Overview

• CHORUS Use Case (UC) Typology
  • What? Why? How?

• Leveraging the UC Typology
  • Identification of test subjects
  • Formulation of simulated work tasks
  • Benchmarking & evaluation
    • Establishing benchmarks
    • Relevancy
    • System architecture

• Conclusions

CHORUS Use Case Typology

~ What? ~

• Identified use case information most relevant to MMSE
• Formalized into a typology
• Administered as a survey to produce standard (use case) profiles for projects
CHORUS Use Case Typology

~ Why? ~

- Saves labor
  - Projects don't need to conduct such extensive unique user studies
- Benchmarking & evaluation
  - Enables cross-project evaluations
    - Profiles can be meaningfully compared across projects
  - Helps to identify most relevant performance criteria for a system
Leveraging the Typology

~ Test Subjects ~

• Identify test subjects (potential users)
  • Test with actual potential users that developers are targeting (Borlund 2003)
Leveraging the Typology
~ Test Subjects ~

• Revealed by UC profile sections:
  • **Topical Domain**
    • i.e., management, medicine, art
  • **Content Type**
    • i.e., text, images, music
  • **System & Domain Competence**
    • i.e., novice, professional
  • **User Roles**
    • i.e., consumers, owners, producers
  • **Community Size**
    • i.e., small, medium, large
Leveraging the Typology
~ Simulated Work Tasks ~

• Definition
  • A short description of a context or scenario that would prompt an individual to use the MMSE system

• 2 important functions
  • Allows user to interpret the information need
  • Framework against which relevance is judged

• Example
  • You are a gardener interested in organic techniques for enriching your soil. You've heard that you can re-use kitchen scraps and yard refuse, such as lawn clippings, to amend your soil and reduce your need for chemical fertilizers. You don't have the time to read a book about organic gardening, so you would like to find a short video to quickly get you started on home composting.
Leveraging the Typology
~ Simulated Work Tasks ~

• Revealed by UC profile sections:
  • Goal of Interaction
    • i.e., retrieve content, stream content, monitoring
  • Query Type
    • i.e., explicit, implicit
  • Retrieval Strategy
    • i.e., browse, recommendation
  • Service Platform
    • i.e., desktop, enterprise, Internet
  • Device
    • i.e., personal computer, e-book, mobile device
Leveraging the Typology

~ Simulated Work Tasks ~

• Follow-Up Interviews
  • User characteristics
    • Verification
    • Discover overlooked traits or user groups
  • More realistic simulated work tasks
    • UC survey collects general information
    • Interviews reveal more detail about real information needs
Benchmarking & Evaluation
~ Establishing Benchmarks ~

• Project classification (fundamental market categories)

  (1) Web Search (WS)
  (2) Enterprise Search (ES)
  (3) Library Search (LS)
  (4) Personal Search (PS)
  (5) Personalized TV (PTV)
  (6) Monitoring, Detection & Alert (MDA)
Benchmarking & Evaluation
~ Establishing Benchmarks ~

Generic UC attributes:
Benchmarking & Evaluation

~ Establishing Benchmarks ~

- Each attribute/value set enumerated in the typology
- Administered as six survey questions
- Example (a project profile indicates:)
  - Developing indexing technologies for well-organized repositories (i.e., controlled and complete metadata) of multimedia.
    - Most likely generic UC: PTV
  - Verified by follow-up interview
Performance Baselines

~ Establishing Benchmarks ~

• **Performance Criteria (baselines)**
  - For each relevancy measure in each UC category
  - Baseline established by evaluating systems in each UC category

• **Evaluation**
  - Projects would try to meet or exceed the criteria set by previous benchmarks

• **In other words**...
  - Projects would aim for performance criteria considered important by their targeted users
Benchmarking & Evaluation

~ Relevance ~

• Indicates a relationship
  • Algorithmic (system relevance)
    • RELATION: query and retrieved object
    • CRITERION: comparative effectiveness
  • Topical (subject relevance)
    • RELATION: topic expressed in a query and the topic covered by retrieved objects
    • CRITERION: “aboutness”
  • Cognitive (pertinence)
    • RELATION: the state of knowledge and cognitive information need of a user, and texts retrieved
Benchmarking & Evaluation

~ Relevance ~

• Situational (utility)
  • RELATION: The situation, task, or problem at hand, and the retrieved information objects.
  • CRITERION: usefulness in decision making, appropriateness of information in resolution of a problem, reduction of uncertainty, etc.

• Motivational (affective relevance)
  • RELATION: The intents, goals, and motivations of a user, and retrieved information objects
  • CRITERION: satisfaction, success, accomplishment, etc.
Benchmarking & Evaluation

~ Relevance ~

• Historically
  • Simplistic and intuitive
  • Tracked only one relationship (algorithmic)

• Each relevancy measure has a baseline for each UC category
  • Users value each relevancy differently depending on who they are and why they're using the system
  • These baselines are then used as the performance criteria for evaluating projects.
Benchmarking & Evaluation
~ Relevance Measures ~

• Relative Relevance (RR)
  • Quantifies relation between objective relevancies and subjective assessments
    • Introduces subjective performance assessments into traditional (recall/precision) measures
    • Important for evaluating dynamic, contextual systems as well as capturing the new relevancies

• Consistency by abstraction
  • Different kinds of subjective and objective relevance assessments can be associated across many users and
Benchmarking & Evaluation

~ Experiments ~

• Environmental control
  • Even with consistent metrics across project evaluations, results are likely to vary due to confounding factors
  • IR evaluations are essentially experiments
  • Achieved the same as classic experimental control

• Minimal system architecture
  • For search components
  • Specifies an analytics pipeline for content processing
  • Controls extra variables
References


Thank you!

- Paul King & Yiannis Kompatsiaris
- CERTH-ITI / Multimedia Knowledge Laboratory
- http://mklab.iti.gr