Text-based annotation of scientific images using Wikimedia categories

Frieda Josi[1][0000−0001−7124−780X], Christian Wartena[1][0000−0001−5483−1529], and Jean Charbonnier[1][0000−0001−6489−7687]

University of Applied Sciences and Arts Hanover,
Expo Plaza 12, 30539 Hanover, Germany
christian.wartena@hs-hannover.de

Abstract. The reuse of scientific raw data is a key demand of Open Science. In the project NOA we foster reuse of scientific images by collecting and uploading them to Wikimedia Commons. In this paper we present a text-based annotation method that proposes Wikipedia categories for open access images. The assigned categories can be used for image retrieval or to upload images to Wikimedia Commons. The annotation basically consists of two phases: extracting salient keywords and mapping these keywords to categories. The results are evaluated on a small record of open access images that were manually annotated.

Keywords: Scientific image search · Text annotation · Wikipedia categories.

1 Introduction

In order to increase the reuse of scientific images from open access journals, we collect scientific images, make them available in a search engine and upload high quality images to Wikimedia Commons. A beta version of the NOA scientific image search, using the categories extracted by the proposed method is available under: `http://noa.wp.hs-hannover.de` [1].

Metadata such as author and disciplines can be adopted from the papers the images are taken from. However, publishers do not provide no specific metadata for the individual images. In this paper we present a method for extracting detailed categories for each image from its caption and from text fragments referring to the image.

After discussing related work, we will present the data we have used (section 3) and a method that is based on extracting keywords from the captions and related text and assigning categories on the base of the keywords, for which categories are known (section 4). In sections 5 and 6 we present an evaluation carried out on the basis of 100 images from open access journals that were uploaded to Wikimedia Commons and annotated manually.

2 Related Work

Relevant work for using Wikipedia titles and categories focuses on linking to the Wikipedia articles [2] or using article titles for indexing images [3]. The categories of Wikipedia articles were used much less frequently. An example of the use of the Wikipedia category system for annotation is the work by Wartena an Brussee [4].

The extraction of key terms is a well studied area with numerous publications. Popular algorithms are those described by Frank et al. [5], Turney [6] and by Mihalcea and Tarau [7]. Leong et al. [8] explicitly use keyword extraction to describe images. The method described by Frank et al. [5] and Turney [6] uses various features to determine the suitability of a word as keyword. The most important feature still is the inverse document frequency (idf) that was already proposed by Salton [9]. Besides idf we use the distributional similarity of a keyword with the entire text. This method was proposed by [10].

The matching of the extracted terms to the Wikipedia article title is described in Mihalcea and Csomai [2]. Classification of text based on the classification of key terms found in the text was e.g. done by Wartena and Sommer [11].

3 Data Records and Wikipedia Categories

For the development of the annotation method, 397 data records were used. Each data record contains the caption and the sentences referring to the image. These images have been published in open access journals by Copernicus, Hindawi, Frontiers and Springer Open. We use the XML markup provided by the publishers to identify references to each image. We use the whole sentence containing the image reference as a context for the image. The image captions used have an average length of 308 words and 1881 characters. Table 1 shows an overview of the number of words in captions and the complete data record, include sentences referring to the image[1].

Table 1. Number of words in caption and referring sentences in the development, the evaluation data record and in the entire database (application).

Data record	Text	Size	Average	Min	Max
Develop	Caption	397	54	2	503
	Caption+ref. sent.		308	13	2274
Evaluation	Caption	100	46	3	404
	Caption+ref. sent.		326	10	2938
Application	Caption	2,9M	81	0	5268
	Caption+ref. sent.		405	0	43817

[1] For an example of an extremely long capture see Fig. 5 in `http://dx.doi.org/10.1002/ece3.2579`. Also some parsing errors resulted in long captions.

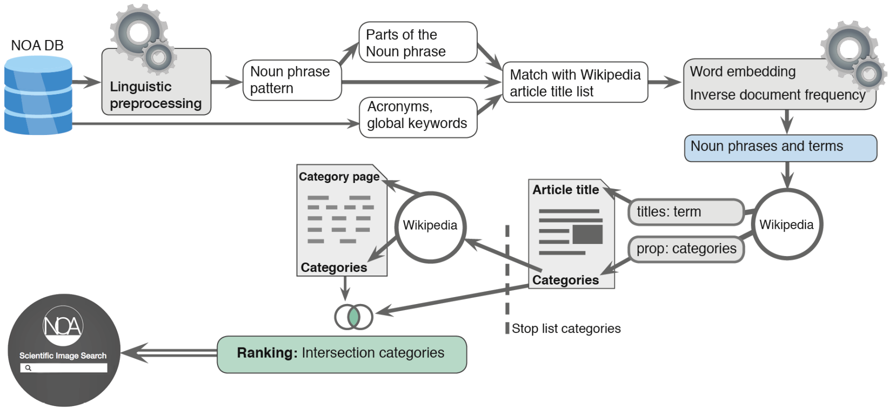


Fig. 1. Text-based annotation model with Wikipedia categories

The whole corpus of the NOA image search engine contains 870,840 articles with images. In total, there are 2.9 M scientific images from the subjects medicine, science, health sciences, biology, technology, chemistry and more. We used the whole corpus to compute the idf-values of all terms.

For evaluation we collected 100 images that were not part of the development set and that seemed to be more or less interesting for the Wikimedia community. i.e. we excluded charts, microscopy images, etc.

For the categorization of the images we used the categories of Wikipedia. The category system in the Wikimedia Foundation projects consists of categories created by volunteers as well as categories from existing norm data. In this way, the norm data of LCCN, and VIAF can be included in the articles of Wikipedia [12]. The article pages in Wikipedia are organized using categories. The categories include theme categories, object categories, structural categories and metacategories [13]. A page can be assigned to several categories. It has to be noted, that Wikipedia categories can be used in Wikimedia Commons, but in principle Wikimedia Commons and the Wikipedia of each language has its own independent category system.

4 Annotating scientific images

Our text-based annotation method consists of two phases: extracting Wikipedia terms and assigning categories[2]. The process flow is shown schematically in more detail in Fig. 1 and described in the following sections.

[2] Our source code will be released together with all developed source codes of the NOA project.

4.1 Term and Noun Phrases Extraction

The linguistic preprocessing (tokenization and part of speech tagging and lemmatization) is carried out with the open source Natural Language Toolkit (NLTK) [14] using the Wordnet lemmatizer for lemmatization.

In order to find words and phrases that are used as the title of a Wikipedia article we search noun phrases according to a simple regular expression over POS tags. Using the syntax for chunking grammars described in [14] we define a noun phrase as:

$$\mathrm{NP} : (<\mathrm{CD}>)?(<\mathrm{JJ}>)* <\mathrm{N(N|P)}.*> + \tag{1}$$

using POS tags from the Penn Treebank Tagset [15]. Thus CD stands for a cardinal number, JJ for adjective and all tags starting with NN or NP for various types of common nouns and proper nouns.

Next we lookup each noun phrase found in Wikipedia (either using the Wikipedia API in the development phase or the SQL-Dump in the application to a larger data set). If the phrase is not found we split the phrase into the first word and the remaining tail. If the first word is a noun, it is looked up in Wikipedia. The tail also is looked up and recursively split until the phrase is found in Wikipedia or no words remain. Thus it is ensured, that only the longer (and more specific) phrase is used if it is a title in Wikipedia, but that the smaller phrases are still used if the longer phrase is not found. E.g., if we find *Greenhouse gas* we don't use the term *gas*, since the whole phrase is found in Wikipedia.

In order to match phrases to Wikipedia titles, we exclude words from a common list of stopwords and we pluralize words if the singular form was not found. All names of albums, magazines, musical groups and films are excluded from matching.

We extract key phrases in this way from the caption, but also from each sentence with a reference to the image. In addition we take all global keywords from the paper (provided by the authors) if they are found in Wikipedia and all expanded acronyms (see [16] for details on acronym resolution in NOA), if they could be expanded automatically and if the complete expansion was found in Wikipedia.

4.2 Term ranking

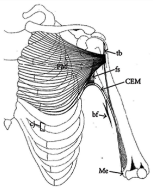
After we have found phrases in the text we have to rank them and select the most promising ones. In absence of suitable annotated data that can be used for training, we use only two features: idf and the similarity of the word embedding of the key phrase with the word embedding of the caption.

The inverse document frequency is calculated with:

$$idf(N) = log\frac{\text{Number of data records in the corpus}}{\text{Number of data records containing N}} \tag{2}$$

In many cases idf alone is not sufficient. In the caption or in the referring context highly specific words can occur that are not related to the image at all.

Table 2. Example of Wikipedia terms found for an image, with source, idf-value and distance of context vector to the caption for each term. Source: r= Referring Context, c = Caption.

<table>
<tr><th>Wikipedia Terms</th><th>source</th><th>idf</th><th>cos</th><th>Wikipedia Terms</th><th>source</th><th>idf</th><th>cos</th></tr>
<tr><td><b>axillary fascia</b></td><td>r</td><td>20.0</td><td>0.72</td><td>inch</td><td>r</td><td>10.9</td><td>0.33</td></tr>
<tr><td>griffith university</td><td>r</td><td>18.1</td><td>0.20</td><td>upper limb</td><td>c</td><td>10.8</td><td>0.65</td></tr>
<tr><td><b>brachial fascia</b></td><td>c</td><td>17.5</td><td>0.77</td><td>humerus</td><td>r</td><td>10.4</td><td>0.62</td></tr>
<tr><td>quartus</td><td>r</td><td>15.7</td><td>0.35</td><td>continuation</td><td>r</td><td>10.2</td><td>0.24</td></tr>
<tr><td>medical literature</td><td>r</td><td>14.4</td><td>0.26</td><td>fascia</td><td>c</td><td>10.0</td><td>0.75</td></tr>
<tr><td>common name</td><td>r</td><td>13.9</td><td>0.26</td><td>nomenclature</td><td>r</td><td>9.4</td><td>0.15</td></tr>
<tr><td><b>deep fascia</b></td><td>r</td><td>13.9</td><td>0.75</td><td>depiction</td><td>r</td><td>9.4</td><td>0.23</td></tr>
<tr><td><b>epicondyle</b></td><td>r</td><td>12.7</td><td>0.76</td><td>rib</td><td>r</td><td>9.3</td><td>0.59</td></tr>
<tr><td>joint capsule</td><td>r</td><td>12.4</td><td>0.58</td><td>informed consent</td><td>r</td><td>9.3</td><td>0.59</td></tr>
<tr><td>queensland</td><td>r</td><td>12.2</td><td>0.16</td><td>wood</td><td>r</td><td>9.2</td><td>0.24</td></tr>
<tr><td>cadaver</td><td>r</td><td>11.4</td><td>0.40</td><td>septum</td><td>r</td><td>9.1</td><td>0.56</td></tr>
<tr><td>axilla</td><td>r</td><td>11.3</td><td>0.56</td><td>thorax</td><td>c</td><td>9.1</td><td>0.58</td></tr>
<tr><td><b>biceps</b></td><td>r</td><td>11.1</td><td>0.69</td><td>...</td><td>...</td><td>...</td><td>...</td></tr>
<tr><td>tubercle</td><td>r</td><td>10.9</td><td>0.57</td><td>number</td><td>r</td><td>2.3</td><td>0.19</td></tr>
<tr><th colspan="8">Image</th></tr>
<tr><td colspan="8">
</td></tr>
</table>

Also the global keywords from the paper might be very good, but sometimes too general for the picture we want to describe. Thus we want to know, how well each term fits to the captions as a whole.

We computed word embeddings for all words in the corpus that occur at least 5 times. We trained an word2vec model using our corpus with a window size of 5 using the CBOW model, an embedding size of 300 and a minimum word occurrence threshold of 5. We removed all tokens from our data that are in the NLTK stopword list or that have less than two characters.

Now, for each word we have an abstract vector of 300 dimensions. We represent both, a key phrase and the caption, by the average vector of all words (excluding stop words) that they consist of. Now the cosine between the vector of the key phrase and the vector of the abstract is a measure for the degree to which the key phrase is representative for the caption.

Table 2 shows the extracted noun phrases and terms for the first image from the article with the DOI: 10.1155/2016/5402081. Also the idf and the cosine between the vectors for the phrase and the caption are given. Here we see, that e.g. *Griffith University* has a very high idf value, but is not very representative for the caption. Here the cosine between the term vector and the caption vector is a much better indicator for relevance.

We will evaluate three variants, as shown in Table 3. In the first variant we use the top five phrases according to idf, in the second variant we use five phrases with the highest cosine similarity and finally we combine both criteria, by using five phrases with the highest cosine similarity taken from the 15 most specific phrases.

The five keywords for our example selected by the third variant are set in bold face in Table 2.

Table 3. Variants of the evaluation

Variant 1: Idf
Variant 2: Cosine
Variant 3: Idf + Cosine

4.3 Category Filtering

We want to assign categories to the images, not keywords (or article titles). The categories assigned to articles corresponding to the best keywords for each image are candidates, but we use also the upper categories of each category as a candidate.

Before selecting the most relevant categories, we remove a large number of categories that are not interesting for our purpose such as *Category:Systems* or categories with meta information for Wikipedia internal purposes, like categories for articles missing references etc. To do so, we filtered out all categories that are classified as hidden category in Wikipedia and all categories that are classified as container category. Futhermore we use a list of regular expression for category names that are filtered out, e.g. all categories that contain the word *Wikipedia* or *stub* or *disambiguation*. Finally, a stop list of further categories is used, containing categories like *Category:Nothing*, *Category:Self* or also meta information like e.g. *Category:ISO_basic_Latin_letters*.

4.4 Category Ranking

We assume that the categories of the articles themselves are more likely to be appropriate than their upper categories and we assume that a category that is the category of two keywords is better than a category that comes only with one keyword. To formalize this intuition, we define the number of keywords associated with a category c at different levels. We say a keyword k is associated at level 0 with c if c is a category of (the Wikipedia article with title) k and c is identical with k. E.g. the article *Fascia* has the category *Fascia*, so the category *Fascia* is associated at level 0 with the keyword *fascia*. We say c is associated at level 1

Table 4. Ranking of categories for the example in Table 2

Category	Value	Category	Value
Fascia	3.0	Limbs (anatomy)	0.4
Muscular system	1.6	Muscles by action and location	0.4
Musculoskeletal system	1.6	Joints	0.4
Soft tissue	1.2	Elbow	0.4
Connective tissue	1.2	Forearm	0.4
Tissues (biology)	1.2	Muscles by location	0.4
Elbow flexors	1.0	Flexors	0.4
Forearm supinators	1.0	Upper limb anatomy	0.4
Muscles of the upper limb	1.0	Muscles by action	0.4
Shoulder flexors	1.0	Shoulder	0.4
Skeletal system	1.0	Organ systems	0.4
Medical Subject Headings	0.8	Dance science	0.4

with k, if c is a category of k but not identical to k; c is associated at level l with k, if it has a subcategory that is associated with k at level $l-1$ (and c is not associated with k at level $l-1$). We denote the number of keywords associated with c at level l as $r_l(c)$. Finally, we define the weight $w(c)$ as:

$$w(c) = \sum_{l=0}^{n} w_n \cdot r_l(c) \tag{3}$$

Where $n \in \mathbb{N}$ and w_n are the level weights. In absence of training data to determine optimal weights, in the following we let $n = 2$ and $w_0 = 1.2$, $w_1 = 1.0$ and $w_2 = 0.4$. For our example this ranking gives the weights shown in Table 4.

It might seem obvious to take the number of relations between the categories as a feature as well, as was done e.g. by [17]. Our experiments indicated however, that this results in a massive preference for categories from areas that are worked out very well. In most cases these are unspecific general areas and high weights for categories with many connections to other found categories therefore suppress specific and precise categories. On the other hand, as we will see below, categories introduced by several keywords usually are quite adequate.

5 Evaluation

The images for the evaluation were uploaded to Wikimedia Commons and manually annotated with categories.[3] The selection criteria for the upload of the images to Wikimedia were a higher probability of reuse, images from current papers, graphics and photos but no schematics and only images with the copyright

[3] The images are available on Wikimedia Commons at the following link: `https://commons.wikimedia.org/w/index.php?title=Special:ListFiles/Sohmen&ilshowall=1`.

Table 5. Examples of semantically related categories used for the semantic evaluation

Commons Category	**Wikipedia Category**
Molecular biology	Molecular modeling
Temperature comparisons	Thermodynamics
Cochlear implants	Hearing
Robotics	Robots
Infectious disease control	Infectious diseases

license cc-by-2.5, cc-by-3.0 or cc-by-4.0. The categories were assigned by project members and also by other Wikipedia users. The image, the caption and the title of the paper were used to get information about the image and to select suitable categories. Only existing Wikimedia Commons categories were used. In total the images have received 264 categories.

Since the gold standard now is annotated with categories from Wikimedia Commons, while our method predicts categories from Wikipedia, the evaluation was not automated but done manually in order to allow for slight differences in the names of the categories like *soil* (Commons) and *soils* (Wikipedia) or *heart* and *heart (organ)*. The scope of literal consistency includes the singular and plural form of a category [18] and addition of scope notes.

Even when we allow for these small differences, we are too strict: if the gold standard has the category *robotics* and the algorithm proposes *robots*, this is of course not completely wrong. Thus, in addition on the *literal evaluation*, we also did a *semantic evaluation*, in which such broader semantically equivalent and related categories also were counted as matches. Further examples of pairs that were counted as equivalent are given in Table 5. This type of evaluation is similar to the semantic evaluation introduced in [17]. Of course, the results from this evaluation are subjective to a certain degree, but will nevertheless able make a division between useful and completely wrong categories instead of only considering literal identity.

For evaluation we assign always the five categories with the highest rank according to the used ranking variant. As measures for the quality of the results we use precision, recall and the harmonic mean of precision and recall (F1).

6 Results

The results of the evaluation are shown in the Table 6. Interestingly, the results for the semantic evaluation are very similar for all methods, while the variant using only idf is clearly inferior to the other variants for the literal evaluation.

7 Conclusion and future work

The overall result of precision and recall both around 0.4 does not seem to be very high, but is comparable to other systems using such a high number of possible

Table 6. Evaluation Results

Method	Literal			Semantic		
	Prec.	**Rec.**	**F1**	**Prec.**	**Rec.**	**F1**
Variant 1	0.036	0.015	0.021	0.42	0.36	0.39
Variant 2	0.054	0.059	0.057	0.40	0.40	0.40
Variant 3	0.053	0.058	0.055	0.42	0.40	0.41

categories. In fact we should compare the results rather to keyword extraction systems than to classification systems. We have applied the proposed method (variant 3) to all images in our database. Here, in total 66,873 different categories were used for 2,889,463 images (each receiving 5 categories). As usual a few categories (e.g. *Proteins* and *Gene Expression*, both over 50,000 times) were used extremely often, while most others were assigned only a few times.

Overall we can conclude that the proposed method gives useful results that, however still can and should be improved.

We see that using word embeddings is much more useful than using idf. It turns out that the cosine similarity between the aggregated semantic vector from key phrase words and caption words is a very effective method to filter out phrases that are found in Wikipedia but that are completely unrelated to the main topic of the caption text. Since many captions are very short or do not contain any words that are found in Wikipedia, we need to include sentences referring to the image as well to get enough candidates. This might, however also be a source that introduces less relevant words. The word embeddings help to filter these words out while keeping the useful ones.

One of the main problems for developing a method to assign good Wikipedia categories of images based on their captions is the absence of larger amounts of data for training and evaluation. Thus e.g. we could not learn optimal values for the weight constants in formula 3. Another problem is the quality of the manual assigned categories. People without much domain knowledge tend to assign categories that are mentioned literally in the caption of the image. Thus any method that does the same will be preferred. For future work we hope that more and more training data will become available if we upload images to Wikimedia Commons and the images get used on Wikipedia.

A further improvement can be expected when we use a better ways to find a representation of phrases and captions than just averaging the word vectors, as e.g. was done by Schltterer et al.[19].

Another direction that we want to explore is to find a more principled and data driven way to distinguish between categories that are useful to describe images (independent of any actual image) and categories that are not.

Acknowledgment

The presented work was developed within the NOA Project - Automatic Harvesting, Indexing and Provision of Open Access Figures from the Fields of Engineering and Technology Using the Infrastructure of Wikimedia Commons and Wikidata - funded by the DFG under grant number 315976924. NOA is a cooperative project of the Hochschule Hannover and the Technische Informationsbibliothek Hannover. We would like to thank the NOA project team.

References

1. J. Charbonnier, L. Sohmen, J. Rothman, B. Rohden, and C. Wartena: NOA: A Search Engine for Reusable Scientific Images Beyond the Life Sciences," in *Advances in Information Retrieval*, ser. Lecture Notes in Computer Science. Springer, Cham, Mar. 2018, pp. 797–800. [Online]. Available: `https://link.springer.com/chapter/10.1007/978-3-319-76941-7\_78`
2. R. Mihalcea and A. Csomai: "Wikify Linking Documents to Encyclopedic Knowledge," in *Proceedings of the Sixteenth ACM Conference on Conference on Information and Knowledge Management*, ser. CIKM '07. New York, NY, USA: ACM, 2007, pp. 233–242. [Online]. Available: `http://doi.acm.org/10.1145/1321440.1321475`
3. O. Medelyan, I. H. Witten, and D. N. Milne: Topic indexing with Wikipedia," vol. AAAI Technical Report WS-08-15, pp. 19–24, 2008. [Online]. Available: `http://researchcommons.waikato.ac.nz/handle/10289/1776`
4. C. Wartena and R. Brussee: Instanced-Based Mapping between Thesauri and Folksonomies," in *The Semantic Web - ISWC 2008*, ser. Lecture Notes in Computer Science, A. Sheth, S. Staab, M. Dean, M. Paolucci, D. Maynard, T. Finin, and K. Thirunarayan, Eds. Springer Berlin Heidelberg, Oct. 2008, pp. 356–370. [Online]. Available: `http://link.springer.com/chapter/10.1007/978-3-540-88564-1\_23`
5. E. Frank, G. W. Paynter, I. H. Witten, C. Gutwin, and C. G. Nevill-Manning: "Domain-Specific Keyphrase Extraction," in *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, ser. IJCAI '99. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1999, pp. 668–673. [Online]. Available: `http://dl.acm.org/citation.cfm?id=646307.687591`
6. P. D. Turney: "Learning Algorithms for Keyphrase Extraction," *Inf. Retr.*, vol. 2, no. 4, pp. 303–336, May 2000. [Online]. Available: `http://dx.doi.org/10.1023/A:1009976227802`
7. R. Mihalcea and P. Tarau, "Textrank: Bringing order into text," *Proc. 9th Conference on Empirical Methods in Natural Language Processing (EMNLP'04)*, 2004. [Online]. Available: `http://ci.nii.ac.jp/naid/20001460576/`
8. C. W. Leong, R. Mihalcea, and S. Hassan: "Text Mining for Automatic Image Tagging," in *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, ser. COLING '10. Stroudsburg, PA, USA: Association for Computational Linguistics, 2010, pp. 647–655. [Online]. Available: `http://dl.acm.org/citation.cfm?id=1944566.1944640`
9. G. Salton, A. Wong, and C. S. Yang: "A Vector Space Model for Automatic Indexing," *Commun. ACM*, vol. 18, no. 11, pp. 613–620, Nov. 1975. [Online]. Available: `http://doi.acm.org/10.1145/361219.361220`
10. C. Wartena, R. Brussee, and W. Slakhorst: "Keyword Extraction Using Word Co-occurrence," in *TIR 2010 - 7th International Workshop on Text-Based Information Retrieval, in Conjunction with DEXA 2010*, Oct. 2010, pp. 54–58.
11. C. Wartena and M. Sommer: "Automatic classification of scientific records using the German Subject Heading Authority File (SWD)," Oct. 2012. [Online]. Available: `https://serwiss.bib.hs-hannover.de/frontdoor/index/index/docId/328`

12. J. Voss, S. Bausch, J. Schmitt, J. Bogner, V. Berkelmann, F. Ludemann, O. Löffel, J. Kitroschat, M. Bartoshevska, and K. Seljuzki: "Normdaten in Wikidata," May 2014. [Online]. Available: `https://serwiss.bib.hs-hannover.de/frontdoor/index/index/docId/438`
13. Wikimedia Foundation: Wikipedia:Categorization, Jan. 2018, page Version ID: 821464874. [Online]. Available: `https://en.wikipedia.org/w/index.php?title=Wikipedia:Categorization\&oldid=821464874`
14. S. Bird, E. Klein, and E. Loper: *Natural Language Processing with Python*, 1st ed. Beijing ; Cambridge Mass.: O'Reilly and Associates, Jun. 2009.
15. English Penn Treebank tagset with modifications | Sketch Engine. [Online]. Available: `https://www.sketchengine.eu/english-treetagger-pipeline-2/`
16. J. Charbonnier and C. Wartena: "Using word embeddings for unsupervised acronym disambiguation," in *Proceedings of the 27th InternationalConference on Computational Linguistics*. Santa Fe: Association for Computational Linguistics, 2018, to appear.
17. L. Gazendam, C. Wartena, V. Malais, G. Schreiber, A. d. Jong, and H. Brugman: "Automatic Annotation Suggestions for Audiovisual Archives: Evaluation Aspects," *Interdisciplinary Science Reviews*, vol. 34, no. 2-3, pp. 172–188, Sep. 2009. [Online]. Available: `https://doi.org/10.1179/174327909X441090`
18. M. Iivonen: Consistency in the selection of search concepts and search terms," *Information Processing & Management*, vol. 31, no. 2, pp. 173–190, Mar. 1995. [Online]. Available: `http://linkinghub.elsevier.com/retrieve/pii/030645739580034Q`
19. J. Schlötterer, C. Seifert, and M. Granitzer: Supporting Web Surfers in Finding Related Material in Digital Library Repositories," in *Research and Advanced Technology for Digital Libraries*, ser. Lecture Notes in Computer Science. Springer, Cham, Sep. 2016, pp. 434–437. [Online]. Available: `https://link.springer.com/chapter/10.1007/978-3-319-43997-6\_38`