The Importance of being Grid
Chemnitz University of Technology at Grid@CLEF

Maximilian Eibl, Jens Kürsten,
Chemnitz University of Technology
Faculty of Computer Science, Media Computer Science
09107 Chemnitz, Germany
eibl@informatik.tu-chemnitz.de, jens.kuersten@informaik.tu-chemnitz.de

Abstract

This paper describes the participation of the Chemnitz University of Technology at Grid@CLEF 2009. We integrated the *CIRCO* framework into our *Xtrieval* framework and performed 15 runs in the three languages German, English, and French. For each language we used two different stemmers and two different retrieval models. One run one was a fusion run combining the results of the four other experiments. Whereas the different runs demonstrated that the impact of the used retrieval technologies is highly depending on the corpus, the merged approach produced the best results in each language.

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.3.4 - Systems and Software

General Terms

Measurement, Performance, Experimentation, Grid-Retrieval

Keywords

information retrieval, audiovisual media, data fusion, merging, Xtrieval, AMOPA

1 Introduction

In 2006 we started participating at CLEF. We hoped to gain in-depth insight into the effects of the different retrieval techniques in order to apply them to our real world problem: an archive for audiovisual media. Thus, we participated at several different tasks and achieved results from acceptable to very good. By now we can claim we got some kind of gut instinct how to configure our system in order to produce good results. But we did not get much closer to gain knowledge about the impact of retrieval techniques based on hard facts.

This contribution begins with explaining our motivation, the retrieval of audiovisual media in a TV-archive. It then provides a summary of the experiments done by the Chemnitz University of Technology at CLEF during the last four years. The final section discusses our results at the Grid-task and gives an outlook to future work.

2 Motivation: The Project sachsMedia

Saxony has a unique TV landscape in Europe. With 60 broadcasting stations 30% of the German local TV stations reside in Saxony. In particular, the district Chemnitz is very strong here. With 165 stations, altogether, the eastern German states are covered considerably dense. In the western German states 37 local TV stations reside.

The local TV stations are an important producer and deliverer of information which are covered by stations broadcasting nationwide. For example, the public broadcasting station "Mitteldeutscher Rundfunk" (MDR) covers three federal states - Saxony, Saxony-Anhalt and Thuringia - at once, and produces only half an hour broadcasting time per day for information from Saxony. According to several media studies, local television stations are the most important suppliers of local news and information – ahead of local radios, newspapers and local Internet offers.

In order to enable the local TV stations to cooperate the project *sachsMedia - Storage, Retrieval and Distribution of Audiovisual Media*[1] is creating tools for an archive of audiovisual media which can be jointly used by the TV stations. Within this archive both raw material as well as produced and broadcasted material is stored by every cooperating TV station. This material needs to be described as comprehensively as possible in order to be easily searchable. On the one hand, the description - or annotation - of the material is carried out intellectually according to principles of documentation. On the other hand automatic annotation will shift the load of annotation from human to the machine as far as possible. Here, sophisticated methods of multimedia retrieval will be implemented like object recognition and automated speech recognition.

It is commonplace in multimedia retrieval, that the actual search is not done using the original material but done using some kind of textual representation of the material. This textual representation can be produced intellectually: Ideally, some documentation specialist watches a video and describes it using a given vocabulary or classification which can be used for retrieval purposes afterwards. Reality looks a bit different: somebody describes a video by his/her own words. This is far from the ideal world but it is better than nothing. In the world of local TV storage and retrieval it looks like this: the producer of a video stores it on a common tape and puts the tape in a huge cupboard. Then he/she types some words describing the tape in a huge Excel-sheet. This annotation process is not designed for retrieval purposes but due to financial thoughts: proper intellectual annotation is expensive.

Another way to get textual annotations is an automatic analysis of the material. Within textual retrieval, this is not too complicated and every days practice. Just have a look at the web search engines. Nevertheless, intellectual annotation usually leads to better retrieval results even in the strictly textual domain. The major problem we face here is the switch between different media. Especially visual media are extremely hard to describe textually. For example, an article *about* the painting Mona Lisa will certainly contain the word "painting" and "Mona Lisa". They are easy to be extracted and therefore easy to be searched for successfully. But, the painting *itself* does not tell us anything about the person displayed. Thought we know the name of the painting, we only know because somebody told us – in words. Somebody annotated it for us. Here, the expensive intellectual annotation seems to be way superior to automated approaches.

In order to enable optimal retrieval and inexpensive but complete annotation, the gap between intellectual and automated annotation needs to be overcome. Our approach is holistic. We use every kind of annotation and combine them to a full description. Figure 1 demonstrates the approach:

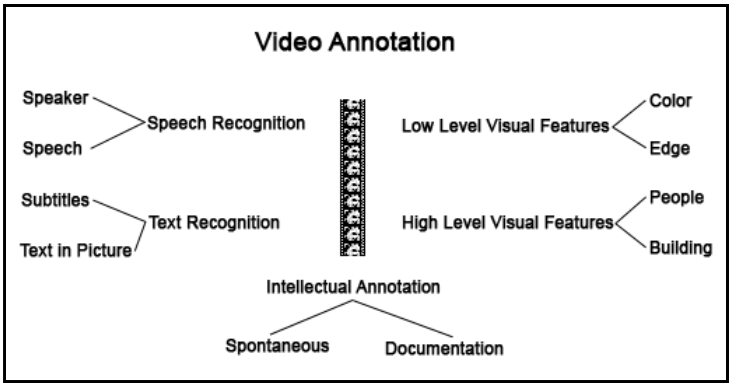

Fig 1: Video annotation approach used in sachsMedia

The video stream is taken and analyzed using text and speech recognition. This produces a large amount of textual information as it is also used in some CLEF tasks which are based on the outputs of automatic speech recognition (ASR) systems. On the other hand visual features of different levels are extracted. They range from low level features like color histograms to high level features like face detection. The third source of metadata is intellectual annotation. Here, user interfaces are created which support the user in providing annotations without being distracted from his original work – which is in the case of local TV-station producing videos.

For some well developed elements of the annotation process we rely on commercial and open source products. Software for speech recognition for example is well developed and there is neither need nor capacity to

[1] The project sachsMedia is funded by the *Unternehmen Region*-program of the *German Federal Ministry of Education and Research*. For more information about the project see: http://www.tu-chemnitz.de/informatik/Medieninformatik/Sachsmedia/

build up an own system. Others like the low-level visual features can easily be implemented in our system. Yet others like speaker recognition are well enough described in literature to guarantee a useful implementation. Thus, for the single annotation tasks we rely on previous work, open source and commercialized products as far as possible. Nevertheless, there is much research left to do. For high-level features like recognition of people and places suitable processes are still needed. Text recognition in video streams needs to be implemented. And the proper way of intellectual annotations needs to be defined.

The final task is to configure the interplay of the single annotations. Speaker recognition by audio for example can support the people recognition by video and vice versa. Buildings found by picture recognition can be named by intellectual annotation or text recognition. The ways of metadata becoming interwoven are manifold. In order to have a closer look at the impact of the different metadata to the retrieval process and the dependencies and interactions between these metadata we are developing a highly flexible retrieval framework which will described in the following.

3 The Xtrieval and AMOPA Frameworks

In 2005 we started conceptualizing and implementing a flexible framework for information retrieval purposes. It is a general finding in information retrieval, that the performance of retrieval systems highly depends on hardly generalizable aspects like for example corpora: Retrieval methods that perform well with one corpus do not necessarily work at all when applied to another corpus. After all that is the reason for installing different tracks in evaluation campaigns like CLEF[2] and TREC[3].

The general idea was to create a framework which is highly flexible and adjustable concerning information retrieval technologies. The framework needed to provide interfaces to combine different state-of-the-art text retrieval techniques on the one hand and to evaluate and integrate new methods for multimedia retrieval on the other hand. An in-depth description of the framework design is given in [1].

The framework, named *Xtrieval*, implements a Java-based object-orientated API specification providing interfaces to all methods necessary for all possible designs of retrieval systems. By this, the framework is able to exchange, evaluate, and combine different components of other retrieval systems. In a first implementation *Apache Lucene*[4] was integrated but by now also *Terrier*[5] and *Lemur*[6] are included in practice. The framework supports not only the integration of these and other toolkits but also allows combining their retrieval results on the fly.

Thus, the framework provides a realm of possible configurations. In order to conveniently adjust the system to different corpora we created a Graphical User Interface (GUI) (see figure 2). This GUI provides a general configuration interface that supports the user in setting all parameter driven classes. Thus, all parameters of each class can be changed during runtime without any changes in the source code of the project. A second interface incorporates methods for calculating and visualizating recall-precision graphs. Additional functions to load and save relevance assessments in popular formats (e.g. TREC) are provided as well.

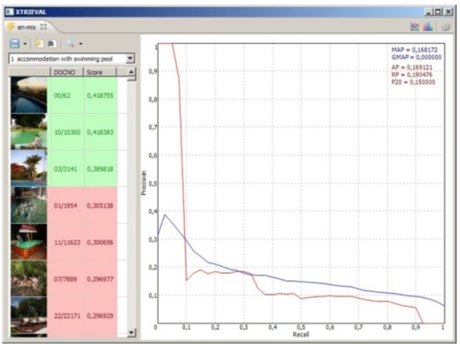

Fig 2: GUI for Evaluation

The GUI can be used to configure the three main components: indexing, retrieval and evaluation (see figure 3). A general programming interface is able to convert every structured data collection into an internal representation which is then used for the application of transformation and tokenization procedures like for example different stemming algorithms. The pre-processed data is than passed forward to a programming interface which allows connecting indexing libraries like *Lucene*. In order to integrate the full amount of metadata of audiovisual data we created the framework *AMOPA* which is presented later on.

Probably the most important interface of the *Xtrieval* framework allows the flexible use of retrieval algorithms. Queries are pre-processed according to the needs of different toolkits. It is also possible to combine searches in different indexes and to fuse these results into one result set by for example Sum-RSV, Product-RSV, and Z-Score.

---

[2] http://www.clef-campaign.org/
[3] http://trec.nist.gov
[4] http://lucene.apache.org/
[5] http://ir.dcs.gla.ac.uk/terrier/
[6] http://www.lemurproject.org/

Finally the evaluation component is capable to store and reload experiments and their complete parameter sets. This enables us to repeat experiments at a later date. It provides several measures to compare retrieval output to assessments. Additionally, it is possible to load and store relevance assessments in the TREC format. Figure 3 demonstrates the basic architecture of *Xtrieval*:

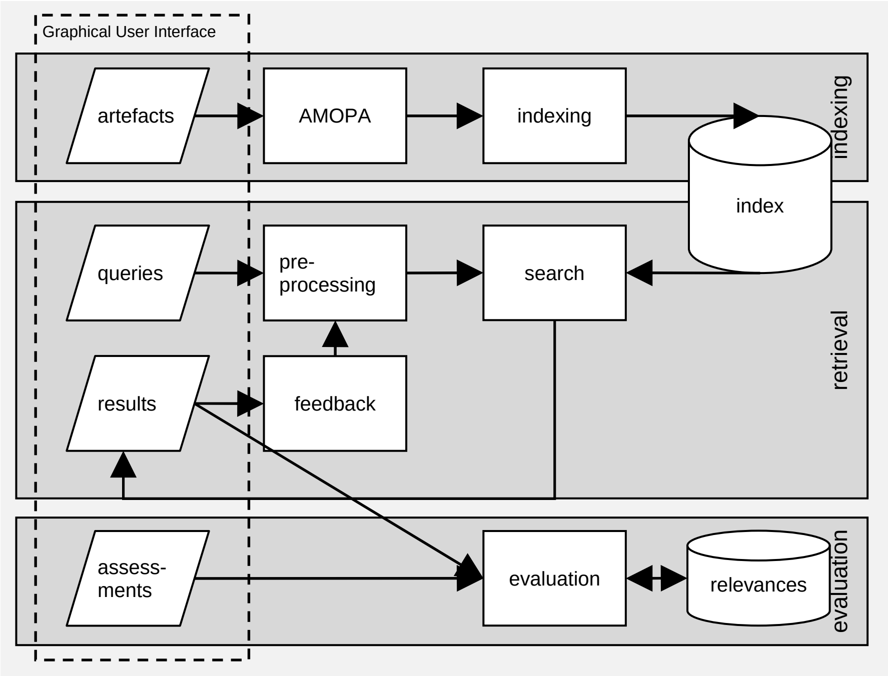

Fig 3: The Xtrieval framework

For practical reasons (video analysis tool are written in C, *Xtrieval* in Java) we built for the automated annotation tasks a separate framework called *AMOPA-Automated MOving Picture Annotator*. *AMOPA* uses the *FFMPEG*[7] library to read video stream and perform first low level methods. Access for Java code to the C library *FFMPEG* is provided by the library *FFMPEG-Java*, which is part of the *Streambaby*[8] project. The actual analysis is performed by *AMOPA* and organized in process chains. This concept allows us to exchange and reorder processes very easily. A detailed description of AMOPA is given in [2]. Figure 4 demonstrates the basic concept:

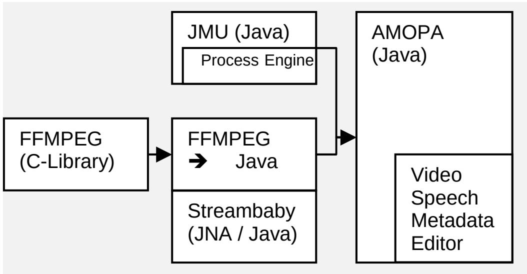

Fig 4: The AMOPA framework [2]

4 Lessons learned

In 2009 we participated the 4th time at CLEF. Table 1 gives a summary of our experiences with different CLEF tasks and provides short insight into the experiences of other participating groups. As one can see our system performed quite different over the years. Performance seems to be highly depending on the underlying corpus.

[7] http://ffmpeg.org/
[8] http://code.google.com/p/streambaby/

Table 1: Our experiences with CLEF

Task	Our configuration and findings	Other participant's experiences	Ref.
2006 Domain Specific	We used Apache Lucene and implemented a combination of suffix stripping, stemming, decompounding, and blind feedback. This rather simple approach performed very well.	Considering that it was our first time at CLEF, our system performed astonishingly well. In some tasks it even outperformed all other systems.	[3] [4]
2006 ImageCLEF	Same approach as above. The search focused on the textual representation. A simple color histogram was the only content based element. The system performed well. The color histogram improved ranking slightly.	Most groups relied on text-only runs, only some 30% implemented some kind of visual technique, usually as simple as ours. One group implemented manual feedback and by this could improve performance significantly.	[5] [6]
2007 Domain Specific	We used a structured index, an unstructured index, and a combination of them both combined by z-score data fusion. Interestingly, the unstructured index outperformed the structured one and the data merging approach. The integration of a thesaurus-based query expansion did not improve performance (the thesaurus was taken from Open Office). Google Translate and PROMT were used for translation because they performed best in some preliminary runs.	Compared to other groups our system did not perform as well as in the previous year. The by far best results came from Xerox. The group used lexical entailment (i.e. the probability that one term entails another) to provide query expansion. Similar terms from corpus documents in relation to query terms were defined by a language modeling approach.	[7] [8] [9]
2007 ImageCLEF	Same approach as in the Domain Specific Track. The best MAP increased from .2436 in 2006 to .3175 in 2007.	Interestingly the system improved compared to 2006. It even produced the best MAP of all participants. Here again Xerox performed very well (see above).	[10] [11]
2008 Ad-Hoc	We used different stemming approaches for German and English and combined the results in the retrieval stage by applying our implementation of the z-score operator. We also used a standard top-k pseudo-relevance feedback algorithm in the retrieval stage. The performance of our monolingual experiments was slightly below the average for the German and French collection and in the top 5 for the English collection. Our bilingual experiments performed very well (at least in the top 3) for all target collections. Here again, merging different approaches improved the results significantly. Due to a mistake in the setting we had to reconfigure the system and to produce the final runs within only a few hours. At least this proves the flexibility of the framework.	The results of the different groups are quite interesting: • Monolingual EN: unine - .3754 MAP (next best MAP: .3623) • Monolingual FR: unine - .3327 MAP (next best MAP: .3088) • Monolingual DE: opentext - .3571 MAP (next best MAP: .3377) • Bilingual EN: chemnitz - .3415 MAP (next best MAP: .2824) • Bilingual FR: chesire - .1884 MAP (next best MAP: .1754) • Bilingual DE: jhu-apl - .1898 MAP (next best MAP: .1851) It is rather noticeable that nearly each task was dominated by another group. One could think the reason for that are different language-specific technologies. But then, should not be at least the same languages be dominated by the same groups? Striking is the bad performance of all groups in the bilingual French and German tasks.	[12] [13]
2008 Domain Specific	The Xtrieval-framework was implemented using combinations of the Porter and the Krovetz stemmers for English and the Snowball stemmer and an N-Gram based decompounding approach for German. For multilingual retrieval we made use of the Google AJAX language API. In addition to pure translation, a combination of automatic translation and language mappings as provided by the bilingual translation tables was employed. We used a blind feedback approach that was combined for some runs with query expansion based on thesaurus terms. Interestingly it	Not all groups that participated at Ad-hoc also participated at DS. Still the performance of the six participating groups was far more homogeneous at DS: • Monolingual EN: chemnitz - .3891 MAP (next best MAP: .3770) • Monolingual DE: unine - .4537 MAP (next best MAP: .4367) • Monolingual RU: unine - .1815 MAP (next best MAP: .1400) • Bilingual EN: chemnitz - .2285 MAP (next best MAP: .2285) • Bilingual DE: chemnitz - .3702 MAP (next best MAP: .2231) • Bilingual RU: chemnitz - .0882 MAP (next best MAP: .0857)	[14] [15] [16]

	was found that such use of the controlled vocabulary did not benefit the retrieval effectiveness.		
2008 VideoCLEF	The Xtrieval framework was adapted for the classification of the Video ASR data. We regarded the task as a text classification problem. Terms from Wikipedia categories served as training data for text classifiers. For the text classification the Naive-Bayes and kNN classifiers from the WEKA toolkit were used. The translation of the feeds to English (translation task) was done using Google's AJAX language API. The evaluation of the classification task showed worse results than our preliminary tests mad us expect with a precision between 10 and 15 percent. Interestingly, we could not improve the quality of the classification by using the provided metadata. Translating the RSS Feeds performed well.	The task was performed the first time at CLEF 2008. Five groups participated. Most of the groups took the task as a text classification problem and implemented different classifiers (SVM, Naive Bayes, k-NN) – as we did as well. Interestingly the group performing best approached the task as an information retrieval problem. In 2009 we used this approach as well and performed way better than in 2008.	[17] [18]
2008 ImageCLEF photo	Our thesaurus based query expansions works well in improving GMAP and BPREF, but deteriorates the improvements gained by the addition of content-based image retrieval. The baseline (text-only) scored a MAP of .0998. The combination of text and content based image retrieval improved the MAP by 37 percent to .1364. After applying the query expansion to both experiments the MAP for the text-only retrieval increased to 0.1081, but the MAP for the combined text and content based retrieval decreased to .1140.	Interestingly our system changed again for the worse. Several other systems (especially DCU) performed way better.	[19] [20]
2008 wiki-pediaMM	We used the same setup as for the ImageCLEF-task: Apache Lucene, a customized analyzer with positional stopword removal and Snowball stemmer. For the content-based image retrieval we used Caliph & Emir calculating additional MPEG-7 descriptors (scalable color, edge histogram, color layout and dominant color descriptor). Our text-only baseline reached a MAP of .2166. By adding the provided content-based image features and the four MPEG-7 descriptors, the MAP decreased to .2138. After the preprocessing of the topics with query expansion our highest MAP of .2195 was achieved. The inclusion of concepts scored the worst MAP of .2048, but retrieved 23 more relevant documents than any other of our experiments.	The University of Alicante produced very good results using a classical text retrieval engine and a procedure to decompose compound file names in camel case notation into single terms and a module that performs geographical query expansion. The Peking University performed best with a MAP of .4333 for their text only run which was based on QE where the expansion terms are selected from a knowledge base that is (semi-) automatically constructed from Wikipedia. Interestingly other approaches like the integration of content based features worsened the results dramatically. Taking only MAP into account fusing text and concepts in general produced the best performing runs. Merging text and content based approaches results in almost identical MAP as the text-only baseline.	[21] [22]
2008 QAST	Here we participated at manual speech transcription tasks. We used the Stanford Named Entity Recognizer for tagging named entities and the CRFTagger - Conditional Random Fields Part-of-Speech (POS) Tagger for English. The passage retrieval was done with the Xtrieval framework. For the classification of the question hand-crafted patterns were implemented. Compared to other participants the system performed worst for factual questions and best for definitional questions.	In both tasks (T1a, T4a) LIMSI (Laboratoire d'Informatique et de Mécanique des Sciences de l'Ingénieur) dominated the factual questions. They used a sophisticated set of enrichment tools ranging from morphological analysis to synonyms and extended NEs as well as a specific index for known acronyms.	[23] [24]

5 Grid retrieval in 2009

The *Xtrieval* framework was used to prepare and run our text retrieval experiments for the *Grid Experiments Pilot Track*. The core retrieval functionality is provided by *Apache Lucene*, the *Lemur* toolkit, and the *Terrier* framework. This allowed us to choose from a wide range of state of the art retrieval models for all kinds of text retrieval experiments. Our main goal in this first Grid experiment was to provide strong baseline experiments, which could be used as reference for evaluation of sophisticated new retrieval approaches.

In order to participate at the *Grid@CLEF* track the *CIRCO* framework [25] had to be integrated into *Xtrieval*. Since one of the main design concepts of the *Xtrieval* framework was flexibility towards enhancements only a small number of classes had to be rewritten: two classes that are used to process the token streams during indexing and another class that writes the processed token stream in the index format of the used retrieval core. Since the integration of the *Lemur* and *Terrier* retrieval toolkits into *Xtrieval* had been done lately we did not have the time to test and debug the integration. Thus, we decided to adapt the *Lucene* indexing class only.

Ten collections in five European languages, namely Dutch, English, French, German and Italian were provided for the *Grid Experiment Pilot Track*. For our participation we chose to run experiments on the English, French and German collections, which included six text collections in total. Table 2 shows the used collections and the provided fields which were taken for indexing. Table 3 shows some indexing statistics.

Table 2: Fields Used for Indexing

Collection	*Indexed Fields*
DE: Spiegel 1994/5	LEAD, TEXT, TITLE
DE: Frankfurter Rundschau 1994	TEXT, TITLE
DE: German SDA 1994	KW, LD, NO, ST, TB, TI, TX
EN: LA Times 1994	BYLINE, HEADLINE, TEXT
FR: Le Monde 1994	CHA1, LEAD1, PEOPLE, SUBJECTS, TEXT, TIO1
FR: French SDA 194	KW, LD, NO, ST, TB, TI, TX

Table 3: Index Statistics and CIRCO Output per Language

Lang	Stemmer	# Docs	# Terms	# Distinct Terms	Avg. Doc Length	# Chunk Files	Compressed File Size (MB)
DE	Snowball	225,371	28,711,385	3,365,446	127.40	225	15,695
DE	N-gram Decomp.	225,371	63,118,598	840,410	280.07	225	19,924
EN	Snowball	113,005	20,210,424	685,141	178.85	114	14,293
EN	Krovetz	113,005	20,701,670	704,424	183.19	114	14,293
FR	Snowball	87,191	12,938,610	1,130,517	148.39	88	7,329
FR	Savoy [27]	87,191	13,262,848	1,239,705	152.11	88	7,323

We performed 15 runs, five for each language German, English, and French. For each language we used two different stemmers and two different retrieval models. One run one was a fusion run combining the results of the four other experiments. Table 4 provides the general configuration of each experiment as well as the retrieval performance in terms of mean average precision (MAP) and geometric mean average precision (GMAP). Please note the French run *cut_fr_3*. This run was corrupted while submitting. We did a separate evaluation for this run: *cut_fr_3** is not part of the official statistics but shows the correct results.

All in all, merging models and stemmers brings the best results for all three languages. Comparing the models and stemmers leads to the following conclusions:

- *German*: BM25 performs better than VSM. N-gram performs better than Snowball.
- *English*: The results in English are vice versa: VSM performs (slightly) better than BM25. Snowball performs (slightly) better than Krovetz.
- *French*: Here the results are even more confusing: VSM performs (especially in conjunction with Savoy) better than BM25. In conjunction with VSM Snowball performs better but in conjunction with BM25 Savoy is superior.

Table 4: Results Overview

Lang	ID	Core	Model	Stemmer	# QE docs / tokens	MAP	GMAP
DE	cut_de_1	Lucene	VSM	Snowball	10 / 50	.4196	.2023
DE	cut_de_2	Terrier	BM25	Snowball	10 / 50	.4355	.2191
DE	cut_de_3	Lucene	VSM	N-gram	10 / 250	.4267	.2384
DE	cut_de_4	Terrier	BM25	N-gram	10 / 250	.4678	.2682
DE	**cut_en_5**	**both**	**both**	**both**	**10 / 50 & 250**	**.4864**	**.3235**
EN	cut_en_1	Lucene	VSM	Snowball	10 / 20	.5067	.3952
EN	cut_en_2	Terrier	BM25	Snowball	10 / 20	.4926	.3314
EN	cut_en_3	Lucene	VSM	Krovetz	10 / 20	.4937	.3762
EN	cut_en_4	Terrier	BM25	Krovetz	10 / 20	.4859	.3325
EN	**cut_en_5**	**both**	**both**	**both**	**10 / 20**	**.5446**	**.4153**
FR	cut_fr_3	Lucene	VSM	Snowball	10 / 20	.0025	.0000
FR	*cut_fr_3**	*Lucene*	*VSM*	*Snowball*	*10 / 20*	*.4483*	*.3060*
FR	cut_fr_1	Terrier	BM25	Snowball	10 / 20	.4538	.3141
FR	cut_fr_5	Lucene	VSM	Savoy [27]	10 / 20	.4434	.2894
FR	cut_fr_2	Terrier	BM25	Savoy [27]	10 / 20	.4795	.3382
FR	**cut_fr_4**	**both**	**both**	**both**	**10 / 20**	**.4942**	**.3673**

Thus, some results demonstrate a better performance for VSM, some results show superiority of BM25. The results for the stemmers are similarly unpredictable. But it seems that this uncertainty can be overcome by data fusion: As table 4 demonstrates, for each language the merging of the retrieval models produced the best results. In our framework, merging is done by the z-score operator [28]. The results for the merged experiments are shown in figure 5:

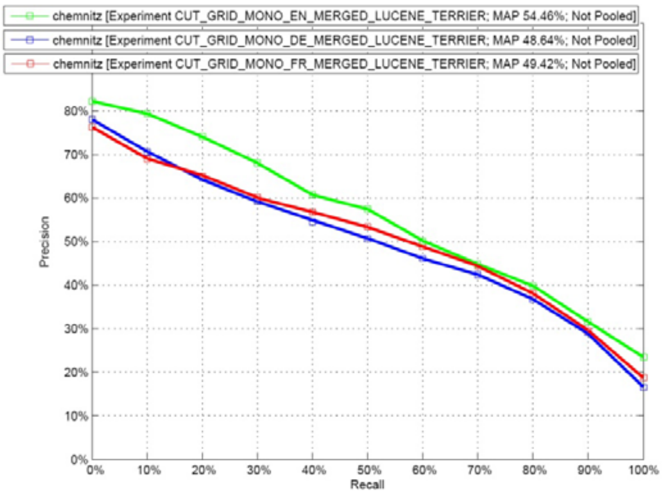

Fig 5: Results for the merged experiments

These results confirm our findings described in section 2: the impact of retrieval techniques are highly depending on the corpus and quite unpredictable. All in all, while participating at CLEF we developed a decent gut instinct in configuring the retrieval system to produce good and very good retrieval results. But in fact the configuring task is still at bit like stumbling in the dark. The exact effects of retrieval mechanisms remain enigmatic. We still do not have strong rules which let us predict the retrieval quality. And so we never know whether or not there is a better configuration we did not predict. Having such rules would enable us to automatically configure a retrieval engine in accordance to the corpus.

The basic idea of Grid@CLEF is to compare the outcomes of different retrieval approaches. For this intermediate data from the indexing process of each participating system is made available to other participants. This year, next to us only the Cheshire group took part. Except for the German runs the differences between the systems are rather small. The next step is to feed our system with the Cheshire intermediate data in order to figure out the effects of the different approaches.

It is our belief, that far more experiments are needed before we get even close to such rules. The Grid@CLEF track is exactly the platform the community needs to answer this question. We will certainly take part again next year.

Acknowledgements

We would like to thank Jaques Savoy and his co-workers for providing numerous resources for language processing. Also, we would like to thank Giorgio M. di Nunzio and Nicola Ferro for developing and operating the DIRECT system.

This work was partially accomplished in conjunction with the project *sachsMedia,* which is funded by the *Unternehmen Region*-program[9] of the *German Federal Ministry of Education and Research*[10].

References

All Web references were retrieved on August 24th, 2009.

[1] Thomas Wilhelm. Entwurf und Implementierung eines Frameworks zur Analyse und Evaluation von Verfahren im Information Retrieval. Chemnitz University of Technology, Faculty of Computer Science, Diploma Thesis. http://archiv.tu-chemnitz.de/pub/2008/0117/data/da.pdf

[2] Marc Ritter (2009). Visualisierung von Prozessketten zur Shot Detection. In: Eibl, Maximilian; Kürsten, Jens; Ritter, Marc: Workshop Audiovisuelle Medien WAM 2009: Archivierung, pp.135-150. http://archiv.tu-chemnitz.de/pub/2009/0095/data/wam09_monarch.pdf

[3] Maximilian Stempfhuber and Stefan Baerisch (2006). Domain-Specific Track CLEF 2006: Overview of Results and Approaches, Remarks on the Assessment Analysis. In: A. Nardi, C. Peters and J.L. Vicedo (Ed.) Working Notes for the CLEF 2006 Workshop, 21-23 September, Alicante, Spain. http://www.clef-campaign.org/2006/working_notes/workingnotes2006/stempfhuberOCLEF2006.pdf

[4] Jens Kürsten and Maximilian Eibl (2006). Monolingual Retrieval Experiments with a Domain-Specic Document Corpus at the Chemnitz Technical University. In: A. Nardi, C. Peters and J.L. Vicedo (Ed.) Working Notes for the CLEF 2006 Workshop, 21-23 September, Alicante, Spain. http://www.clef-campaign.org/2006/working_notes/workingnotes2006/kuerstenCLEF2006.pdf

[5] Paul Clough, Michael Grubinger, Thomas Deselaers, Allan Hanbury, and Henning Müller (2006). Overview of the ImageCLEF 2006 photographic retrieval and object annotation tasks. In In: A. Nardi, C. Peters and J.L. Vicedo (Ed.) Working Notes for the CLEF 2006 Workshop, 21-23 September, Alicante, Spain. http://www.clef-campaign.org/2006/working_notes/workingnotes2006/cloughOCLEF 2006.pdf

[6] Thomas Wilhelm and Maximilian Eibl (2006). ImageCLEF 2006 Experiments at the Chemnitz Technical University. In: A. Nardi, C. Peters and J.L. Vicedo (Ed.) Working Notes for the CLEF 2006 Workshop, 21-23 September, Alicante, Spain. http://www.clef-campaign.org/2006/working_notes/ workingnotes2006/wilhelmCLEF2006l.pdf

[7] Vivien Petras, Stefan Bärisch, and Maximillian Stempfhuber (2007). The Domain-Specific Track at CLEF 2007. In: A. Nardi and C. Peters (Ed.) Working Notes for the CLEF 2007 Workshop, 21-23 September, Budapest, Hungary. http://www.clef-campaign.org/2007/working_notes/ petras_overview-ds-CLEF2007.pdf

[8] Jens Kürsten and Maximilian Eibl (2007). Domain-Spefic Cross Language Retrieval: Comparing and Merging Structured and Unstructured Indices. In: A. Nardi and C. Peters (Ed.) Working Notes for the CLEF 2007 Workshop, 21-23 September, Budapest, Hungary. http://www.clef-campaign.org/2007/ working_notes/kuerstenCLEF2007.pdf

[9] Giorgio Maria Di Nunzio and Nicola Ferro (2007). Appendix C: Results of the Domain Specific Track: In: A. Nardi and C. Peters (Ed.) Working Notes for the CLEF 2007 Workshop, 21-23 September, Budapest, Hungary. http://www.clef-campaign.org/2007/working_notes/AppendixC.pdf

[10] Michael Grubinger, Paul Clough, Allan Hanbury, and Henning Müller (2007). Overview of the ImageCLEFphoto 2007 photographic retrieval task. In: A. Nardi and C. Peters (Ed.) Working Notes for the CLEF 2007 Workshop, 21-23 September, Budapest, Hungary. http://www.clef-campaign.org/2007/working_notes/grubingerCLEF2007_Overview.pdf

[11] Thomas Wilhelm, Jens Kürsten, and Maximilian Eibl (2007). Experiments for the ImageCLEF 2007 Photographic Retrieval Task. In: A. Nardi and C. Peters (Ed.) Working Notes for the CLEF 2007 Workshop, 21-23 September, Budapest, Hungary. http://www.clef-campaign.org/2007/working_notes/ wilhelmCLEF2007.pdf

---

[9] http://www.unternehmen-region.de

[10] http://www.bmbf.de/en/

[12] Eneko Agirre, Giorgio M. Di Nunzio, Nicola Ferro, Thomas Mandl, and Carol Peters (2008). CLEF 2008: Ad Hoc Track Overview. In: Working Notes for the CLEF 2008 Workshop, 17-19 September, Aarhus, Denmark. http://www.clef-campaign.org/2008/working_notes/adhoc-final.pdf

[13] Jens Kürsten, Thomas Wilhelm, and Maximilian Eibl (2008). CLEF 2008 Ad-Hoc Track: On-line Processing Experiments with Xtrieval. In: Working Notes for the CLEF 2008 Workshop, 17-19 September, Aarhus, Denmark. http://www.clef-campaign.org/2008/working_notes/kuersten-ah-paperCLEF2008.pdf

[14] Vivien Petras, Stefan Baerisch (2008). The Domain-Specific Track at CLEF 2008. In: Working Notes for the CLEF 2008 Workshop, 17-19 September, Aarhus, Denmark. http://www.clef-campaign.org/ 2008/working_notes/overview-ds-clef-2008.pdf

[15] Giorgio Maria Di Nunzio and Nicola Ferro (2008). Appendix D - Results of the Domain Specific Track. In: Working Notes for the CLEF 2008 Workshop, 17-19 September, Aarhus, Denmark. http://www.clef-campaign.org/2008/working_notes/AppendixD.pdf

[16] Jens Kürsten, Thomas Wilhelm, and Maximilian Eibl. (2008) The Xtrieval Framework at CLEF 2008: Domain-Specic Track. In: Working Notes for the CLEF 2008 Workshop, 17-19 September, Aarhus, Denmark. http://www.clef-campaign.org/2008/ working_notes/kuersten-ds-paperCLEF2008.pdf

[17] Martha Larson, Eamonn Newman, and Gareth Jones (2008). Overview of VideoCLEF 2008: Automatic Generation of Topic-based Feeds for Dual Language Audio-Visual Content. In: Working Notes for the CLEF 2008 Workshop, 17-19 September, Aarhus, Denmark. http://www.clef-campaign.org/2008/ working_notes/Larson_overviewCLEF_VideoCLEF.pdf

[18] Jens Kürsten, Daniel Richter, and Maximilian Eibl (2008). VideoCLEF 2008: ASR Classification based on Wikipedia Categories. In: Working Notes for the CLEF 2008 Workshop, 17-19 September, Aarhus, Denmark. http://www.clef-campaign.org/2008/ working_notes/kuersten-videoclef-paperCLEF2008.pdf

[19] Thomas Arni, Paul Clough, Mark Sanderson, and Michael Grubinger (2008). Overview of the ImageCLEFphoto 2008 Photographic Retrieval Task. In: Working Notes for the CLEF 2008 Workshop, 17-19 September, Aarhus, Denmark. http://www.clef-campaign.org/2008/working_notes/ ImageCLEFphoto2008-final.pdf

[20] Thomas Wilhelm, Jens Kürsten, and Maximilian Eibl (2008). The Xtrieval Framework at CLEF 2008: ImageCLEF photographic retrieval task. In: Working Notes for the CLEF 2008 Workshop, 17-19 September, Aarhus, Denmark. http://www.clef-campaign.org/2008/working_notes/wilhelm-imageclefphoto-paperCLEF2008.pdf

[21] Theodora Tsikrika and Jana Kludas (2008). Overview of the wikipediaMM task at ImageCLEF 2008. In: Working Notes for the CLEF 2008 Workshop, 17-19 September, Aarhus, Denmark. http://www.clef-campaign.org/2008/working_notes/wikiMMoverview-paperCLEF2008.pdf

[22] Thomas Wilhelm, Jens Kürsten, and Maximilian Eibl (2008). The Xtrieval Framework at CLEF 2008: ImageCLEF Wikipedia MM task. In: Working Notes for the CLEF 2008 Workshop, 17-19 September, Aarhus, Denmark. http://www.clef-campaign.org/2008/working_notes/wilhelm-imageclefwiki-paperCLEF2008.pdf

[23] Jordi Turmo, Pere Comas, Sophie Rosset, Lori Lamel, Nicolas Moreau, and Djamel Mostefa (2008). Overview of QAST 2008. In: Working Notes for the CLEF 2008 Workshop, 17-19 September, Aarhus, Denmark. http://www.clef-campaign.org/2008/working_notes/turmo_QAST2008-overview.pdf

[24] Jens Kürsten, Holger Kundisch, and Maximilian Eibl (2008). QA Extension for Xtrieval: Contribution to the QAst track. In: Working Notes for the CLEF 2008 Workshop, 17-19 September, Aarhus, Denmark. http://www.clef-campaign.org/2008/working_notes/ kuersten-qast-paperCLEF2008.pdf

[25] Iadh Ounis, Christina Lioma, Craig Macdonald, and Vassilis Plachouras (2007). Research directions in terrier: a search engine for advanced retrieval on the web. Novatica/UPGRADE Special Issue on Next Generation Web Search, pages 49-56, 2007.

[26] Nicola Ferro and Donna Harman (2008). Dealing with multilingual information access: Grid experiments at trebleclef. Post-proceedings of the Fourth Italian Research Conference on Digital Library Systems (IRCDL 2008), pages 29-32, 2008.

[27] Jaques Savoy (1999). A Stemming Procedure and Stopword List for General French Corpora. In: Journal of the American Society for Information Science, 1999, pp. 944-982.

[28] Jaques Savoy (2004). Data Fusion for Effective European Monolingual Information Retrieval. In: Working Notes for the CLEF 2004 Workshop, 15-17 September, Bath, UK. http://www.clef-campaign.org/2004/working_notes/ WorkingNotes2004/22.pdf