FlickLing: a Multilingual Search Interface for Flickr

Víctor Peinado, Javier Artiles, Julio Gonzalo, Emma Barker*, and Fernando López-Ostenero
NLP & IR Group, ETSI Informática, UNED
c/ Juan del Rosal, 16, E-28040 Madrid, Spain
victor@lsi.uned.es, javart@gmail.com, julio@lsi.uned.es
e.barker@dcs.shef.ac.uk, flopez@lsi.uned.es

Abstract

In this paper, we present FlickLing, a multilingual search interface for Flickr designed and implemented for the CLEF 2008 interactive task.

FlickLing consists of two interfaces which allow to perform monolingual and multilingual image retrieval over the Flickr database, retrieving results with images annotated in different languages. From a given query entered by the user, FlickLing performs automatic term-by-term translation into up to six languages, and provides assistance for interactive query refinement and translation.

With the goal of collecting a large search log, Flickling works as an online competitive game, where users have to find as many images as possible to obtain the highest individual and team scores. Here we describe the user logs generated, which has been used as a data source for the iCLEF 2008 log analysis task and contain over 5,000 complete search sessions made by over 200 users with a wide variety of language skills.

Categories and Subject Descriptors

H.3 [**Information Storage and Retrieval**]: H.3.1 Content Analysis and Indexing; H.3.3 Information Search and Retrieval; H.4 [**Information Systems Applications**]: H.4.m Miscellaneous

General Terms

interactive information retrieval, cross-language information retrieval

Keywords

iCLEF, Flickr, log analysis, multilingual image search, user studies

1 Introduction

This paper presents FlickLing[1], a multilingual search interface for Flickr designed and implemented by the UNED NLP & IR Group (nlp.uned.es) for the CLEF 2008 interactive task [1].

*Department of Computer Science, University of Sheffield, Regent Court, 211 Portobello Street, Sheffield S1 4DP, UK.

[1]FlickLing is available at http://soporte1.lsi.uned.es/flickling.

Flickr[2] is a popular online community where users share and organise their personal image collections, annotated with titles, descriptions and tagged with terms freely chosen by users (folksonomies). According to Wikipedia, as of November 2007, it already claimed to host more than 2 billion images.

Flickling was designed to collect a large search log of multilingual image searches, which serves as the input data for the log analysis shared task at iCLEF 2008. FlickLing consists of two search modes (mono and multilingual) which allow to retrieve Flickr images annotated annotated in different languages. From a given query, FlickLing is able to automatically translate it into several languages (remembering the user's preferred term translations) and offer the user mechanisms to refine the query and improve the translations provided by the system. In addition, Flickling is offered to users as an online game with ranks for the best individual users and the best teams.

As target, Flickling users were given raw images (without annotations) and the goal was to find in the Flickr database as many images as possible in order to obtain the highest score for them and their teams. To do that, users can launch monolingual and multilingual searches, manipulate the automatic translations or refine their queries. When a user finds the target image, she obtains 25 points. At any time, users can quit and stop searching. When they do that, the system offers some hints to help finding the image. If users accept the hint, their score is penalised. Teams and users are ranked according to their score, precision (percentage of found images with respect to the images seen) and average time spent for each successful search session. The most challenging aspect of the task, besides the difficulty to describe the content of the picture and handling multiple languages, is that users don't know *a priori* which language(s) were used to annotate the image when it was uploaded into Flickr.

In the remainder of the paper we explain the search functionalities of Flickling, the query translation facilities, the questionnaires posed by the system and the search logs syntax and content.

2 Search Functionalities

The goal of Flickling was to collect large usage logs that reflect users' behaviour when facing a multilingual search task. Therefore, some restrictions were observed when designing the interface: (i) all standard multilingual search facilities should be present; (ii) non-standard or overly innovative interface aspects should be avoided, to avoid testing a particular approach to multilingual search interfaces rather than a general study of users' behaviour when searching multilingual information; (iii) as we were seeking for spontaneous, non-controlled users (versus controlled populations under laboratory conditions), the interface should be intuitive and usable without pre-search training; (iv) Flickling should work robustly at least with the two main web browsers, IE and Firefox.

In the remainder of this section we focus on the interface functionalities of Flickling that follow the above requisites.

2.1 Search Modes

The search engine has two search modes:

- A standard search mode with no translation facilities, as shown in Figure 1.
- A multi-language interface which automatically translates query terms into the target languages enabled by the user, as shown in Figure 2. It also shows translation suggestions and allows the user to add and remove query terms.

  In this multilingual interface, the user selects the source language she's writing in (his preferred language by default) and the languages in which she wishes to translate her query (all target languages by default). The user can enable or disable individual target languages just by clicking on them (see Figure 3).

[2]Flickr is available at `http://www.flickr.com`.

Figure 1: FlickLing Monolingual Interface

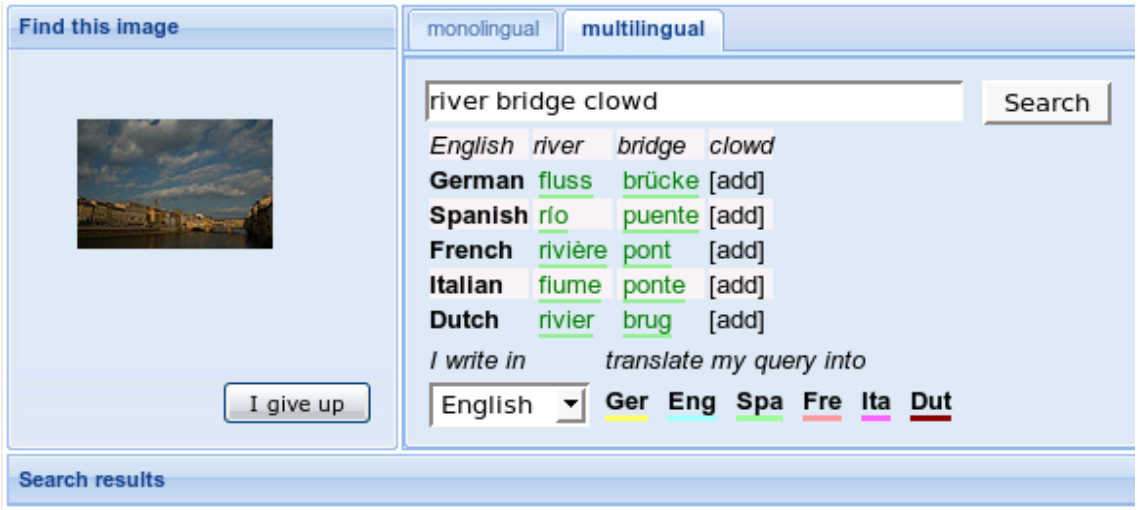

Figure 2: FlickLing Multilingual Interface

Figure 3: Multilingual Interface: selecting source and target languages

The user can personalise her own dictionary by using two different functionalities. First, the user can modify the choice of term translations initially made by the system (see Figure 4). Every click to enable or disable a suggested translation gives, respectively, a positive or negative score to each translation. These scores associated to each possible translation are used to sort the possible translations according to the user's preferences, as explained in Section 3. Second, the user can add new translations to her personal dictionary (see Figure 5).

2.2 Ranking of Results

Both interfaces return a ranking containing the most relevant images for the query, as shown in Figure 6. The associated titles and tags are also shown, since these texts (along with the image descriptions) are the indexed fields against which the textual queries are launched. Immediately above the ranking, the system displays a (multilingual) set of suggested tags related to the query.

When the user clicks on a suggested term, or a term from one of the retrieved queries, the system offers two alternatives to refine the query: (i) adding the term to the current query and (ii) launching a new query with the selected term. When the selected term is in a foreign language, the system tries to translate it into the user's preferred language to facilitate the relevance feedback process (see Figure 7).

On the right of each thumbnail in the ranking there are two different buttons:

- The second icon (the document page) can be used to display the complete image description (if there is such), as shown in Figure 8.
- The first icon (the bell) is used to point at the target image once it is found.

Anytime, the user can decide to quit and stop searching. Next to the target image there's a button to give up. Everytime the user intends to give up, the system will tempt her with a hint. The first hint is always the language in which the image is annotated; after that hint, search becomes a bilingual or monolingual problem for the user. The rest of hints (up to five) are always keywords to find the target image (see Figure 9). This hint mechanism was introduced, after initial testing of the task, to avoid making the task too difficult to engage potential users.

Figure 4: Multilingual Interface: manipulating translation suggestions in the personal dictionary

Figure 5: Multilingual Interface: adding new translations to personal dictionary

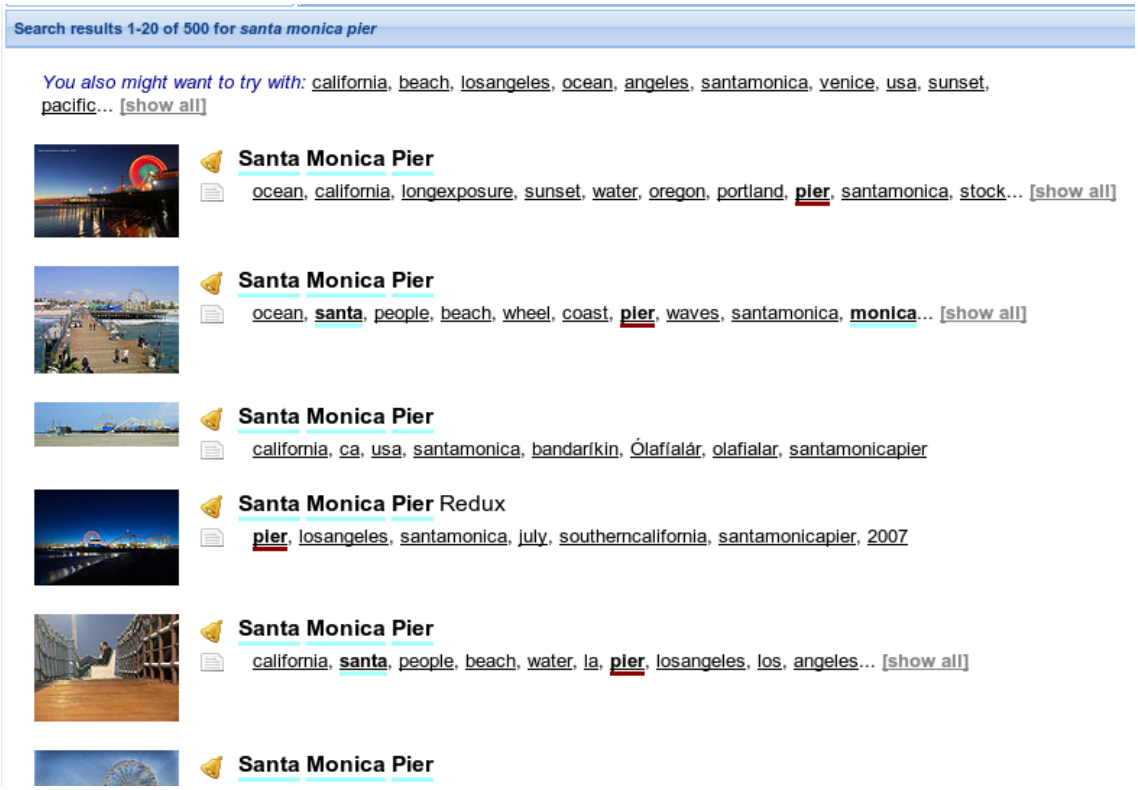

Figure 6: Ranking of results

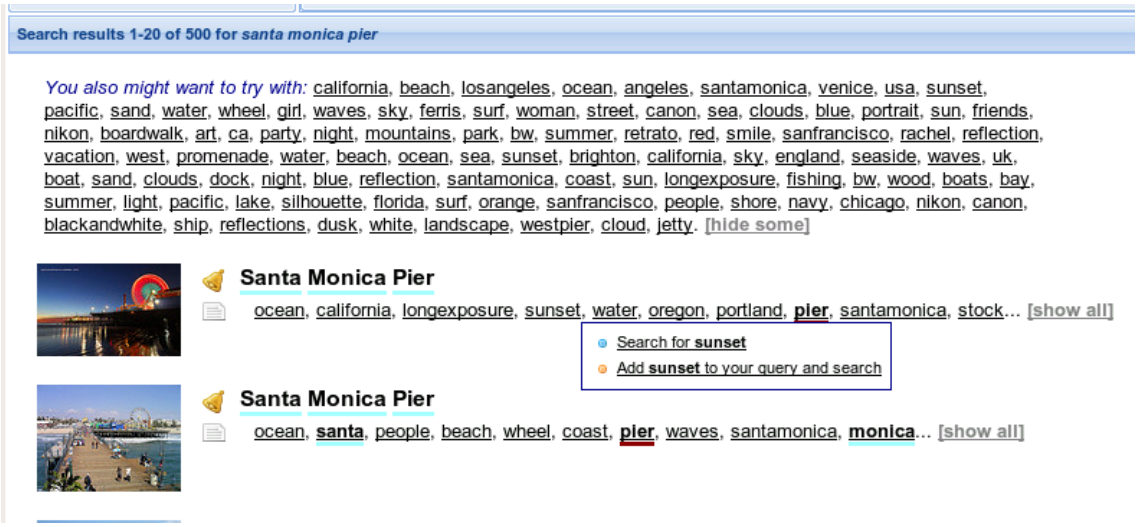

Figure 7: Related terms and image tags can be used to refine or launch new queries

Figure 8: The document icon allows to access the image description

Figure 9: The first hint provided by the system is always the annotation language of the target image

2.3 User profiling and interface localisation

All Flickling users register before starting their searches. This is needed to track the game-like features of the interface (scores, images already seen, etc.) and also to enrich the search logs (user's language abilities, search history, etc.). The user profile focuses on language abilities, and records native language(s), preferred interface language, and skills in each of the task languages (active knowledge, passive knowledge or unknown). The interface has been localised in the six languages handled by the multilingual search facility (EN, ES, IT, FR, NL, DE).

In addition, users may register as belonging to a certain group; this is designed to allow for studies which focus on specific user profiles, and also to engage users in group-oriented competition. The list of groups is predefined, and everyone wanting to create a new group must submit a request to the Flickling system manager.

At any point during a session, a user may see his personal scores (points, average precision, average time spent per image, etc.) and compare them with all other users. A similar rank is also maintained for groups.

3 Translations and Dictionary Managing

3.1 Dictionaries

FlickLing handles translations in up to six languages: Dutch, English, French, German, Italian and Spanish. We built our dictionaries from the freely available XDXF Dictionaries[3]. The sources required some simple normalisation and symmetrisation before building the translation tables. XDXF Dictionaries were chosen because, in spite of their poor quality in some domains, they were general purposes dictionaries and offered a reasonable coverage between the six languages considered in FlickLing (EN, ES, DE, NL, IT, ES).

[3]http://xdxf.revdanica.com/down

Since we didn't have linguistic tools for all languages involved we decided to match query terms with dictionary entries using stemming. Each dictionary entry was stemmed using a Snowball implementation[4]. This decision have obvious pros and cons: inflected terms can be translated without the need for lemmatisers, but occasionally the dictionary happens to group translations from different words under the same entry. This adds semantic ambiguity that can eventually be fixed via user's manual intervention.

3.2 Term-to-term translation strategy

Once our dictionaries are built, FlickLing automatic translation provides a term-to-term translation. From a given query, the system performs a simple tokenisation (using blank spaces and punctuation as token limits) and each token is stemmed just as with the dictionary entries (see Section 5 for examples on each of these steps).

The translation strategy was a delicate issue. Search is performed via the Flickr API, which permits full boolean queries but does not support term weighting or synonym operators. The best option is then launching the boolean query `(term1 OR term1translation1 OR term1translation2 OR ...) AND (term2 OR term2translation1 OR term2translation2 OR ...) AND ...` In this situation, retaining all translations - as many effective cross-language IR systems do - is not a reasonable strategy.

Therefore we decided to filter translations using two criteria: one is query independent, and the other one takes the query context into account. The system orders the candidates according to:

1. The edition distance between the term and the translation candidate, which is a useful heuristic in many practical situations.

2. The matching with the related terms provided by Flickr. This proved to be a very useful heuristic, because many Flickr users add tags in more than one language to their images. Therefore, Flickr related terms often suggest translations for the query terms, even if Flickr does not explicitly support cross-language suggestions. When a potential translation for a query term appears in Flickr suggested terms, this is a good indicator that it is a good translation in the context of the query.

3. The preferences of the user's personal dictionary. If the user has previously selected a dictionary translation for a given term, or added a new translation, this is taken as evidence in favour of such translation. This is a non-contextual hint, because the user might have manipulated the term translations in a different search scenario (i.e. with a different query context).

The best candidate is chosen by the multilingual interface as preferred translation, while the others are shown (if requested) as translation suggestions (see Figure 4 for a graphical example).

4 Image set

Around 100 images were selected as the target for the Flickling game; there were shown to every user in a randomised order. We researched for images in the six languages of the interface, although the final set of images is not totally balanced, because frequently images annotated in languages other than English also contained some descriptors in English. In the final set, approximately 60% of the images has some English descriptor. Fortunately this had only a relative impact on the multilingual search of the task, first because the target language is not know a priori by the user, and second because the language skills of the actual set of users in the search log (around 300) are very heterogeneous: out of 5101 search sessions, in 2809 cases the target language was an

[4]We used the PyStemmer module, a Snowball-based stemmer for English, German, Norwegian, Italian, Dutch, Portuguese, French, Swedish. See `http://sourceforge.net/projects/pystemmer` for further details.

active language for the user (i.e. she could make queries in that language), in 726 it was passive (i.e. she could partially understand the language but not make queries) and in 1566 was unknown (completely unfamiliar).

Rather than selecting images randomly from Flickr, we wanted to maintain some element of experimental control and topic variation. The following points were considered during selection of the images:

- There should be sufficient text/tags accompanying an image to facilitate the task (i.e. we required rich annotations where possible).
- Ideally we wanted diverse topics in the test set and required roughly equivalent subject/topics in the different language groups, so the aim was to get at least one instance of a subject/topic group, for each of the language sets.
- When collecting images in different languages but with the same subject/topic, we aimed to find images with a similar visual perspective.
- The known item task should not be too hard: queries for finding images were manually recorded and an independent search carried out to check that the images were not too hard to find. This requisite proved hard to meet, because it was highly dependent from the source/target language combination. Some images were easy to find in a monolingual setting but hard to retrieve in a cross-language setting.

A difficulty when selecting images was the large size of the Flickr database. Even using the right keywords, some images might be buried in the ranking simply because there were lots of similar images. A strategy to alleviate this problem was to search images by combining pairs of common tags which were not likely to appear simultaneously in the same image, such as "phone cow". This method was rather successful and, in general, once the user finds the appropriate keywords the image tends to be in the first ten positions of the ranking.

5 User Logs

iCLEF participants could participate in the track doing basically two things: a) interactive experiments recruiting their own set of users and; b) search log analysis and data mining using the Flickling log file distributed by the organisation. In both cases, search logs are the essential source of information. Therefore, how search logs are recorded is an essential part of the Flickling interface.

Each log line corresponds to a unique functionality, even though each interaction may generate more than one log line, as the examples below will show. Log lines have five different fields, separated by the pipeline symbol ('|'):

```
ASCII time|username|session ID|method/functionality|message returned or additional information
```

Some examples illustrating the logs are the following:

1. Successful registration:

   ```
   Tue May 20 15:41:54 2008|user_2|NjDh5AF2JXi05l5|register|succes
   ```

2. Successful login:

   ```
   Tue May 20 15:41:56 2008|user_2||login2|success
   ```

3. A new search session is created and a new target image is assigned:

   ```
   Tue May 20 15:41:56 2008|user_2|NjDh5AF2JXi05l5|getTargetImg|new ActiveSearch created
       photoid 7983304
   Tue May 20 15:41:56 2008|user_2|NjDh5AF2JXi05l5|getTargetImg|success photoid 7983304
   ```

4. A monolingual search in launched and retrieved from the system's cache. Note that time is stopped while querying the database:

```
Tue May 20 15:41:59 2008|user_12|u8ygLdcq7yGK121|search|launch query 'big ben' (100:20)
Tue May 20 15:41:59 2008|user_12|u8ygLdcq7yGK121|pauseTime|time set on pause
Tue May 20 15:41:59 2008|user_12|u8ygLdcq7yGK121|search|500 results retrieved from
    QueryCache
Tue May 20 15:42:00 2008||u8ygLdcq7yGK121|playTime|time set on play mode
```

5. A cross-language search from Spanish into other available languages (each language is denoted by a number) is launched and retrieved from Flickr's database. Again, time is stopped while querying:

```
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|pauseTime|time set on pause
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|launch query 'saltando
    naranja' (0:20)
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|tokens:saltando, naranja
```

Initial query is tokenised and every single query term is stemmed and looked up in the system's dictionaries. Translations and suggestions for each query term are ordered according to: 1) edition distance among all possible translations; 2) matchings with Flickr's related terms and; 3) preferences from users' personal dictionaries:

```
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    for 'saltando' 2->0|sprung|
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    updated with PersonalDict 'saltando'||
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    for 'saltando' 2->1|jump jumping|startle jump start leap bound spring bulge bulk
    leaping bounce jumping
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    updated with PersonalDict 'saltando'||
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    for 'saltando' 2->3|saut|
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    updated with PersonalDict 'saltando'||
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    for 'saltando' 2->4|salto|balzo lancio saltare
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    updated with PersonalDict 'saltando'||
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    for 'saltando' 2->5|sprong|
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    updated with PersonalDict 'saltando'||
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    for 'naranja' 2->0|orange|apfelsine
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    updated with PersonalDict 'naranja'||
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    for 'naranja' 2->1|orange|orangeness orange orange tree
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    updated with PersonalDict 'naranja'||
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    for 'naranja' 2->3|orange|
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    updated with PersonalDict 'naranja'||
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    for 'naranja' 2->4|arancia|arancio arancione
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    updated with PersonalDict 'naranja'||
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    for 'naranja' 2->5|sinaasappel|
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translations and suggestions
    updated with PersonalDict 'naranja'||
```

The final query is structured using Flickr syntax (boolean query), as in:

```
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|translatedQuery
    '(sprung OR jump OR jumping OR saltando OR saut OR salto OR sprong) AND
    (orange OR orange OR naranja OR orange OR arancia OR sinaasappel)'
Tue May 20 15:45:54 2008|user_1|s51DItDdyW81j2M|clsearch|query not found in DB,
    let's use Flickr's API
Tue May 20 15:45:56 2008|user_1|s51DItDdyW81j2M|clsearch|saved QueryCache
Tue May 20 15:45:56 2008|user_1|s51DItDdyW81j2M|clsearch|500 results retrieved
    from Flickr
```

6. The user asks for some hints:

```
Tue May 20 15:48:35 2008|user_5|6l4sFWVeK5ETXBi|getPhotoHint|1 hints showed for
    photo 363836021
Tue May 20 15:51:12 2008|user_5|6l4sFWVeK5ETXBi|getPhotoHint|2 hints showed for
    photo 363836021
```

7. The user finds her first image and wins 10 points (after having asked for some hints). Score, precision and average time are updated accordingly.

```
Tue May 20 15:44:45 2008|user_5|6l4sFWVeK5ETXBi|log|click on found it 704702302:
    success
Tue May 20 15:44:45 2008|user_5|6l4sFWVeK5ETXBi|foundImg|photo 704702302 at
    98.9274973869 secs
Tue May 20 15:44:45 2008|user_5|6l4sFWVeK5ETXBi|foundImg|score updated + 10:
    S:10 P:1.0 T:98.9274973869
```

8. The user gives up searching photo 7983304:

```
Tue May 20 15:49:47 2008|user_2|NjDh5AF2JXi05l5|giveUp|photo 7983304 at
    104.460667133 secs
Tue May 20 15:49:47 2008|user_2|NjDh5AF2JXi05l5|giveUp|score updated:
    S:0 P:0.0 T:0.0
```

9. The user fills in the questionnaire after giving up. Each question has a numeric identifier.

```
Tue May 20 15:49:47 2008|user_2|NjDh5AF2JXi05l5|log|giveUpQuestionnaire:0:false;
    1:false;2:false;3:false;4:false;5:true;comments:I can't recognize the animal
    in the image, because the thumb is to small and/or I'm so bad in biology.
```

For a more comprehensive description of the user logs generated by FlickLing and the mappings of the questionnaires and language identifiers, see the help file distributed among the iCLEF participants[5].

6 Questionnaires

After either finding or giving up on a target image, the user is asked to fill in a small post-image questionnaire with questions about her impressions about the task performed (see Figure 10). The questionnaire after finding an image is the following:

1. It was easy.
2. It was hard because of the size of the image set
3. It was hard because the translations were bad
4. It was difficult to describe the image
5. It was hard because I didn't know the language in which the image was annotated
6. It was hard because of the number of potential target languages

[5]The help file is available at `http://nlp.uned.es/iCLEF/2008/iclef-logs-help.txt`.

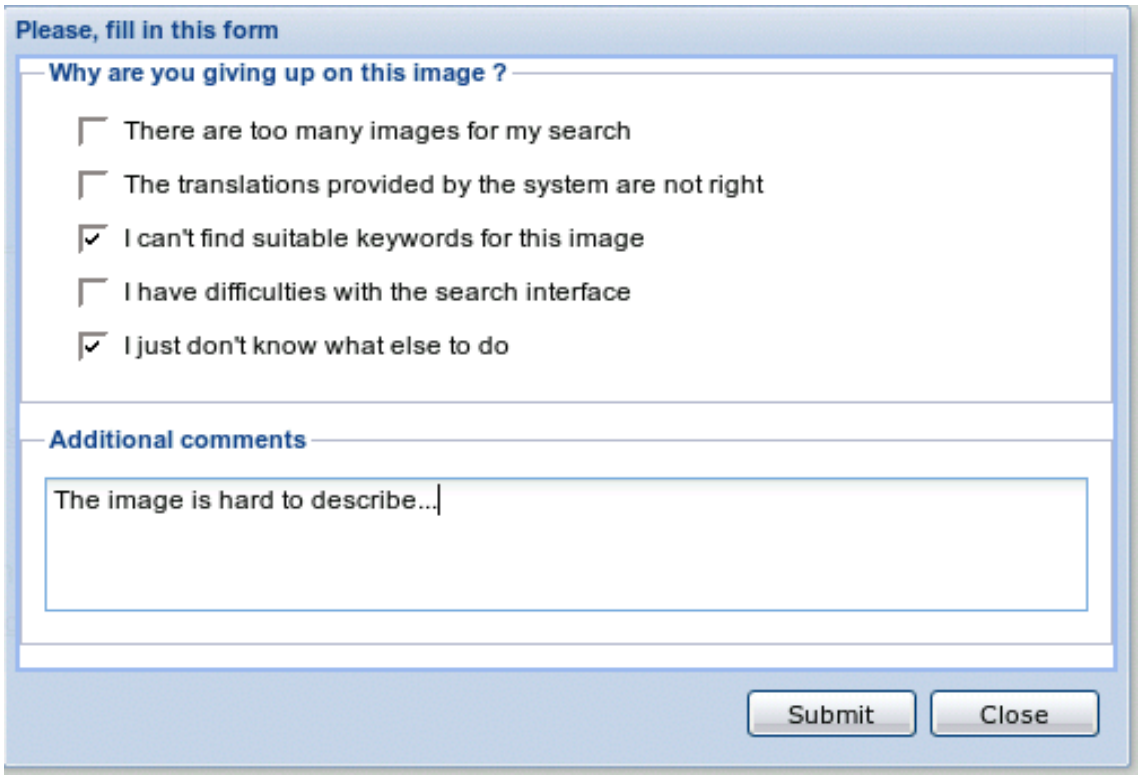

Figure 10: Post-image questionnaire after giving up

7. It was hard because I needed to translate the query

If the user gives up before finding the image, the questionnaire is then:

1. There are too many images for my search
2. The translations provided by the system are not right
3. I can't find suitable keywords for this image
4. I have difficulties with the search interface
5. I just don't know what else to do
6. Additional comments

Lastly, regardless the results of her searches, after searching 15 images, the user was shown a "final" questionnaire[6]. In this case, the questionnaire asks the user to describe her overall experience giving details about her behaviour, the most challenging aspects of the task, the problems found and her opinions about the interface functionalities, to name a few. The complete questionnaire is the following:

1. Do you need to search information in foreign languages in your daily life?
2. Do you often use image search facilities?
3. The search task you performed was:

[6]After this final questionnaire, the user can continue playing at will. We established this threshold after the 15th search session as a good indicator of an extensive experience using FlickLing.

- clear
- easy
- familiar
- interesting
- relevant to you

4. Did you find multilingual search capabilities useful to find images in Flickr?
5. Would you now prefer to use a multilingual search facility for your own image searches?
6. Which, in your opinion, are the most challenging aspects of the task?
   - Selecting/finding appropriate translations for the terms in my query
   - Handling multiple target languages at the same time
   - Finding the target image in very large sets of results
7. Which interface facilities did you find most useful?
   - The automatic translation of query terms
   - The possibility of improving the translations chosen by the system
   - The additional query terms suggested by the system ("You might also want to try with...")
   - The assistant to select new query terms from the set of results
8. Which interface facilities did you miss?
   - Detection and translation of multi-word expressions
   - Bilingual dictionaries with a better coverage
   - A system able to select the translations for my query terms better
   - More support to decide what the possible translations mean and therefore which ones are really appropriate
   - The possibility to search according to visual features of the images ("search images that look like this", "search only B/W images ", "search only for dark images")
   - The classification of search results in different tabs according to the image caption language(s)
   - An advanced search mode giving more control on how Flickr is queried
   - Other (specify)
9. How did you select/find the best translations for your query terms?
   - Using my knowledge of target languages whenever possible
   - Using additional dictionaries and other online sources
   - I did not pay attention to the translations, I just trusted the system

The results of this survey are discussed in the iCLEF 2008 overview paper [1].

7 FlickLing Components and Technology Used

FlickLing search engine consists of two basic components:

- the graphical user interface which controls the search engine features (queries, ranking of results, etc.), functionalities related to cross-linguality (translations, suggestions, related term, etc.) and the game-like features of the task (flow of images, users ranking, etc.)
- a set of web services working behind that are in charge of accessing Flickr, managing the system's databases and generating user logs.

For the development of FlickLing only free software tools have been used. The user interface is implemented entirely in Java and Google Web Toolkit (GWT)[7]. GWT is an open source Java development framework for creating AJAX-based interfaces[8]. It is a web development technique for creating interactive Web applications using HTML, Javascript, CSS and XML web applications. GWT provides many functionalities, tools and performance improvements that are difficult to manage when developing at a lower level. The Flickling GUI has made extensive use of features like application localisation, history management and asynchronous remote procedure calls, to name a few. In order to extend the graphical capabilities of GWT we used the GWT-Ext library[9].

On the other hand, FlickLing web services are implemented in Python using the TurboGears[10] web application framework. The communication with Flickr is rather simple (FlickLing only interacts with Flickr in order to launch queries and retrieved results) and it is established through the Beej's Python Flickr API[11]. The system's databases are structured as MySQL[12] tables and they store users' information (e.g. registration information, language skills, target images shown, score, etc.), the system cache (e.g. queries launched to Flickr and their associated relevant results, image descriptions, etc.), dictionaries and the pool of target images.

Finally, data interchange between the user interface and the web services is performed through JSON[13] objects.

8 Conclusions

In this paper we have presented FlickLing, a search interface for Flickr designed and implemented by the NLP & IR Group at UNED to support the CLEF 2008 interactive task [1]. The interface allows users to search Flickr images in six language simultaneously, offering automatic query translation and support for interactive query translation and refinement. Flickling features multilingual image search as a competitive online game, where a "hall of fame" score ranking is intended to engage a large number of users and record a large search logs which serves as the raw data for the log analysis task at iCLEF 2008.

The user logs generated by the search engine in a period of around 60 days have been distributed among the iCLEF participants. They contain information from more than 5,000 searches (user/image pairs) conducted by more than 200 users from five continents and with a wide variety of language skills. This is, to our knowledge, the largest cross-language controlled search log available for scientific research, and we hope that it will permit to advance in the knowledge of multilingual search processes from a user's perspective.

[7] See `http://code.google.com/webtoolkit`.
[8] AJAX stands for *Asynchronous JavaScript And XML*.
[9] See `http://www.gwt-ext.com`.
[10] See `http://www.turbogears.org`.
[11] See `http://flickrapi.sourceforge.net`.
[12] See `http://www.mysql.com`
[13] JSON stands for *JavaScript Object Notation*. See `http://json.org` for further information.

Acknowledgements

This work has been partially supported by the Regional Government of Madrid under the MAVIR Research Network (S-0505/TIC-0267) and the Spanish Government under project Text-Mess/INES (TIN2006-15265-C06-02). We would also like to thank Valentín Sama, Paul Clough, Jussi Karlgren and Carol Peters as early testers of preliminary versions of Flickling.

References

[1] Gonzalo, J., Clough, P., Karlgren, J.(2008): Overview of iCLEF 2008: search log analysis for Multilingual Image Retrieval. In Borri, F., Nardi, A. and Peters, C. CLEF 2008 Workshop Notes.