

Linking user-generated video annotations to the web of data

Michiel Hildebrand¹ and Jacco van Ossenbruggen^{1,2}

¹ VU University Amsterdam, The Netherlands

² CWI Amsterdam, The Netherlands

Abstract. In the audiovisual domain tagging games are explored as a method to collect user-generated metadata. For example, the Netherlands Institute for Sound and Vision deployed the video labelling game *Waisda?* to collect user tags for videos from their collection. These tags are potentially useful to improve the access to the content within the videos. However, the uncontrolled and often incomplete tags allow for multiple interpretations, preventing long term access. In this paper we investigate a semi-automatic process to define the interpretation of the tags by linking them to concepts from the Linked Open Data cloud. More specifically, we investigate if existing web services are suited to find a number of candidate concepts, and if human users can select the most appropriate concept from these suggestions in the context of the video. We present a prototype application that supports this process and discuss the results of a user experiment where this application is used with different data sources.

1 Introduction

Recently, the Netherlands Institute for Sound and Vision collected a large number of user-generated annotations with the video labelling game *Waisda?* [9]. With these time-based user tags the institute aims to improve access to the content within their videos. Currently, the professional annotations only describe videos as a whole, while their users predominately place orders for fragments within a video [7]. However, the uncontrolled nature of the user tags make it difficult to guarantee long term access. In general, the tags allow for multiple interpretations: named entities, such as persons, typically contain only part of the name (e.g. only the first or last name), and (ii) the subject terms are only available in the vocabulary used by the players of the game, which might not coincide with the vocabulary of a searcher.

A typical solution on the Semantic Web to define the interpretation of a textual value is by linking it to a concept defined by a vocabulary publicly available on the web. This process is also known as “reconciliation”³. Typically, this is a semi-automatic process where a reconciliation service suggests a number

³ The term was coined by the authors of the data cleaning tool Freebase Gridworks, now Google Refine.

of candidate concepts and the user selects the most appropriate one. In this paper, we investigate if this approach can be applied to the tags of the user-generated video annotations from *Waisda?*.

More specifically, our research questions are:

1. what is the coverage of different reconciliation sources with respect to *Waisda?* tags?, and,
2. to what extent can human users use these services to quickly and correctly select the most appropriate concept in context of a video?

To investigate these questions, we focus on two services. First, Freebase⁴ provides a reconciliation service for their community build database, including structured information from Wikipedia. Second, the semantic layer of Europeana⁵ provides a reconciliation service for several controlled vocabularies of (cultural) institutions, including the in-house vocabulary of the Netherlands Institute for Sound and Vision. To test reconciliation for the *Waisda?* tags against these datasources, we implemented a user interface that allows the user to (i) select a data source to reconcile against, and (ii) select a concept from this source that is appropriate in context of the video. We then conducted a small user experiment where four participants used this interface to reconcile the *Waisda?* tags associated with a historic newsreel, and analyzed the results.

In the remainder of this paper we first describe in Section 2 the *Waisda?* video labelling game, discuss the nature of the collected tags and demonstrate how reconciliation of these tags could enable long term access. Next, in Section 3, we briefly review related work on tag reconciliation. Our approach for reconciliation of user-generated video annotations is explained in Section 4. We describe the user experiment with our prototype implementation in Section 5, and discuss the lessons learned from this experiment in Section 6. Finally, we wrap-up the paper in Section 7 by discussing the limitations of our approach and directions for future work.

2 *Waisda?* video labelling game

Waisda? is a multiplayer game where users describe videos by entering tags [9]. The development of the game was initiated and guided by the Netherlands Institute for Sound and Vision and developed by a Web development company. In the first pilot project the game was used to annotate digitized historic newsreels as well as more recent TV episodes from a Dutch broadcaster. The homepage of *Waisda?* contained four channels, each continuously streaming videos from a predefined category. A player starts a game by selecting one of the channels, and plays against other players that joined the same channel. Figure 1 shows a screenshot of the game with a video from a Dutch TV episode. During the game the user can score points by entering tags in the textfield, shown below the video.

⁴ <http://www.freebase.com>

⁵ <http://semanticweb.cs.vu.nl/europeana/>



Fig. 1. Screenshot of the *Waisda?* labelling game, where a user is tagging an episode of a Dutch TV talk show.

Based on the principles of Louis van Ahn’s ESP image labelling game [14], points are scored when two players enter the same tag. As in the Yahoo! Video Tag game [13] this notion of the same tag is extended to a 10 second time-interval. The tags added by the user are shown below the textfield, and the points scored with a tag are indicated by the different colors. The ranking of the players in the current game are displayed to the right of the video.

In the first pilot with *Waisda?* more than 420,000 tags were added to 612 videos, an average of almost 700 tags per video. Table 1 shows a list of tags assigned to a newsreel about the 1937 visit of the Dutch royal family to the coronation of George VI. We will use this example throughout the rest of the paper. As the game targeted a Dutch audience the tags are primarily in Dutch. They contain person names, such as “Elisabeth” and “Juliana”, locations, such as “England”, and subject terms, such as “paarden” (horses in Dutch) and “beefeater”.

Searching within videos Based on previous work [4] we expect that the time-based annotations collected in *Waisda?* can be used to support users with the task of finding objects within a video. For example, the time-based tags allow the user to directly navigate to the point in the video showing “beefeaters”, or find specific content by searching through the tags. By reconciling the tags against concepts this functionality can be enhanced in several ways. The screenshot in

gouden koets, wegrijden, elisabeth, god save the king, hye park, westminster abbey, abbey, priester, geestelijken, Hyde, militairen, kanonnen, beefeater, regen, hek, paarden, zwart, straat, tocht, aankomst, kerk, intocht, stoet, koets, kroning, mensenmassa, parade, rust, juliana, koning, kroon, niets, engeland, bernhard, park, troon

Table 1. Example list of *Waisda?* tags assigned to the George VI coronation newsreel from 1937, containing person names, locations and subject terms.

Figure 2 shows a prototype that supports search and browsing within a video, which is using the tags that were reconciled in the user experiment described in Section 5. In the middle it contains a video player and to the left the reconciled tags. The reconciliation has uniquely identified the tags, as shown by the full name for the persons and locations. For example, the concept with label “Prince Bernhard of Lippe-Biesterfeld” instead of the tag “Bernhard”. The type information available for the concepts enables a categorization of the tags, as shown on the left side of the screenshot. The rich information of the concepts extends the possible ways to find tags. For example, by reconciliation to Freebase the subject term “beefeater” can also be found by the synonym “yeamon warders”. The mappings from the Dutch version of WordNet, Cornetto, to the English version enable multilingual access. Finally, the links to the concepts provide rich background information for each tag, as shown to the right of the video in the screenshot of Figure 2.

3 Related work

This work is inspired by Google Refine, a tool to clean and transform tabular data⁶. Among other operations it allows the user to reconcile data values to the concepts from an external source. By default Google Refine suggests topics (locations, persons, movies etc.) from Freebase. The task of the user is then to select the most suited suggestion in context of the table row in which the value occurs. For our task, Google Refine is not directly suited as the *Waisda?* tags have to be reconciled in context of the video.

Google Refine requires a reconciliation service to be able to return a ranked list of candidates for a given string. Most Linked Data providers do not yet support this service. Therefore, Maali et.al. investigated how this reconciliation can be supported by existing Linked Data services [8], such as the standard query language for the Semantic Web SPARQL [11] or the Semantic Web search engine Sindice [12]. Their evaluation showed that it is feasible to use these sources for reconciliation, and they implemented extensions for Google Refine to access these services. It should, however, be noted that reconciliation interfaces require high precision results because they can show only a limited number (typically

⁶ <http://code.google.com/p/google-refine/>

DE KRONINGSPLECHTIGHEDEN

De kroning van de nieuwe Britse koning, George VI, te Londen. SHOTS: - aankomst prinses Juliana en prins Bernhard per schip te Harwich (Groot-BrittanniA*) waar zij worden ontvangen door vertegenwoordiger van Britse koninklijke huls en Nederlandse gezant te Londen, waarna zij voor pers poseren; - lange, feestelijke stoet met oa koels waarin George VI en zijn vrouw Elizabeth, opweg naar Westminster Abbey; - aankomst bij Westminster Abbey; - kroningsplechtigheden in kerk; - vertrek vanaf kerk, gevolgd door tocht door Londen waarin naast koets met koning en

The screenshot displays a video player interface. On the left, there is a search bar and a list of tags categorized by type: Person/Organization (Juliana of the Netherlands, Prince Bernhard of Lippe-Biesterfeld), Location (England, Hyde Park, Westminster Abbey), and Subject (aankomst, antigeluid, hek, intocht, godshuls). The video player shows a scene with Prince Bernhard and Queen Juliana. A progress bar indicates the video is at 00:28 of 05:07. Below the video, a thumbnail for 'bernhard 0:20' is visible. On the right, a sidebar provides background information for Prince Bernhard of Lippe-Biesterfeld, including his date of birth (Jun 29, 1911), place of birth (Jena), and date of death (Dec 1, 2004). It also includes a brief biography and a list of awards: Noble person, Award Winner, Deceased Person.

Fig. 2. Screenshot of prototype with enhanced access to tagged video. To the left of the video the reconciled tags are categorized by their type. To the right of the video background information of the current content is shown.

in the range from three to ten) of suggestions to the user to choose from. In general, such a high precision is only reached by using additional restriction on the type of the results, e.g. persons. In our task this information is not available.

There are several approaches to link tags to concepts fully automatically. For example, for pictures on [Flickr.com](#) [1, 10] or videos on [Youtube.com](#) [3]. We believe that such automatic techniques could also be applied to link the user tags from *Waisda?* However, it is unlikely that the precision of such algorithms will reach the quality standards of an archive for long term preservation. Therefore, the need for human assessment remains. Furthermore, this work focuses on the use of existing reconciliation services to find candidate concepts. We hope our use case inspires the integration of advanced ranking algorithms into such services. Finally, the interactive approach presented in this paper also gives us the opportunity to collect a golden standard, which we can use for the evaluation of more advanced suggestion algorithms in future work.

4 Linking video annotations to concepts

In this section we present our semi-automatic approach to link the tags collected in *Waisda?* to concepts from the Linked Data cloud. At the backend we use existing reconciliation services to collect a number of candidate concepts, and at

a front-end we provide a graphic user interface that allows the user to select the appropriate candidate.

4.1 Reconciliation API

The reconciliation API, as introduced in Google Refine, is a web service that links textual values to database identifiers. The service is intended to be used semi-automatically, where the service suggests a number of candidate concepts and the human user selects the concepts appropriate for her case. The algorithm to match a query to a concept is not specified in the API. Typically, it performs a fuzzy string match between the query and the textual labels of the concepts, for example the name of a location or person.

The main parameters of the service are straightforward: a query object with one or more textual values, an optional specification of the type of the results, and a limit on the number of results returned⁷. The service returns for each query an identifier, the name, all available types, the score produced by the algorithm and a boolean that is true when the service is confident enough to indicate the match as the right candidate.

4.2 Linked Open Data sources

The Linked Open Data initiative has inspired many data providers to publish their datasets on the Web [2]. Popular sources are DBPedia⁸, containing structured data from Wikipedia, and Geonames⁹ as a source for geographic locations. We expect that these sources are useful for reconciliation of the *Waisda?* tags. However, the data providers do not provide a reconciliation service. Recently, Talis launched an initiative to provide several services for several datasets from the Linked Data cloud, including reconciliation¹⁰. However, after exploration we observed that the ranking of the results was not sufficient for our purposes. Freebase provides an alternative to these data sources, is also available as Linked Data and provide a public reconciliation service.

Europeana provides another useful Linked Data source for our tag set. The semantic layer of Europeana contains a large number of controlled vocabularies from different (cultural) institutes. Within the reconciliation service each data-source can be accessed separately. In particular, we consider two sources to be relevant for the tags in *Waisda?*. First, the Netherlands Institute for Sound and Vision uses an in-house thesaurus for the documentation of audiovisual content. This so-called GTAA thesaurus (Dutch acronym for Common Thesaurus Audiovisual Archives) contains approximately 160,000 terms in six facets: subjects, locations, person names, organization names, maker names and genres.

⁷ The full specification is available at: <http://code.google.com/p/google-refine/wiki/ReconciliationServiceApi>

⁸ <http://www.dbpedia.org>

⁹ <http://www.geonames.org>

¹⁰ <http://www.kasabi.com>

DE KRONINGSPLECHTIGHEDEN

De kroning van de nieuwe Britse koning, George VI, te Londen. SHOTS: - aankomst prinses Juliana en prins Bernhard per schip te Harwich (Groot-Brittannië) waar zij worden ontvangen door vertegenwoordiger van Britse koninklijke huis en Nederlandse gezant te Londen, waarna zij voor pers poseren; - lange, feestelijke stoet met oa koets waarin George VI en zijn vrouw Elizabeth, opweg naar Westminster Abbey; - aankomst bij Westminster Abbey; - kroningsplechtigheden in kerk; - vertrek vanaf kerk, gevolgd door tocht door Londen waarin naast koets met koning en



Fig. 3. Screenshot of tag reconciliation interface.

Second, Cornetto, a WordNet-like lexical semantic database of Dutch that contains 70,000 synsets [15].

4.3 User interface

The screenshot in Figure 3 shows the user interface of the prototype application. On the right it contains the current video, on the left it contains the list of tags assigned to this video and in between them a list of frames for the currently selected tag. The user starts a session by selecting a reconciliation source, shown below the tag list. The reconciliation is performed by sending a request to the corresponding service. When the service returns results for a tag an icon is added to the right of the tag. Selecting such a tag then shows these concepts below the tag. To support the user in identifying the concepts, they are represented by their label and their types. The selection of a tag also brings up the frames that are annotated by this tag, and clicking a frame will play the video starting at this frame. Finally, when the user selects a concept a request is sent to the server to update the database and the tag is highlighted in the interface.

5 User experiment

We performed a user experiment to test the coverage of the different data sources for the *Waisda?* tags and to what extent users are able to select the most appropriate concept in context of the video. Our assumption is that the users are somewhat familiar with the vocabularies used, and know what kind of concepts

	tags	unique for source	concepts
Freebase	13	3	15
GTAA	12	3	12
Cornetto	22	15	25
total	33		

Table 2. Distribution of reconciled tags over the different sources.

they can expect. In this experiment we focus on a quick and simple method to link tags to concepts. Therefore, we expect the user to select one of the concepts suggested by a reconciliation service, and we do not include functionality to first fix the spelling of tags or manually search within a vocabulary in case no suited concept is found. Furthermore, we expect that a suited concept should be part of the first 10 suggestions provided by the service.

For the experiment we used the video presented in Section 2, a newsreel about the visit of the Dutch royal family to the coronation of George VI. In total the video contained 36 tags. We asked four colleagues from our department to use the prototype interface to select the most appropriate concept for each tag. They were instructed to only select a concept when an appropriate one was among the suggestions. They could use the frames related to a tag and play the video add that time point, but were not forced to do so. The users reconciled the tags against three different sources: GTAA, Cornetto and Freebase. A session started with a short explanation of the interface using a different video. In the experiment the participants performed the following steps: select a datasource, start the reconciliation service by clicking the “Go” button, and once results returned start selecting concepts. When all tags were considered the user continued with the next datasource.

5.1 Results

For 33 (out of 36) tags a concept was selected by one or more participants. The three tags for which no concept was selected all contained a spelling error that prevented the reconciliation services to find the appropriate concept: “hye park”, “elisabeth” instead of “elizabeth”, and “millitairen” instead of “militairen”. Table 2 shows the distribution of these tags over the different data sources. Using Cornetto as the reconciliation source most concepts were selected (22), while for GTAA and Freebase less than half of the tags could be linked.

Table 3 shows the number of tags that were reconciled per participant. These numbers are comparable to the total number of tags reconciled by all participants. The low number of tags selected from GTAA by P3 can be explained by the fact that this participant only selected entities from this source and no subject terms, while the others did select the subject terms. Further manual

participants:	P1	P2	P3	P4	α
Freebase	11	10	9	8	0.78
GTAA	9	12	5	10	1.00
Cornetto	21	20	20	18	0.92
total	32	32	27	29	0.91

Table 3. Number of tags reconciled by the four participants. The final column shows inter-rater agreement using Krippendorff’s alpha.

examination showed that for each source most of the tags were reconciled by 3 or 4 users. For GTAA and Cornetto only one tag was reconciled by only one user, and for Freebase 2 tags. This already gives some intuitive indication that the participants, to a large extent, agreed upon which tags to reconcile.

To formally compute the agreement between our participants, we need a suitable metric. Because we have more than four raters, and missing data for some of the tags, we use Krippendorff’s alpha (α) as a reliability coefficient [5]. Note that $\alpha = 1.0$ indicates perfect agreement (e.g. in the GTAA cases) while $\alpha = 0.0$ indicates a level of agreement that can be completely attributed by chance. The $\alpha = 0.78$ for Freebase concepts can be interpreted as moderately high agreement, but is quite low compared to the other data sources. This can be partly explained by the fact that we treated both truly distinct concepts and closely related concepts as disagreement in the computation of α . All other α s are above 0.9, which is generally interpreted as a high level of agreement. We conclude that for this small experiment, with a coverage of 33 out of 36 tags and a total reliability coefficient of 0.91, users can effectively reconcile tags using the selected services.

To explain the disagreements among the users we manually assessed the selected concepts. For Freebase the participants selected different concepts for three tags. However, on close examination we considered the selected concepts valid alternatives. For example, two users selected for the tag “abbey” the specific location “Westminster Abbey”, while another selected the general concept for “abbeys”. For Cornetto the differences between the selected concepts were very subtle. For example, the tag “hek” is in Dutch used for “gate” as well as “fence”. Both these interpretations were also applicable to the video. Although the participants agreement on all selected concepts from GTAA we found one error. For the tag “bernhard” (the former prince of the Netherlands) three of the four users selected a Dutch interviewer with the name “Prins, Bernhard” (“prins” is Dutch for “prince”). The participants overlooked the type shown below this name indicating that this was an interviewer.

We also analysed how many of the tags were only found in a single source, shown in Table 2 in the column labeled *unique*. This shows that most tags that were reconciled against Freebase and GTAA could also be found in

rank	1	2	3	4	5	6	7	8	9
#concepts	32	9	2	6	0	1	2	1	2

Table 4. Ranks of the selected tags.

another source. More specifically, the general subject terms, for example “park”, were also selected from Cornetto. Most entities, such as persons and locations were, however, only found in GTAA or Freebase. All entities selected from GTAA were also found in Freebase. The majority of the tags selected from Cornetto were only found in this source. However, for some tags there were subject terms selected from GTAA that were not found in Cornetto. All these tags were plural forms of regular words missed by the reconciliation service of Cornetto.

Finally, Table 4 shows for the selected concepts at which position in the list of candidate concepts they were ranked. We excluded the falsely selected concept mentioned above. More than half of the selected concepts were ranked first (32), while 23 concepts were ranked second or lower. Several (6) selected concepts were even ranked sixth or lower.

6 Lessons learned

Although the experiment involved only one video and 36 tags, we can make several observations about the data sources, services and user interface.

- Cornetto is a promising source for subject terms. For most of the subject tags in the experiment the participants could select a concept from Cornetto. Only tags with irregular plural forms were not found. More flexibly string matching techniques, or a better stemming algorithm for Dutch, could help to also find concepts in these cases. However, selecting the most suited concept from Cornetto in the user interface can be time consuming, as the differences between interpretations can be very subtle. In contrast, for GTAA there is usually only one concept used for all interpretations.
- For all the named entities the participants could select a concept from Freebase as well as GTAA. However, the additional information in GTAA is sparser than in Freebase, making it more difficult to identify the right concept. The coverage of Freebase worked well for the video used in the experiment, as the persons and locations are well known. For specific Dutch content Freebase might, however, be less suited.
- Full coverage was prevented by spelling errors and variations in the tags. However, in some cases the video fragment also contained the correctly spelled tag.
- When users know what to expect they adapt their behavior. For example, they quickly noticed that Freebase was most suited for persons and locations, making them only briefly scan the list for subject terms.

- For most tags the context in the video was already clear from the video description or the other tags. However, the video and the frames corresponding to the tags were used in several occasions, and the participants remarked that they were sometimes essential to determine the correct interpretation.

Based on these observations we derive three recommendations for configuring a reconciliation interface for video annotations. First, we recommend to reconcile against multiple data sources at the same time, as the best coverage is acquired by a combination of sources. To effectively use the combined results we recommend that duplicates are merged into a single suggestion. This can, for example, be done using alignments between the concepts in different sources [6]. In addition, the results are best presented in different categories, for example persons, locations and subjects. This allows the user to quickly ignore suggestions from a wrong type. Second, to deal with spelling variations and errors we recommend a preprocessing step that merges similar tags. Third, to simplify the selection of the most suited concept the results from sources that provide high precision should be presented first. The results found in sources that provide high recall should be presented as an alternative. The same approach can be used for a single datasource to distinguish the results found by high precision or high recall string matching techniques.

7 Discussion

We showed that it is feasible to semi-automatically reconcile the user tags collected in *Waisda?* against open data available on the Web. We acknowledge that the small scale of the user experiment prevents us from making more general conclusions. Therefore, we are planning a large scale experiment in the context of a second pilot with *Waisda?*. Based on the lessons learned in this research we will improve our prototype and make it suitable for large scale use. We hope that the integration with *Waisda?* attracts users to reconcile their own annotations or those made by others. We believe, that the data collected in such a large scale experiment can be valuable for the Netherlands Institute for Sound and Vision to improve access to their collection. In addition, such data can be valuable for the research community as a golden standard for concept suggestion algorithms and automatic concept detection.

Acknowledgements

This research was partially supported by the PrestoPRIME project, funded by the European Commission under ICT FP7 (Seventh Framework Programme, Contract No. 231161)

References

1. S. Angeletou, M. Sabou, and E. Motta. Semantically enriching folksonomies with flor. In *In Proc of the 5th ESWC. workshop: Collective Intelligence and the Semantic Web*, 2008.
2. C. Bizer, T. Heath, and T. Berners-Lee. Linked data - the story so far. *Int. J. Semantic Web Inf. Syst.*, 5(3):1–22, 2009.
3. S. Choudhury, J. G. Breslin, and A. Passant. Enrichment and ranking of the youtube tag space and integration with the linked data cloud. In A. Bernstein, D. R. Karger, T. Heath, L. Feigenbaum, D. Maynard, E. Motta, and K. Thirunaryan, editors, *International Semantic Web Conference*, Lecture Notes in Computer Science, pages 747–762. Springer, 2009.
4. R. Gligorov, M. Hildebrand, J. Van Ossenbruggen, G. Schreiber, and L. Aroyo. On the role of user-generated metadata in audio visual collections. In *Proceedings of the International Conference on Knowledge Capture (K-CAP)*, pages 145 – 151. ACM Press, June 2011.
5. A. F. Hayes and K. Krippendorff. Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1):77–89, 2007.
6. M. Hildebrand, J. R. van Ossenbruggen, L. Hardman, and G. Jacobs. Supporting subject matter annotation using heterogeneous thesauri, a user study in web data reuse. *International Journal of Human-Computer Studies*, 67(10):888–903, October 2009.
7. B. Huurnink, L. Hollink, W. van den Heuvel, and M. de Rijke. Search behavior of media professionals at an audiovisual archive: A transaction log analysis. *Journal of the American Society for Information Science and Technology*, 61(6):1180–1197, June 2010.
8. F. Maali, R. Cyganiak, and V. Peristeras. Re-using cool uris: Entity reconciliation against lod hubs. In *In Proceedings of the 4th Linked Data on the Web (LDOW) Workshop at the World Wide Web Conference (WWW)*, 2011.
9. J. Oomen, L. Belice Baltussen, S. Limonard, A. van Ees, M. Brinkerink, L. Aroyo, J. Vervaart, K. Asaf, and R. Gligorov. Emerging practices in the cultural heritage domain - social tagging of audiovisual heritage. In *Proceedings of the WebSci10: Extending the Frontiers of Society On-Line*, Raleigh, US, April 2010.
10. S. E. Overell, B. Sigurbjornsson, and R. van Zwol. Classifying tags using open content resources. In R. A. Baeza-Yates, P. Boldi, B. A. Ribeiro-Neto, and B. B. Cambazoglu, editors, *WSDM*, pages 64–73. ACM, 2009.
11. E. Prud'hommeaux and A. Seaborne. SPARQL Query Language for RDF. W3C Recommendation, 2008.
12. G. Tummarello, E. Oren, and R. Delbru. Sindice.com: Weaving the open linked data. In *Proceedings of the 6th International Semantic Web Conference and 2nd Asian Semantic Web Conference (ISWC/ASWC2007), Busan, South Korea*, pages 547–560, Berlin, Heidelberg, November 2007.
13. R. van Zwol, L. Garcia, G. Ramirez, B. Sigurbjornsson, and M. Labad. Video tag game. In *Proceedings of the 17th International World Wide Web Conference (WWW 2008)*, Beijing, China, April 2008.
14. L. von Ahn and L. Dabbish. Labeling images with a computer game. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 319–326, New York, NY, USA, 2004. ACM.
15. P. Vossen, I. Maks, R. Segers, and H. van der Vliet. Integrating lexical units, synsets and ontology in the Cornetto database. In E. L. R. A. (ELRA), editor, *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, 2008.