

Multi-modal Information Retrieval

experiences from

Context-Aware Image Management, CAIM

Joan Nordbotten

Dept. Of Information and Media Science
University of Bergen, Norway

Multi-modal Information Retrieval - MMIR

1. Modalities of information & data
2. Mobile Multi-modal information retrieval
3. Image context
4. The Information Retrieval (IR) process

The CAIM project framework

1. The technology
2. MMIR from a Nokia 5800 phone
3. MMIR from a PC using [VISI3](#) demonstration

Status of Multi-modal IR in CAIM

1. Evaluations
2. Future research & Development

Multi-modal information - MMI

Definitions

Multi-modal: having or involving several modes, modalities

Mode: 1) a particular form or variety of something
2) a manner of doing something

(Webster, 1974,2003)

Information: "The meaning that a *human* extracts from data
by means of known conventions of the *representation* used."

(IFIP. Gould, 1971)

Multi-modal Data: The *representation* of information
using multiple modes (formats)

(Jcn, 2009)

Data Modalities

- **Numeric** **count, size, age, ...**
- **Text** **document, title, keyword, ...**
- **Image** **painting, drawing, photo, sketch, ...**
- **Audio** **speech, music, signal, ...**
- **Spatial** **point (gps), line, area, direction, ...**
- **Temporal** **date, time, duration (season, period, ...), ...**
- **Sensory?** **Touch, taste, smell**

Multi-modalities:

- **Static** **illustrated text: newspaper, web-page, map, chart, table, ...**
- **Dynamic** **video, film, TV**

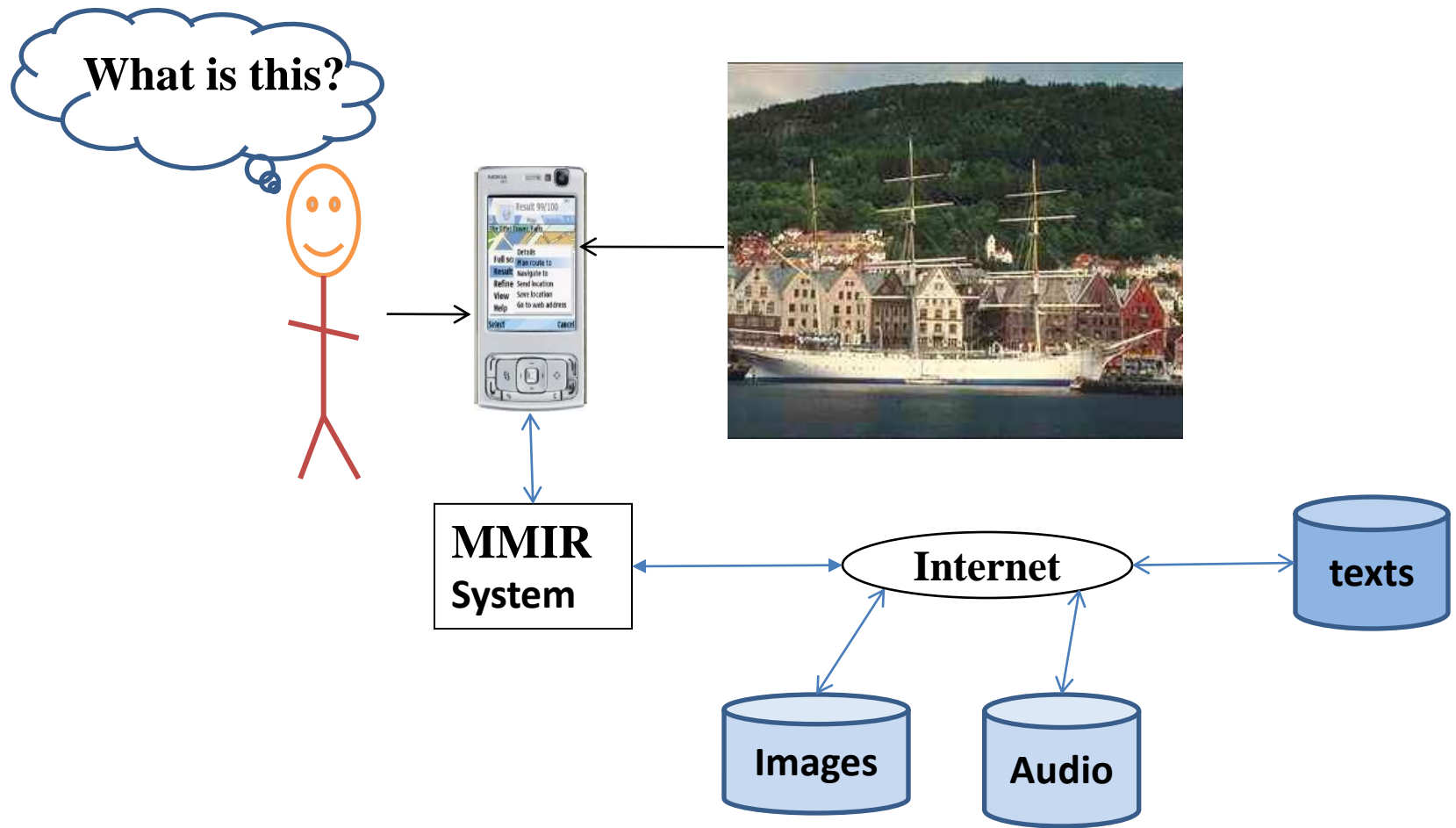


tourist

**Guide
Book**



Mobile MMI Retrieval



- **“*Context* any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, ...”.**
(Dey & Abowd, 2000).
- **“*Context* the information that must be common knowledge between the user and system for communication between them to succeed.”**
(Elgesem & Nordbotten, 2007).

Image context examples

Creation - automatic

- time, date
- gps
- direction
- Camera settings
- ...



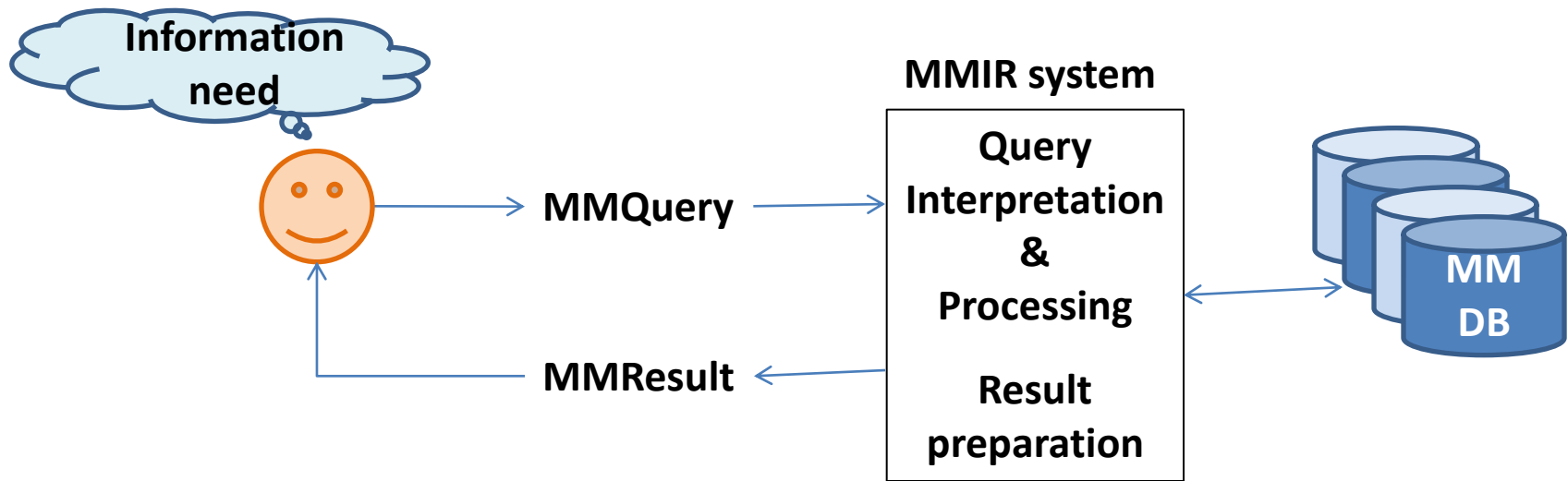
Creation - manual

- Photographer
- Owner
- purpose
- Object(s)
- Object location
- Object age
- Collection(s)
- ...

Contents

- Object ids
- Event/activity
- Descriptors
- Descriptor weights
- Inter-object relationships
- ...

CAIM MMIR – query processing



MM-Query: <image, gps-camera, time, date>

MM-DB: <image, gps-obj(s), text →, audio →, <context data>>

MM-result: <image-set, text, audio, ... >

CAIM prototypes

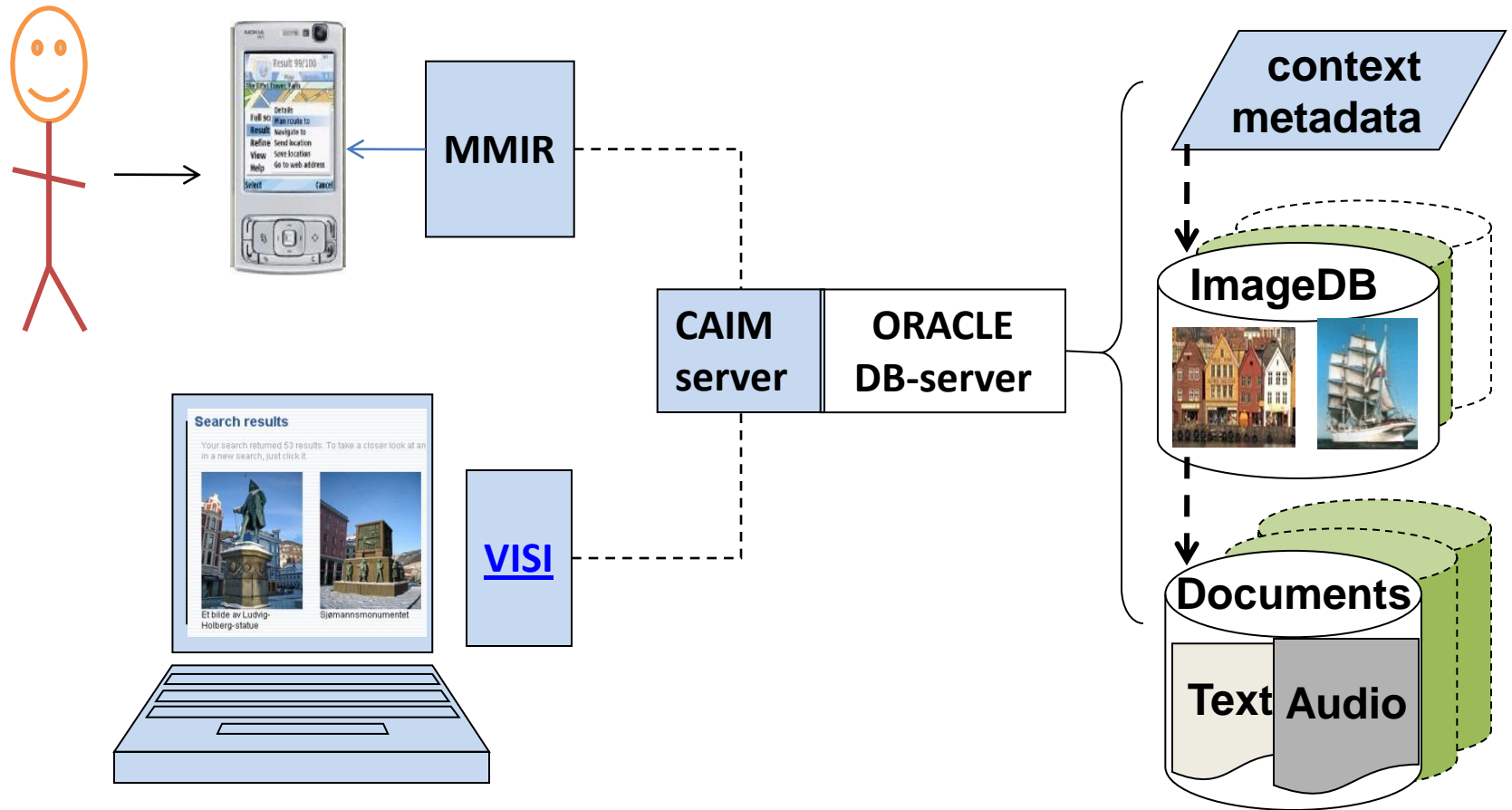


Image Retrieval w/Oracle

Quality \approx Precision

	<i>P%</i>	<i>comments</i>
• TBIR:	70+	image descriptors
• CBIR:	<20	based on color match ☹
• CBIR+TBIR	70	assuming input tags
• CBIR+GPS	50+	lack of direction

Notes:

1. **Small studies in # participants, DB size and query set**
2. **Relevance is very subjective**
3. **Recall is nearly impossible to measure, p.g.a. #2 above and DB size**
4. **2 precision measures used: every 'n'th retrieved and for each relevant image**

Future research & Development

- **Query specification:**
 - **object identification**
 - **sketch**
- **Object recognition:**
 - **Separate multi objects within an image**
 - **Expand use of image context**
 - **Record object location in image**
- **Improve MM merge algorithms**
- ...



The CAIM project, <http://caim.uib.no>

Møller, T. (Sept. 2009). [*Bergen By – a Multi-Modal Image Database*](#). CAIM-TR-9, Dept. of Information and Media Sciences, Univ. of Bergen.

Carlson, Christoph. (Sept. 2009). [*VISI3 – Context Aware Image Retrieval*](#). CAIM-TR 8. Dept. of Information and Media Sciences, Univ. of Bergen.

Hellevang, Mathias (Sept. 2009). [*MMIR3 – Mobile Multimedia Image Retrieval*](#). CAIM-TR 7. Dept. of Information and Media Sciences, Univ. of Bergen

Questions?

Thank you for your attention 😊

MMIR processing

