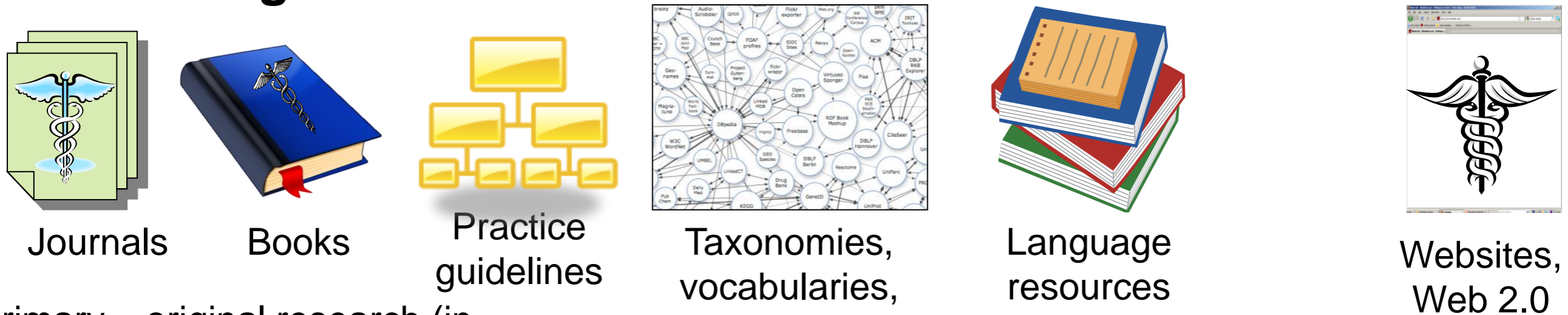
Fachhochschule Westschweiz

University of Applied Sciences  
Western Switzerland

## Patient-specific information



## Knowledge-based information

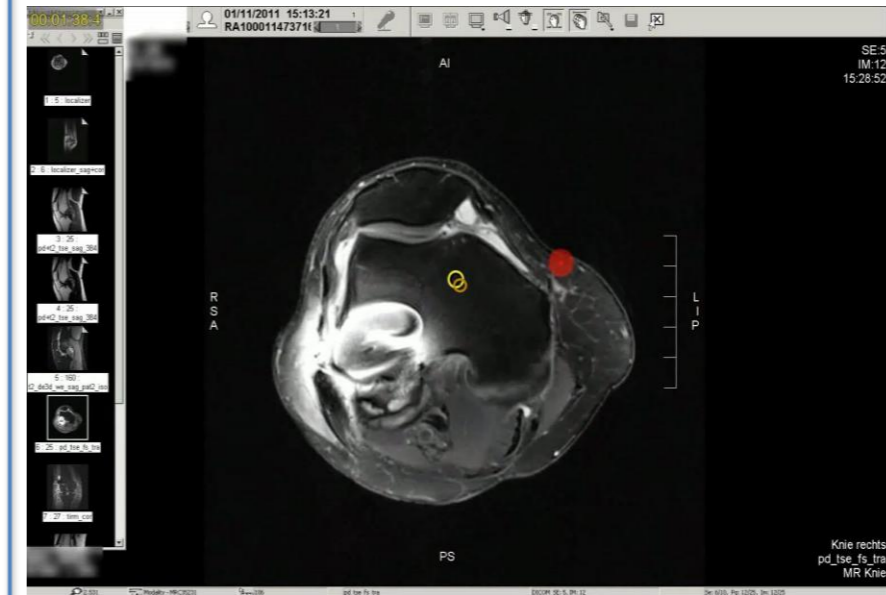
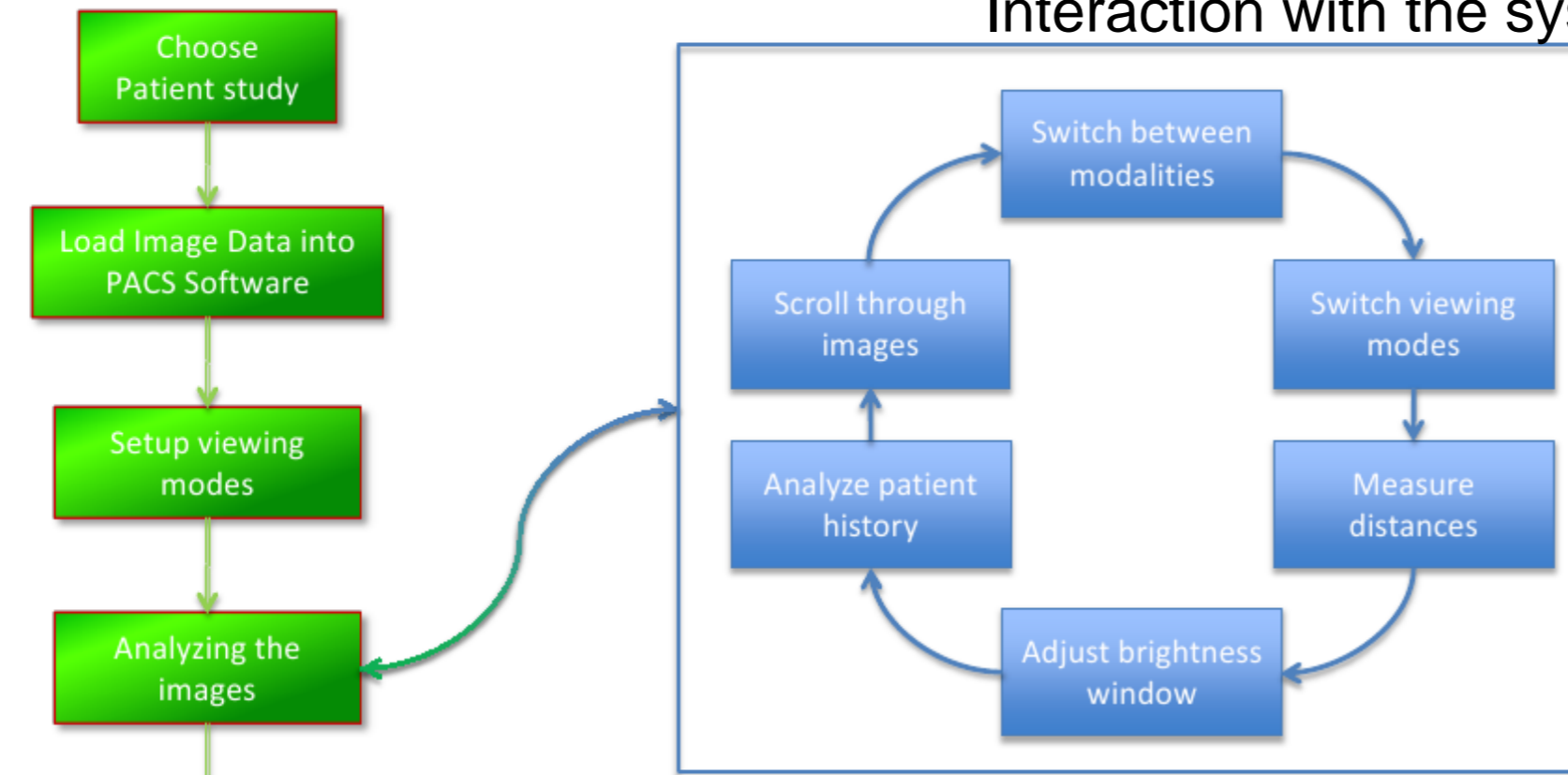


- Primary – original research (in journals, books, reports, etc.)
- Secondary – summaries of research (in review articles, books, practice guidelines, etc.)

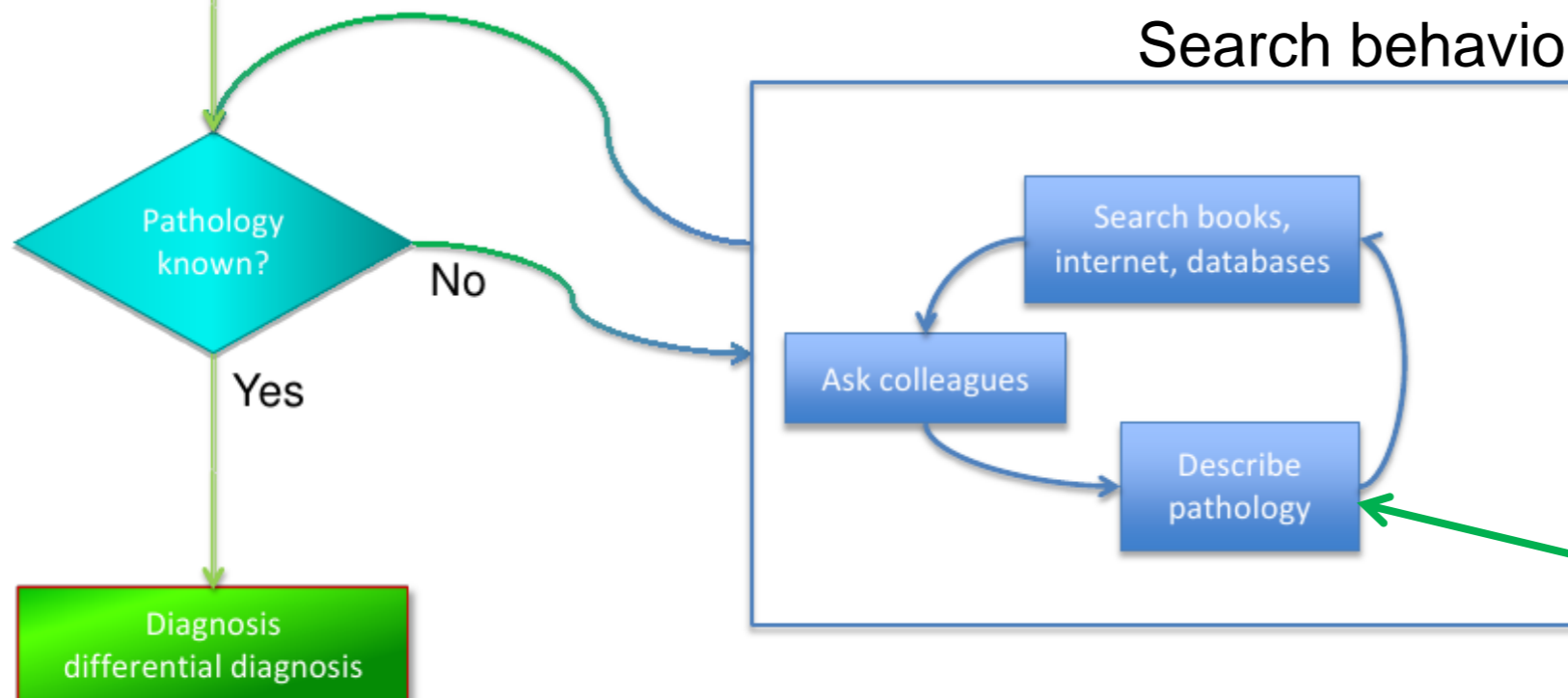
- Health records are currently almost always only accessible by patient ID
- Could even more on optimum treatment be learned by searching and mining the huge archives of health records?
  - Example: Kaiser Permanente and Vioxx
- Feufel et al. “believe that it is **inefficient, and in some cases unethical, to store [population-based data] without installing mechanisms to allow access and publication** in appropriate and useful ways.” (Better Doctors, Better Patients, Better Decisions: Envisioning Health Care 2020, MIT Press, 2011)

# Use Case: Radiology

## Interaction with the system

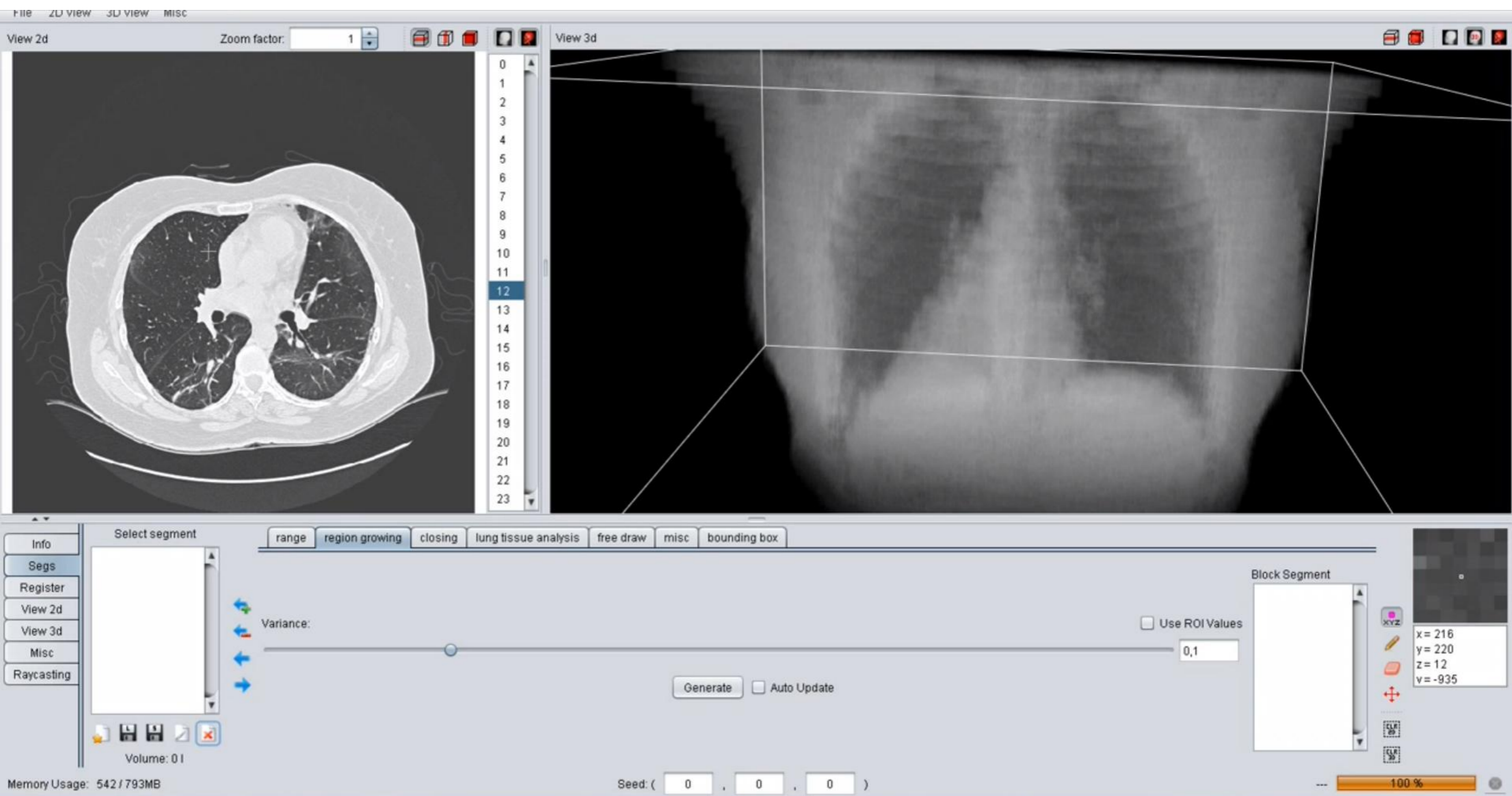


## Search behaviour



Needed as image search not available

# Radiology Example



File 2D view 3D view Misc

View 2d Zoom factor: 1 View 3d

0  
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23

Info  
Segs  
Register  
View 2d  
View 3d  
Misc  
Raycasting

Select segment  
lungobello

range region growing closing lung tissue analysis free draw misc bounding box

Lung Tissue Analysis Parameters

distance between the center of 32x32 blocks: 4

Generate Auto Update

Volume: 0,3661 l

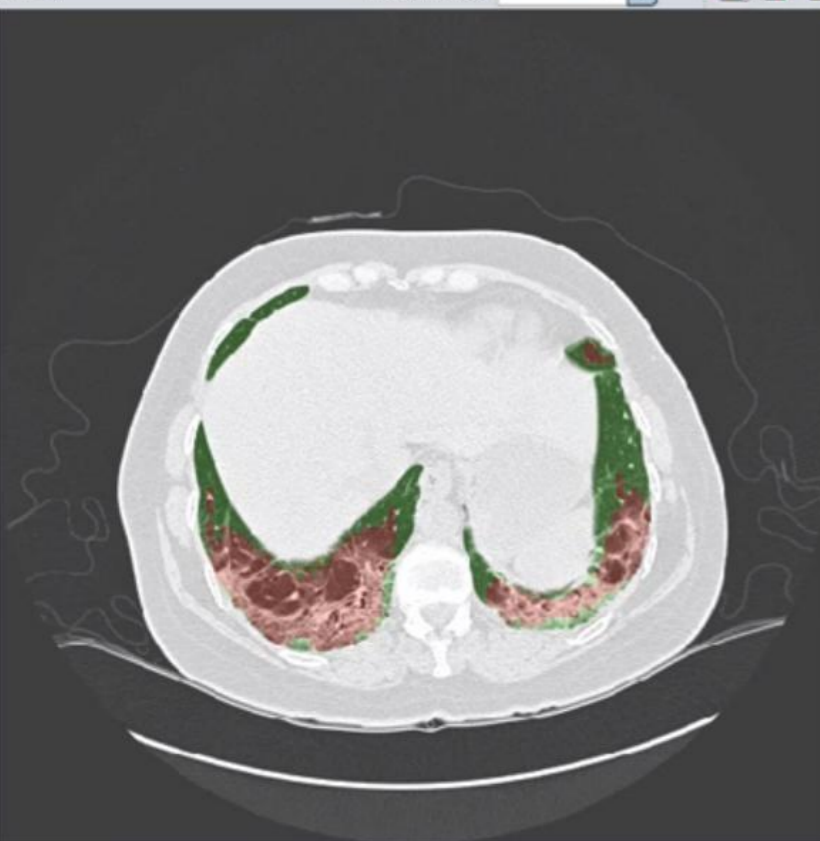
Memory Usage: 471 / 561MB

Seed: ( 0 , 0 , 0 )

100%

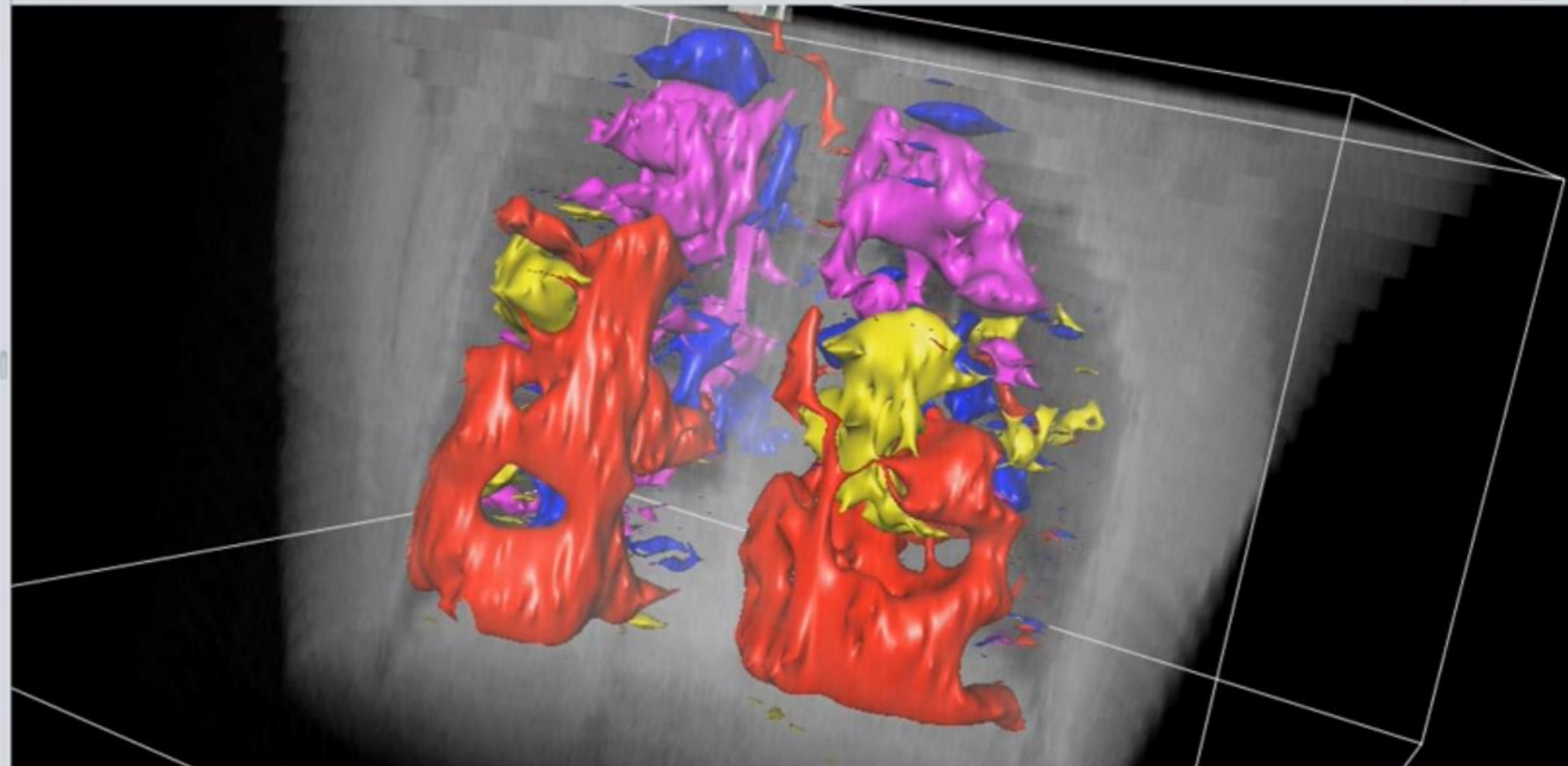
File 2D view 3D view Misc

View 2d Zoom factor: 1



0  
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23

View 3d



segment rendering texture renderer slice planes lights misc

Info  
Segs  
Register  
View 2d  
View 3d  
Misc  
Raycasting

Segment Visualization

- healthy
- emphysema
- ground\_glass
- fibrosis
- micronodules

Options

- Show
- Points
- Marching Cube
- Texture 2d
- Texture 3d

x= ---  
y= ---  
z= ---  
v= ---

Memory Usage: 359 / 561MB Seed: ( 0 , 0 , 0 ) 100 %

http://medgift.unige.ch/talisman/applications/ria/

Fichier Edition Affichage Favoris Outils ?

Favoris Sites suggérés Hotmail Galerie de composants W...

Talisman Ria Application

fr\_FR

Patient list - Talisman RIA Interface

Text to search... Search

Detail

Patient data list

	similarity
1 Diagnosis : <i>Fibrose</i> , Diagnosis : <i>pulmonary_fibrosis</i> , Diagnosis reliability : 100 % , Age : 64 years old , Gender : <i>m</i> , Smoking : <i>true</i> , Quantity : 60 UPA,	98.9815%
1 Diagnosis : <i>fibrose pulmonaire idiopathique</i> , Diagnosis : <i>pulmonary_fibrosis</i> , Diagnosis reliability : 98 % , Age : 66 years old , Gender : <i>m</i> , Smoking : <i>true</i> , Quantity : 60 UPA, Bronchoscopy transbronchial : <i>interstitial inflammation</i> ,	98.9815%
1 Diagnosis : <i>Fibrose de la pyramide basale du lobe inférieur gauche</i> , Diagnosis : <i>pulmonary_fibrosis</i> , Diagnosis reliability : 100 % , Age : 66 years old , Gender : <i>m</i> , Smoking : <i>true</i> , Quantity : 25 UPA, remarks : 1 pack/day currently , Lung : <i>interstitial fibrosis</i> ,	98.9815%
1 Diagnosis : <i>Fibrose interstitielle</i> , Diagnosis : <i>pulmonary_fibrosis</i> , Diagnosis reliability : 75 % , Age : 63 years old , Gender : <i>f</i> , Bronchoscopy transbronchial : <i>interstitial inflammation</i> ,	97.963%
1 Diagnosis : <i>Fibrose pulmonaire idiopathique</i> , Diagnosis : <i>pulmonary_fibrosis</i> , Diagnosis reliability : 100 % , Age : 62 years old , Gender : <i>f</i> ,	96.9445%
1 Diagnosis : <i>Fibrosis</i> , Diagnosis : <i>pulmonary_fibrosis</i> , Diagnosis reliability : 100 % , Age : 70 years old , Gender : <i>m</i> , Smoking : <i>true</i> , Quantity : 60 UPA, quitted since : 2 years,	94.9075%
1 Diagnosis : <i>pneumopathie interstitielle fibrosante</i> , Diagnosis : <i>pulmonary_fibrosis</i> , Diagnosis reliability : 100 % , Age : 70 years old , Gender : <i>m</i> , Smoking : <i>true</i> , Quantity : 60 UPA, Bronchoscopy transbronchial : <i>interstitial inflammation</i> ,	94.9075%
1 Diagnosis : <i>Fibrose pulmonaire secondaire à une sclérose systémique</i> , Diagnosis : <i>pulmonary_fibrosis</i> , Diagnosis reliability : 100 % , Age : 59 years old , Gender : <i>f</i> ,	

1-8 of 124

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 > >>

Terminé Internet 100%



http://medgift.unige.ch/talisman/applications/ria/

Fichier Edition Affichage Favoris Outils ?

Favoris Sites suggérés Hotmail Galerie de composants W...

Talisman Ria Application

Patient data detail view - Talisman RIA Interface fr\_FR

**Patient data** YaDIV

- General info
  - Diagnosis : pulmonary\_fibrosis
  - Age : 70 years old
  - Gender : m
  - Profession : former ambassador
  - Disease duration : 3 years
  - Medication
  - Findings on physical exam
  - Past medical history
  - Occupational history
  - Environmental exposures
  - Host risk factors
  - Laboratory tests
  - Oximetry tests
  - PST (pulmonary function testing)
    - Smear sputum tests :
  - BAL (bronchoalveolar lavage)
  - Biopsy
  - Remarks : Radiology-based diagnosis

Images series Button

Slice thickness	Number of slices	Date of study
1	31	Mon Jan 18 20:09:02 GMT+0100 2010

Terminé Internet 100%

- Search for “similar” patients
  - Similar age, symptoms, treatment, etc.
  - Also potentially similar radiology image features/pathology
  - Learn how these patients were treated and what the outcome was
    - Especially important for rare diseases
  - Fits well into the common practice of consulting with colleagues

# Patientslikeme



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

05/16/11 - 36lbs. 10oz.  
04/11/11 - 37lbs. 10oz.  
03/14/11 - 38lbs. 10oz.  
03/07/11 - 39lbs. 10oz.

#### STANDER CHALLENGE

Since April - 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, gastroesophageal reflux

# (More) Search Tasks in Medicine

- Linking patient records to
  - Guidelines/clinical pathways
  - Literature
- Search within medical records
  - Many medical records are chronological collections of documents
  - It is difficult to find the pertinent information

# Challenges

- Personal health records must be anonymised before being displayed as search results
- The targeted end users potentially don't want the solution!



**ELGA stellt Sie vor den anderen bloß!**

Wollen Sie tatsächlich zum gläsernen Patienten werden?

ELGA ist eine elektronische Krankheitsakte, mit der mehr als 100.000 Personen Zugang zu Ihren Krankheitsdaten bekommen können. Diese Daten könnten dann ganz leicht in die Hände möglicher Arbeitgeber gelangen.

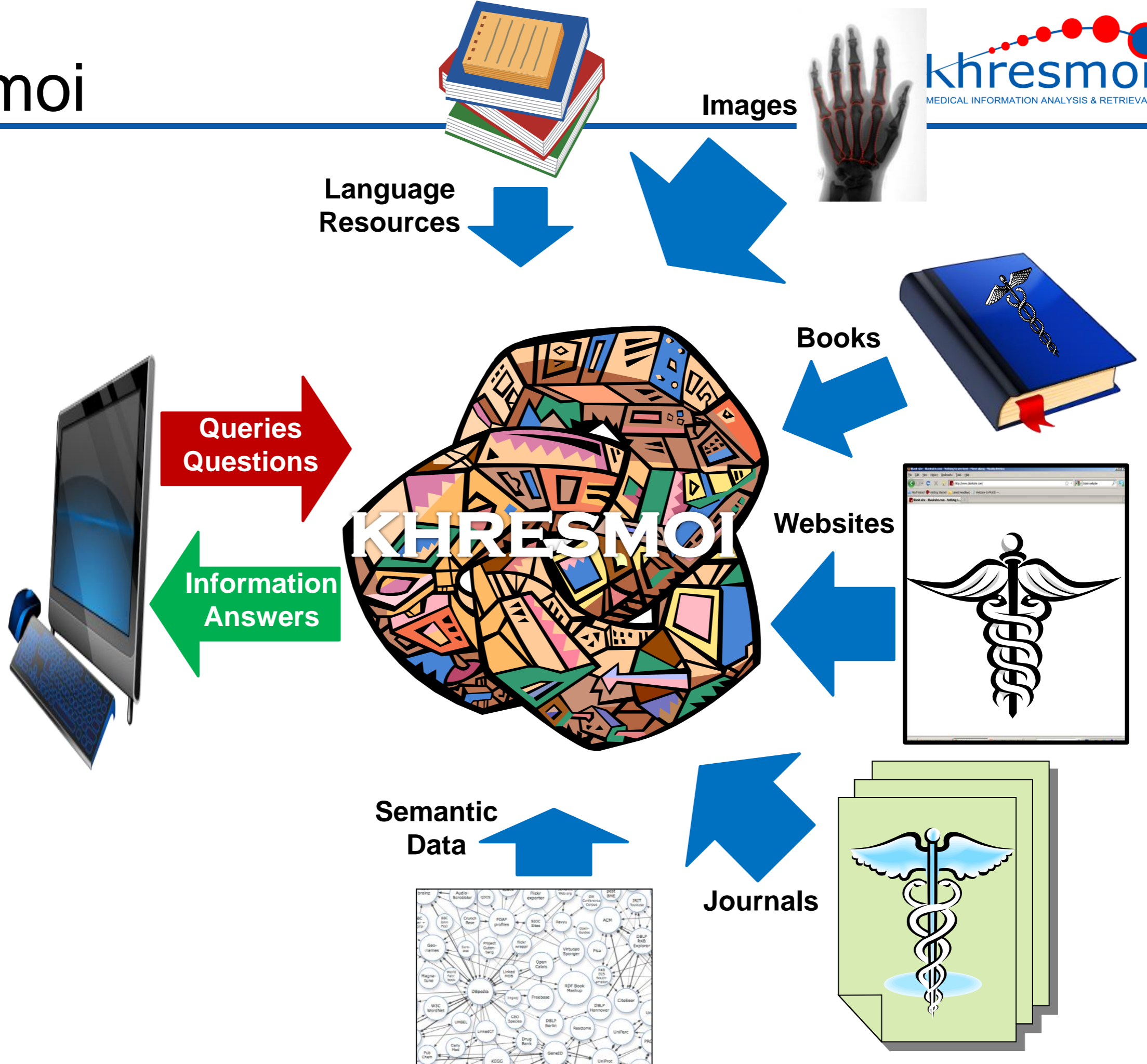
ÄRZTEKAMMER FÜR WIEN

# More challenges

- Data quality and provenance
- Health care providers are not used to entering data that will have secondary use
- The time aspect is of central importance
  - Especially time relative to events – e.g. how long has a person been on a drug
- Information should be delivered at the point of care

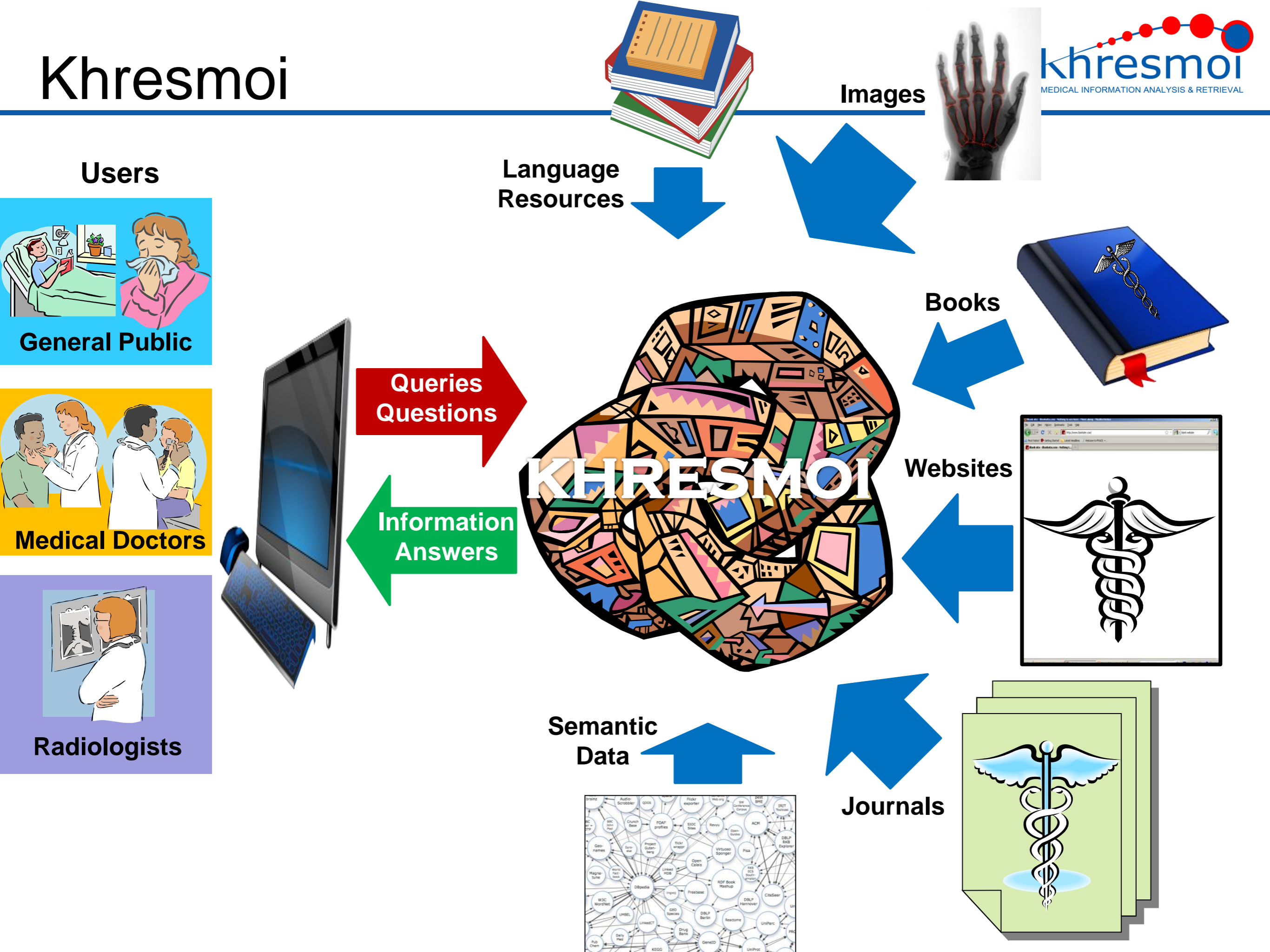


# Khresmoi





# Khresmoi



# Technologies being developed

- Semantic annotation and search
- Medical image search
  - Including high-dimensional image search
- Cross-lingual search and machine translation
- Health knowledge bases
- User tests to validate technologies

- Hospitals have huge archives of patient data that could be put to effective secondary use
  - Stored in many formats
  - More free text than structured data
  - Multimodal information
  - Not hyper-linked
- Putting content and context together
- Privacy preservation is the biggest challenge