INEX 2005 Multimedia Track - Working Document

Roelof van Zwol   Mounia Lalmas   Gabriella Kazai

October 4, 2005

1   Introduction

This guide serves as the working document for the INEX 2005 Multimedia Track and will accumulate all necessary guides.

1.1   Objective

The main objective of the INEX 2005 multimedia track is to provide an evaluation platform for structured document retrieval systems that do not only include text in the retrieval process. Many structured document collections today also contain other types of media, such as images, speech, and video. To include these media types into the retrieval process and to produce a meaningful ranking is far from trivial.

Using the structure of the document as a semantic/logical backbone for the retrieval of multimedia document fragments will allow us to investigate this problem from a new perspective. In the first year of the multimedia track, we plan to provide an evaluation platform for the retrieval of multimedia structured document fragments, rather similar to the methodology used for the INEX Ad Hoc track.

1.2   Task description

The task for the multimedia track is to retrieve relevant document fragments based on an information need with a structured multimedia character. A structured document retrieval approach in that case should be able to combine the relevances of the different media types into a single (meaningful) ranking that is presented to the user. The INEX multimedia track differs from other approaches in multimedia information retrieval, in the sense that it focuses on using the structure of the document to extract, relate and combine the relevances of different multimedia fragments.

The focus for 2005 will be on the combination of text and image retrieval. For this purpose we will use the Lonely Planet document collection, which can be explored at: `http://contentlab.cs.uu.nl/lonelyplanet/Collections/LonelyPlanet/`. The required login and password for this web-site is sent to the participating groups, once the data-handling agreement for the Lonely Planet collection is received by the organization.

2   Topic development guide

This topic development guide is heavily based on the developement guide used for the INEX 2005 Adhoc track[1] Here we will redefine the Topic format and Topic formulation procedure. For clarity we

---

[1]http://inex.is.informatik.uni-duisburg.de/2005/internal/pdf/TD05.pdf

have inserted a part of the section about topic creation criteria.

2.1 Topic creation criteria

Creating a set of topics for a test collection requires a balance between competing interests. The performance of retrieval systems varies largely for different topcis. This variation is ussually greater than the performance variation of different retrieval methods on the same topic. Thus, to judge whether one retrieval strategy is (in general) more effective than another, the retrieval performance must be averaged over a large and diverse set of topics. In addition, to be a useful diagnostic tool, the average performance of the retrieval systems on the topics can be neither too good nor too bad as little can be learned about retrieval strategies if systems retrieve no, or only relevant, documents.When creating topics, a number of factors should be taken into consideration. Topics should:

- be authored by an expert in (or someone familiar with) the subject areas covered by the collection,

- reflect real needs of operational systems,

- represent the type of service an operational system might provide,

- be diverse,

- differ in their coverage, e.g. broad or narrow topic queries,

- be assessed by the topic author.

In addition we add for the multimedia track that the topics should:

- have a clear multimedia character, which should be clearly specified in the topic description and narrative.

Within the multimedia track we'll focus on the content and structure topics, as these allow explicit formulation of the multimedia character in the information request, e.g. NEXI-CAS query.

Consider for example the topic based on the Lonely planet collection below:

Example 1

```
<inex_topic inex_track="MM" query_type="CAS" topic_id="0">
<description>Find images depicting scuba diving activities for destinations with a tropical climate and that discuss exploring the beautiful underwater nature by snorkeling and diving activities</description>
<castitle>//destination[about(.//weather,tropical climate) and about(.//activities, beautiful "underwater nature" snorkeling and diving)]//images//image[about(., scuba diving)]</castitle>
<narrative>We want to go on a scuba diving trip to a destination with a tropical climate and beautiful underwater scenery (nature). Therefore we want to find some images on this subject, so that we can have some pre-holiday fun.</narrative>
</inex_topic>
```

The `description` and `narrative` specify the information need and its explanation. Based on this information need, we can formulate a NEXI-query as given in the `title` section of the topic definition. Our example query above includes both textual and image components of the information need, e.g. `about(.//weather,tropical climate)` is a condition that the weather textual element of a destination should talk about tropical climate, while the `//image[about(., scuba diving)` asks for images that depict scuba diving scenes.

Although the target elements of the above example are images, so far, simple textual retrieval approaches may be sufficient to produce the required output by searching image captions. However, a combination of text and image retrieval systems is encouraged within the track as these may in fact produce better results.

Consider next the following example topic:

Example 2
```
<inex_topic inex_track="MM" query_type="CAS" topic_id="0">
```
<description>Find images depicting scuba diving activities, like in BN5970_6.jpg, for destinations with a tropical climate and that discuss exploring the beautiful underwater nature by snorkeling and diving activities</description>
<castitle>//destination[about(.//weather,tropical climate) and about(.//activities, beautiful "underwater nature" snorkeling and diving)]//images//image[about(., scuba diving) and about(., src: /image/BN5970_6.jpg)]</castitle>
<narrative>We want to go on a scuba diving trip to a destination with a tropical climate and beautiful underwater scenery (nature). Therefore we want to find some images on this subject, so that we can have some pre-holiday fun.</narrative>
```
</inex_topic>
```

In this example, an instance of the *about* clause, `about(., src:/image/BN5970_6.jpg)`, uses a slightly different syntax to the usual two arguments of a path directive and a textual description of the information need. Here the second argument contains a reference to a multimedia object, in this case, the name and path of an image file. This extension of the *about* clause allows for querying by sample.

By expressing both the content and image components of the information need within the same *about* clause, we are effectively overloading its meaning, leaving it to the retrieval system to decide if a text or image search (or both) is required. The reason for doing so is to emphasise the multimedia nature of the track. Using the extended *about* clause, we can specify query constraint for a document fragment (which may be pure text, image, or a combination of multiple media) using a textual descriptions (e.g. `about(//image, scuba diving)`, `about(//destination, scuba diving)`) or using similar images (e.g. `about(//image, src:/image/BN5970_6.jpg)`, `about(//destination, src:/image/BN5970_6.jpg)`).

Various combinations of query conditions will require different strategies where text and image retrieval can be combined. The track's focus is on the combination of the two techniques and therefore we encourage the submission of topics that force systems to implement content-based image retrieval (e.g. `about(//destination, src:/image/BN5970_6.jpg)`).

2.2 Topic format

The topic format for the multimedia track consists of the following fields: a **description**, **castitle**, and **narrative**. The following information should be contained in each of these fields:

- `<description>` A brief description of the information need, specifying any structural, textual, and visual requirements/composition on the content. The description must be precise, concise, and informative, but it must contain the same terms and the same structural requirements that appear in the `castitle`, albeit expressed in natural language.

- `<castitle>` A valid NEXI expression based on the Lonely Planet document collection that contains at least one *about* clause containing at least one image component. The expression is of the form *//A[B]* or *//A[B]//C[D]*.

- `<narrative>` A detailed explanation of the information need and the description of what makes and element relevant or not. The narrative should explain not only what information is being sought, but also the context and motivation of the information need, i.e. *why* the information is being sought and what purpose it may serve. Assessments will be made on compliance to the narrative alone; it is therefore important that this description is clear and precise.

The additional constraints, as defined in the Topic development Guide for the INEX Ad Hoc track apply, for so far they are applicably and not overruled in this additional guide.

2.3 Topic development guidelines

Each participating group will have to submit 3 (CAS) topics by July 15, 2005 using the on-line form provided.

The topic creation process is divided in several steps.

Step 1: Initial topic statement

Create a one or two sentence description of the information you are seeking. This should be a simple description of the information need without regard to retrieval system capabilities or document collection perculiarities. This should be recorded in the Initial Topic Statement field. Record also the context and motivation of the information need, i.e. why the information is being sought.

Step 2: Exploration phase

In this step the initial topic statement is used to explore the collection. Obtain an estimate of the number of relevant elements, then evaluate whether this topic can be judged consistently. We have two search engines available, a text-based system and an image-based system, which can be used by the participants for topic development. These are reachable from: `http://contentlab.cs.uu.nl/~lonelyplanet/`.

Step 2a: Assess the top 25 text fragments

You have to judge the relevancy of the retrieved text fragments (using binary relevance only). Each result should be judged on its own merits. Abandon a search, if there are fewer than 2 or more than 20 relevant text fragments in the result list.

Step 2b: Assess top 25 images

Since most participants will not have an off-the-shelves system available for the multimedia track, we have chosen to do a separate scan for the relevance of the image component. Therefore you'll have to assess the top 25 images and judge their relevance (using binary relevance only). Each result should

be judged on its own merits. Abandon the search, if there are fewer than 2 or more than 20 relevant images in the result list.

Step 2c: Inspect document matching

To assure that the document collection has a reasonable chance of completely fulfilling the text and image-based constraints of the information need a check at document level is needed. In this step, you have to count the number of documents that both satisfy the textual conditions and the image conditions. I.e. if a relevant text fragment was found, and there is an image that belongs to the same XML document, a match is found that corresponds to the information need. Abandon the topic if less than 3 matches are found over the top 25 results for both components.

Step 3: Write description, title, and narrative

During this step you're asked to complete the topic definition by writing the description, title, and narrative. Please obey the instruction that are given for the topic development procedure, as described in the INEX 2005 Guidelines for Topic Development.

Step 4: Topic submission

Once you are finished, fill out the on-line Candidate Topic Submission Form on the INEX website at http://inex.is.informatik.uni-duisburg.de/2005 under Tasks/Tracks → Multimedia → Topics → Submit Topic. Make sure you submit all your candidate topics no later than 15 July, 2005. Thank you.

3   Relevance assessments guidelines

We base the definition of relevance on the definition employed in the INEX ad-hoc track with the exception that we measure exhaustivity only on a binary scale. In addition, reflecting the SSCAS task, we only consider an XML element relevant if it strictly matches the structural conditions specified within the query, i.e. only target elements may be relevant and only if they are contained in an XML document that satisfies the query's containment constraints.

Therefore, a given multimedia fragment is said to be relevant if it "discusses" (or depicts) the topic of request to any degree and if it strictly adheres to the structural conditions requested by the user.

Similarly to the ad-hoc track, the assessment procedure follows the highlighting approach. However, given the binary nature of relevance, the assessment procedure for the multimedia track consists only of a single pass. During this single pass assessors are to highlight multimedia fragments that contain only relevant content, i.e. relevant content that contains no (or only minimal) non-relevant content. In the case of textual content, only relevant text fragments, e.g. words or sentences, should be highlighted. In the case of images, since currently it is not possible to highlight only a part of an image, the whole image should be highlighted if it contains relevant content (regardless of how much of the image may be non-relevant).

The on-line assessment interface can be accessed at https://inex.lip6.fr/2005/xrai-mm/ using your multimedia track login. The assessment system is described in the INEX 2005 Relevance Assessment Guide[2]. The only difference is that the Toolbar has been simplified. Firgure 1 shows a document of the Lonely Planet collection being assessed. The toolbar can be seen in the bottom left corner. The icons from left-to-right are: remove assessments, save assessments, save assessments and move to the previous document, save assessments and move to the next document, show or hide support elements(i.e. retrieval results), set document as assessed, highlight selected text fragment, unhighlight

---

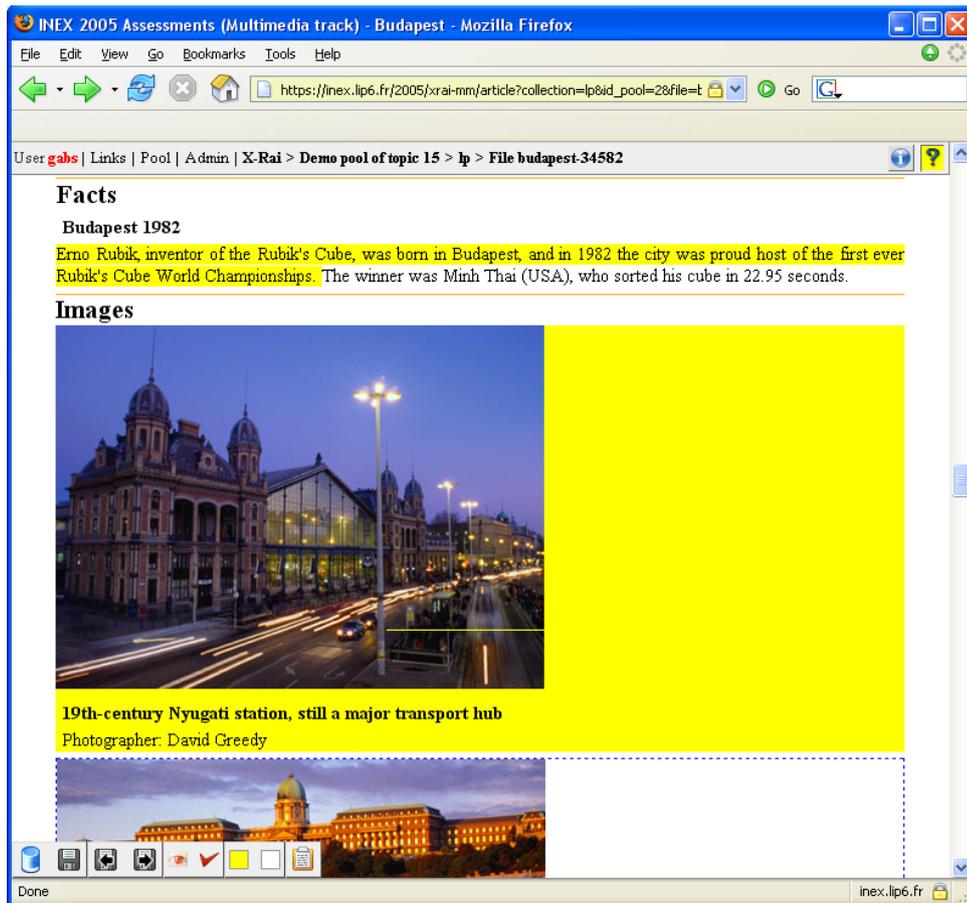[2]http://inex.is.informatik.uni-duisburg.de/2005/internal/pdf/Relevance_Assessment2005.pdf

Figure 1: The on-line assessment system and toolbar

selected text fragment and passage information. Please make sure you set a document as assessed and save it before moving onto the next document to assess.

All highlighted multimedia fragments will be saved as exhaustive, fully specific and strictly satisfying the structural constraints.

Given the small collection and the strict structural match we expect that assessment work will be light. However, the strict relevance criteria may bring up unforeseen questions that we would like to hear about. We would like to ask assessors to keep a log book and record any questions or uncertainties that arised during assessment or any situations when you felt you had to make perhaps not-so-obvious decisions. For each issue or question, please record the topic id and the particular XML element that sparked the question. To help you with this reporting task, the assessment system provides an easily accessible passage information facility (the right-most button of the Toolbar). Simply highlight a passage and click the clipboard button to display all the reference information that you need to include in a log entry. Please then submit your log to us or even make it available to everyone by emailing the list of multimedia participants. Thank you.

4 Acknowledgements

This guide is heavily based on the Topic Development Guide for the INEX 2005 Ad hoc track. We are thankful for the guidelines provided there and thankfully adopt them, for as far as applicable, for the INEX 2005 Multimedia track.

We would like to express our gratitude for the work of Benjamin Piwowarski for providing the necessary modifications to the on-line assessment system.

5 Example

Based on the example introduced earlier, the following should be submitted:

1) Affiliation: Utrecht University, the Netherlands
2) Author (name, e-mail): Roelof van Zwol, roelof@cs.uu.nl
3) Initial topic statement:
We want to go on a scuba diving trip to a destination with a tropical climate and beautiful underwater scenery (nature). Therefore we want to find some images on this subject, so that we can have some pre-holiday fun.
4) Number of relevant text-fragments in TOP 25: 14
5) Top 25 retrieved text fragments (path, document):
```
file:  comoros-and-mayotte-755.xml path:/destination[1]/images[1]
file:  bonaire-3807.xml path:  /destination[1]/images[1]
...
file:  vancouver-18170.xml path:/destination[1]/activities[1]
```
6) Number of relevant image fragments in TOP 25: 7
7) Top 25 retrieved images (image name or path, document):
```
file:  australia-245.xml image:  /images/BN918_1.jpg
file:  belize-590.xml image:  /images/BN9002_8.jpg
...
file:  comoros-and-mayotte-755.xml image:  /images/BN5970_6.jpg
```
8) Matching documents: 4
9) Candidate topic:
```
<inex_topic inex_track=''MM'' query_type="CAS">
<description>Find images depicting scuba diving activities for destinations with
a tropical climate and that discuss exploring the beautiful underwater nature by
snorkeling and diving activities.</description>
```
<castitle>//destination[about(.//weather,tropical climate) and about(.//activities, beautiful "underwater nature" snorkeling and diving)]//images//image[about(., scuba diving) and about(., src:/image/BN5970_6.jpg)]</castitle>
```
<narrative>We want to go on a scuba diving trip to a destination with a tropical
climate and beautiful underwater scenery (nature).  Therefore we want to find some
images on this subject, so that we can have some pre-holiday fun.</narrative>
</inex_topic>
```