A Formative Evaluation of a Comprehensive Search System for Medical Professionals

Veronika Stefanov, Alexander Sachs, Marlene Kritz, Matthias Samwald, Manfred Gschwandtner, and **Allan Hanbury**

Overview

- Search system for medical professionals, under development

- Iterative development and testing

- Iterative test development

- "Black-box" formative evaluation
  - simulated work tasks
  - real users

Motivation

- Medical doctors need rapid and accurate answers

    - A search of MEDLINE takes on average 30 minutes [1]

    - Doctors have on average 5 minutes available for a search [2]

    - 40% of searches do not yield the information required [3]

1. Hersh, W.R., Hickam, D.H.: How well do physicians use electronic information retrieval systems? a framework for investigation and systematic review. JAMA 280(15) (Oct 1998) 1347-52 PMID: 9794316.

2. Hoogendam, A., Stalenhoef, A.F.H., de Vries Robb, P.F., Overbeke, A.J.P.M.: Answers to Questions Posed During Daily Patient Care Are More Likely to Be Answered by UpToDate Than PubMed. J Med Internet Res 10(4) (2008)

3. Ely, J.W., Oshero, J.A., Maviglia, S.M., Rosenbaum, M.E.: Patient-care questions that physicians are unable to answer. JAMA 14 (2007) 407-414

Khresmoi

- 4 year EU-funded Project

- Currently in its 4th year

Khresmoi – multilingual semantic search of medical text and images

Niraj Aswani[i], Thomas Beckers[g], Erich Birngruber[k], Célia Boyer[j], Andreas Burner[k], Jakub Bystroň[h], Khalid Choukri[d], Sarah Cruchet[j], Hamish Cunningham[i], Jan Dědek[h], Ljiljana Dolamic[j], René Donner[k], Sebastian Dungs[g], Ivan Eggel[b], Antonio Foncubierta[b], Norbert Fuhr[g], Adam Funk[i], Alba Garcia Seco de Herrera[b], Arnaud Gaudinat[b], Georgi Georgiev[e], Julien Gobeill[b], Lorraine Goeuriot[f], Paz Gomez[c], Mark Greenwood[i], Manfred Gschwandtner[l], Allan Hanbury[a], Jan Hajič[h], Jaroslava Hlaváčová[h], Markus Holzer[k], Gareth Jones[f], Blanca Jordan[c], Matthias Jordan[g], Klemens Kaderk[k], Franz Kainberger[k], Liadh Kelly[f], Sascha Kriewel[g], Marlene Kritz[l], Georg Langs[k], Nolan Lawson[j], Dimitrios Markonis[b], Iván Martinez[c], Vassil Momtchev[e], Alexandre Masselot[j], Hélène Mazo[d], Henning Müller[b], João Palotti[a], Pavel Pecina[h], Konstantin Pentchev[e], Deyan Peychev[e], Natalia Pletneva[j], Diana Pottecher[c], Angus Roberts[i], Patrick Ruch[b], Alexander Sachs[l], Matthias Samwald[a], Priscille Schneller[d], Veronika Stefanov[a], Miguel Angel Tinte[c], Zdeňka Urešová[h], Alejandro Vargas[j], Dina Vishnyakova[b]

diabetes

Results: 250

Relevance ▼ | Enter filter terms

Choose filters

No Image modality selected ▼ | No Classifications selected ▼ | No Media type selected ▼

Ungrouped
▼ **By Classifications**
   All Items (250)
   Alcohol (3)
   Babies & Children (26)
   Cancer (1)
   Cardiovascular disease (
   Dictionary Definitions (42)
   Disease (2)
     Risk factors (1)
     Treatment (1)
   Drugs & Medical substan
   Health Information (42)
   Healthy Lifestyle (1)
     Exercise & Weight loss
   HIV/AIDS (2)
   Home (20)
   Images (100)
   Influenza (3)
   Men's Health (1)
   Mental health (3)
   Research (150)
   Scientific article (50)
   Smoking (2)
   Weblinks (42)
   Women's Health (8)
   Work (9)
   Young people (5)
By Decade
By Source
By Publisher
By Target audience
By Language
By Media type
By Image modality

☐ All Items (250)

1. **Diabetes in Children and Adolescents - Medpedia**
Diabetes in Children and Adolescents - Medpedia   <\![endif]
wiki.medpedia.com
Babies & Children

2. **Utilization of diabetes medication and cost of testing supplies**
2006 — ...the cost of testing supplies in Canada using information collected by
(DP) in 2001. The diabetes cohort (n = 41,630) included individuals who met the
Surveillance System (NDSS) case definition. An algorithm...
J. A. Johnson, S. L. Pohar, K. Secnik, N. Yurgin, Z. Hirji
Research • Scientific article 👁

3. **Diabetes - Medpedia**
Diabetes - Medpedia   <
wiki.medpedia.com
Babies & Children

4. **Reducing health risk in family members of patients with type 2**
2009 — ...Patients with type 2 diabetes can have an important role in discussing
This study aimed to establish the acceptability to first degree relatives towards t
diabetes intervening as health promoters a...
D. L. Whitford, H. McGee, B. O'Sullivan
Research • Scientific article

5. **Clinical:Insulin Resistance and Prediabetes - Medpedia**
...have "prediabetes." Doctors sometimes call this condition impaired fasting glu
wiki.medpedia.com
Cardiovascular dis

6. **Trends in laboratory testing for diabetes in Ontario, Canada 1**
2009 — ...There are concerns that testing for type 2 diabetes is low and many p
diagnosed. We sought to describe the rates of diabetes-related lab testing in On
adults without diabetes, and to explore the to...
S. E. Wilson, L. L. Lipscombe, L. C. Rosella, D. G. Manuel
Research • Scientific article

Utilization of diabetes medication and cost of testing supplies in Saskatchewan, 2001.

2006 — J. A. Johnson, S. L. Pohar, K. Secnik, N. Yurgin, Z. Hirji — BMC Health Serv Res

http://www.biomedcentral.com/1472-6963/6/159

Keywords: Research, Scientific article

Abstract

The purpose of this study was to describe the patterns of antidiabetic medication use and the cost of testing supplies in Canada using information collected by Saskatchewan's Drug Plan (DP) in 2001. The diabetes cohort (n = 41,630) included individuals who met the National Diabetes Surveillance System (NDSS) case definition. An algorithm was then used to identify subjects as having type 1 or type 2 diabetes. Among those identified as having type 2 diabetes (n = 37,625), 38% did not have records ...

Images
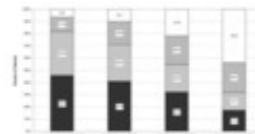
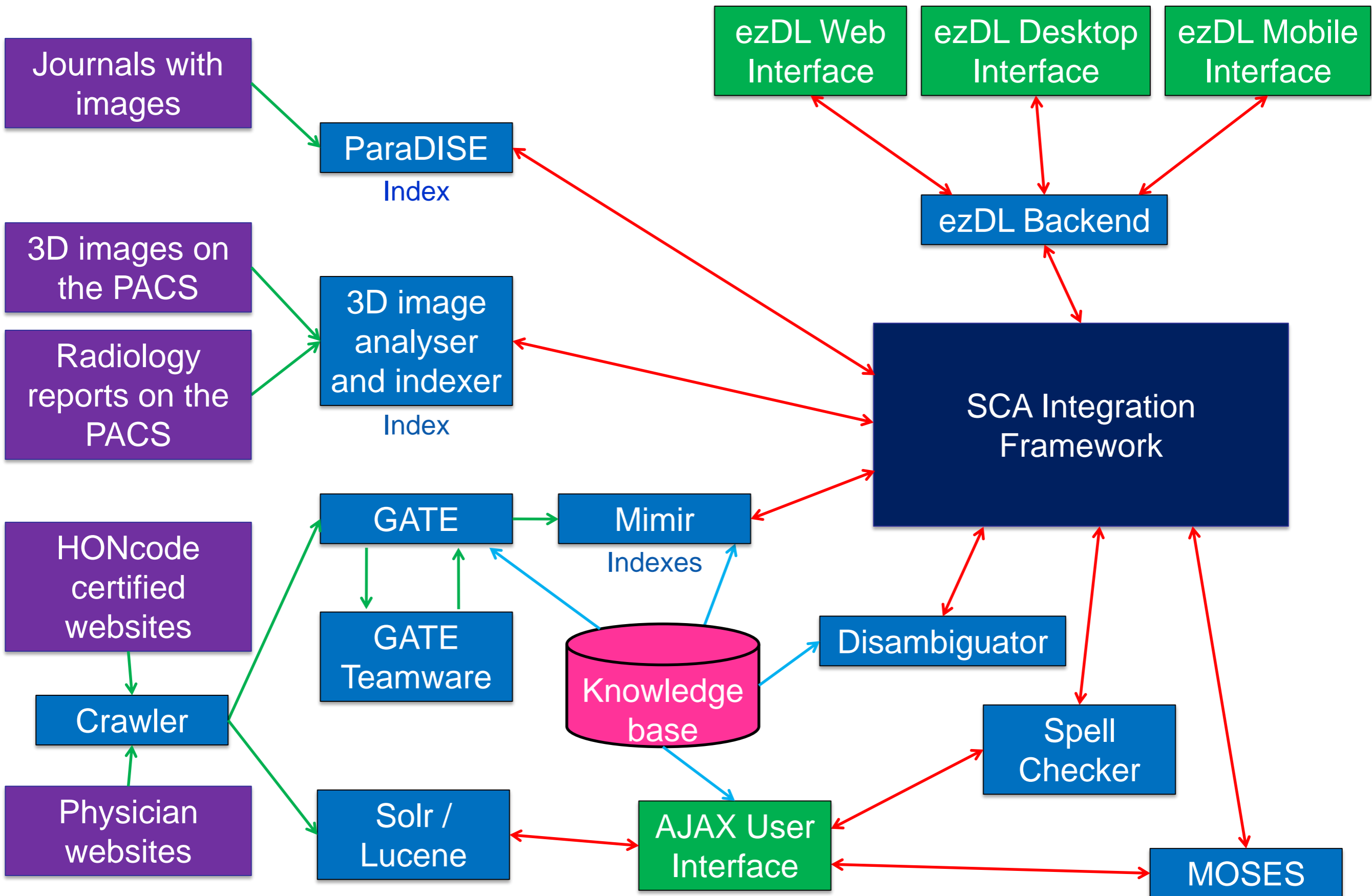Similar Images
Add to query

Frequent terms in this document

algorithm antidiabetic canada case cohort collected cost definition
describe **diabetes** dp drug identified identify included individuals
information medication met national ndss patterns plan purpose
records saskatchewan study subjects supplies surveillance system
testing type

Detail links

http://www.biomedcentral.com/1472-6963/6/159

Comments (0)

Search finished

(Unknown city), AUSTRIA (AT); .00/.00
English (German)

The Khresmoi System

Study Design and Setup

- **4 tasks**
  - Simulated work tasks with right/wrong answers
  - 10 mins for each task,1 hour overall
  - Suggested tools (personal library, sharing, translation)
- **Questionnaires**
  - Demographics
  - Per Task
    - prior knowledge before searching
    - Multiple choice answer + supporting links as documentation
    - Feedback: good/bad/suggestion
  - Overall feedback
    - SUS + custom questions

Example Task (shortened)

Atrial Fibrillation

Is it ok for a 69 year old women with a history of atrial fibrillation
and cardioversion to stop anticoagulation due to recent rhythm stability?

Case scenario:
A 69 year old woman, diagnosed 4 years ago with atrial fibrillation has successfully received cardioversion. That time she felt elevated heart rate and palpitations and is taking oral anticoagulants. Since then she is symptom free. She is health conscious and regularly measures her heart rate, which seems ok. She is otherwise healthy, her heart has a normal structure, only the left ventricle shows a moderate enlargement. She wants to stop oral anticoagulants.

From your knowledge: Is it ok for her to STOP taking oral anticoagulants?
( ) Yes
( ) No
( ) I don't know / I require further information to answer this question

**Ask for prior knowledge before searching**

Please use KHRESMOI to find the (evidence to support your) answer and cite at least 3 websites (or more until you are confident in your answer) that you consider supportive.

People and Data

- In person testing
- Two researchers per participant
  - facilitator and observer (in another room)
- All participants are medical doctors (19 in total)
  - Not students or other surrogates
  - Test users are true target users of the system
- Data collection
  - Questionnaires
  - Observation
  - Logging
  - Self-report
  - Discussions

Test Software

- ## Run the test
  - Display task descriptions and questionnaires in order
  - Prompt users
- ## Record the test
  - Screen
  - Mouse, Keyboard, Window Events
  - Webcam and Audio
- ## Observe the test
  - Comments and observations
  - Add live and later
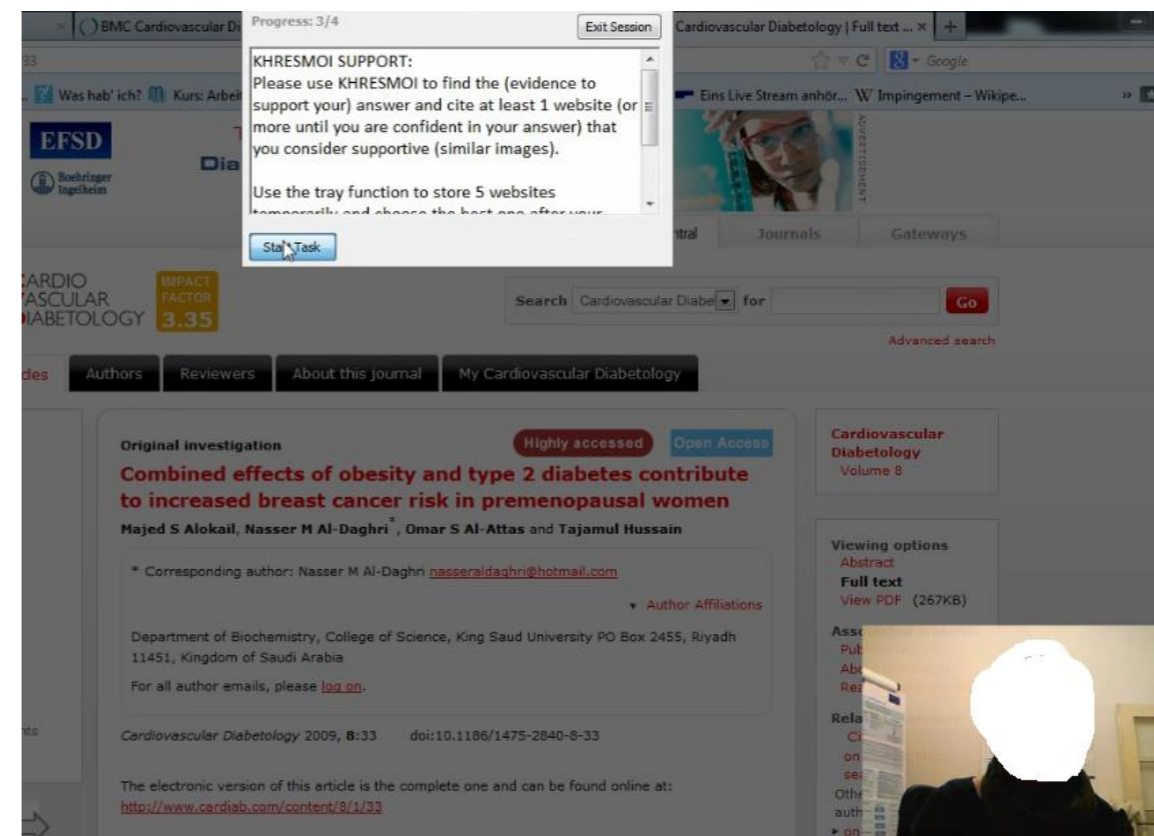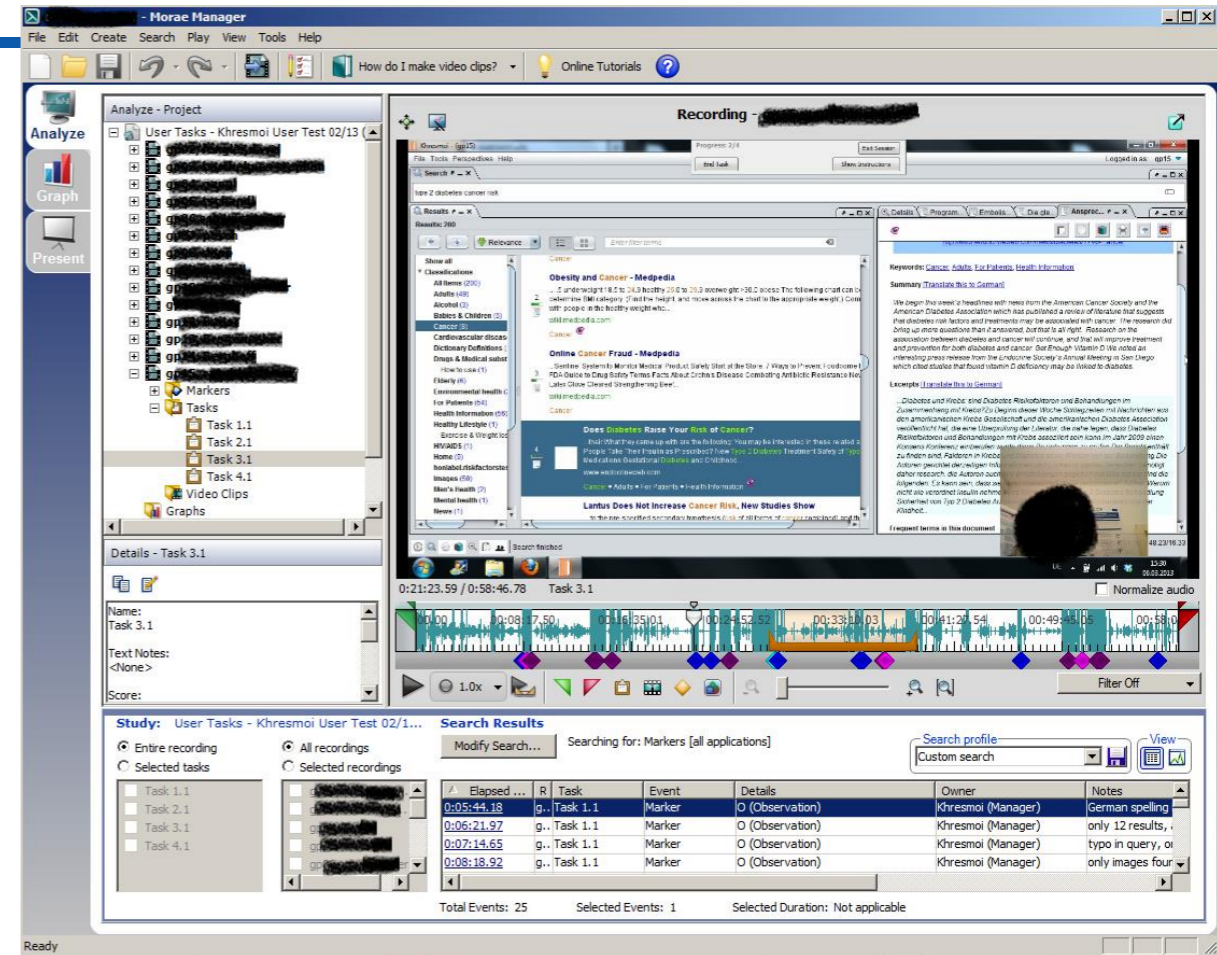
Iterative process

1. Prepare initial versions of tasks, questions, test protocol

2. Run pilot tests with test users, applying changes after each one until no more problems occur

3. Freeze test and run with exactly the same setup x times

4. Evaluate the data from 3

5. Go back to 1 with learnings from 3 and 4

NB: all users participate only once

Pilot tests: Test the test before you test

- Initial system design was based on a large-scale online survey

- Leads to unrealistic expectations with regard to what tests can accomplish

- Pilot tests: Rapid iteration of test design (tasks, questionnaires, schedule)

- In our case: Stable after 5 versions

Outcome:

- **Shorter, simpler, clearer, more focused, more reliable**

14

Findings (1)

- **Multilingual environment**
  - Mixed-language queries are surprisingly common
    - Participants were German speakers, test setup was in English
    - Language detection and spelling correction stumble over mixed language queries
    - Hard to decide whether misspelled English or correct German word
  - Translation of the whole process
    - Either users really depend on the translation, then everything needs to be translated, incl. fulltext of results
    - Or they switch to English anyway

Findings (2)

- Very efficient in uncovering:
  - Bugs and errors in software
  - Additional data source and display requirements
  - Implicit assumptions of developers vs. users
    - E.g., that queries are always in one language
- Finds problems caused by integration of components that are much harder to find with a component-based evaluation
  - Spelling correction vs. translation
  - Filtering results vs. clustering feature

Future work

- **More tests**
  - Smaller "walk in" tests
    - at professional meetings
    - at conferences during coffee breaks
  - Much shorter (20 mins total maximum)

- **More different users (age, specialty, etc.)**
  - Planned for rest of 2013:
    - Pediatricians and dermatologists
    - Specialists of internal medicine
    - Self-employed general practitioners

- **More features to test**
  - Query templates
  - Semantic queries
  - Image search

Summary

- Medical search is a domain with room for improvement

- We employed an iterative approach to perform a formative evaluation

  - With target users (medical doctors) as testers

  - With realistic simulated work tasks

  - With integrated, comprehensive system under test

- Orthogonal to component-based evaluation

- Resources will be made available

- http://khresmoi.eu
- @khresmoi
- Khresmoi LinkedIn Group