DOKTORI (Ph.D.) ÉRTEKEZÉS

NEW METHODS IN WEB INFORMATION RETRIEVAL EFFECTIVENESS

ÚJ MÓDSZEREK A WEB-ES INFORMÁCIÓ-VISSZAKERESÉS HATÉKONYSÁGÁNAK TERÜLETÉN

Skrop Adrienn

Témavezető: Dr. Dominich Sándor

Pannon Egyetem
Műszaki Informatikai Kar
Informatikai Tudományok Doktori Iskola

2006

ÚJ MÓDSZEREK A WEB–ES INFORMÁCIÓ-VISSZAKERESÉS HATÉKONYSÁGÁNAK TERÜLETÉN

Értekezés doktori (PhD) fokozat elnyerése érdekében

a Pannon Egyetem Informatikai Tudományok Doktori Iskolájához tartozóan

Írta:

Skrop Adrienn

Témavezető: Dr. Dominich Sándor

Elfogadásra javaslom (igen / nem)

.....

Dr. Dominich Sándor

A jelölt a doktori szigorlaton …..... % -ot ért el.

Az értekezést bírálóként elfogadásra javaslom:

Bíráló neve: …..... …..... igen / nem

…….....

(aláírás)

Bíráló neve: …..... …..... igen / nem

…….....

(aláírás)

A jelölt az értekezés nyilvános vitáján …..... % - ot ért el.

Veszprém, …….

a Bíráló Bizottság elnöke

A doktori (PhD) oklevél minősítése …..

……...……

Az EDT elnöke

TARTALMI KIVONAT

Az információ–visszakeresés egyik fontos területe az információ-visszakereső módszerek relevanciahatékonyságának a mérése. A relevanciahatékonyság azt jelenti, hogy az információ–visszakereső módszer képes releváns választ adni a felhasználó információigényére. A relevanciahatékonyságot laboratóriumi körülmények között a Cranfield paradigma alapján mérik. A kiértékelés standard tesztkollekciókon a teljesség és pontosság standard mértékek alkalmazásával végezhető el.

A Web–es információ–visszakeresés relevanciahatékonyságának mérésére nem alkalmas a laboratóriumi Cranfield féle mérés, mert a mértékek nem számíthatók ki. Ezért a Web–es információ–visszakeresés relevanciahatékonyságának mérésére új mértékeket kell létrehozni.

A legújabb kutatások azt mutatják, hogy a Web–es keresésnek három válfaja van: navigációs, tájékozódási és tranzakciós. Az egyik legfontosabb navigációs feladat a honlapkeresés. Honlapkeresés során a felhasználó célja egy adott entitás (cég, intézmény, személy stb.) honlapjának megtalálása Web–es keresőmotor segítségével.

Szerző a honlapkeresési hatékonyságot a felhasználók szempontjából vizsgálja, ennek mérésére két új mértéket adott meg: a Pszeudo–pontosságot és az Átlag Pszeudo–rang mértékeket. A Pszeudo–pontosság és az Átlag Pszeudo–rang mértékeket felhasználva, Szerző megadta a MICQ eljárást keresőkérdések honlapazonosító képességének a mérésére.

ABSTRACT

In Information Retrieval (IR) the evaluation of IR systems plays an essential role. The most important type of evaluation of IR systems is retrieval effectiveness evaluation. Retrieval effectiveness evaluation measures how well a given system or algorithm can match, retrieve and rank documents that are relevant to the user's information need. Laboratory testing of IR algorithms is based on the Cranfield paradigm. The Cranfield paradigm uses a test collection and retrieval effectiveness is measured with the standard measures Precision and Recall.

Information retrieval on the Web is different from retrieval in traditional document collections. Thus, the Cranfield type evaluation of Web IR systems is usually not possible: the standard measures cannot be calculated. New or revised methodology and evaluation measures are required. Two new measures called Pseudo Precision and Mean Pseudo Rank are proposed in the dissertation. The measures are based on the Mathematical Reliability Theory and they measure the home page identification capability on the Web. Based on Pseudo Precision and Mean Pseudo Rank the dissertation introduces the MICQ method to measure the home page identification capability of search queries on the Web.

ABSTRAKT

Ein wichtiges Gebiet in Informationswiedergewinnung (*information retrieval*) ist die Messung von der Relevanzwirksamkeit der verwendeten Methoden. Relevanzwirksamkeit bedeutet, daß mit der wiedergewinnenden Methode auf den Informationsbedarf (*query*) der Benutzer relevante Antwort gegeben werden kann. Die Relevanzwirksamkeit wird unter Laborzuständen aufgrund des Cranfield Paradigmas gemessen. In den Standard Testkollektionen kann die Auswertung mit der Anwendung von den Recall und Precision Standardtests durchgeführt werden.

Die Cranfield Labormessung kann bei der Messung der Relevanzwirksamkeit der Informationswiedergewinnung im Web nicht verwendet werden, weil die Maße nicht auszurechnnen sind. Deswegen müssen für die Relevanzwirksamkeit der Wiedergewinnung von Informationen im Web neue Maße geschaffen werden.

Die neuesten Forschungen zeigen, daß das Suchen im Web drei Arten hat: Navigation, Orientierung und Transaktion. Eine der wichtigsten Navigationsaufgaben ist das Suchen von Webseiten, wobei das Ziel des Benutzers das Finden einer Webseite von einer Einheit (Unternehmen, Institution, Person, usw.) mit Hilfe einer Web-Suchmaschine ist.

Die Autorin untersucht die Wirksamkeit des Webseitensuchens aus dem Gesichtspunkt der Benutzer. Für die Messung werden von ihr zwei neue Maße eingeführt: das Pseudo Precision und das Mean Pseudo Rank Maß. Mit Hilfe von den Pseudo Precision und Mean Pseudo Rank Maßen gibt die Autorin das MICQ Verfahren an. Das MICQ Verfahren mißt im Web die Wirksamkeit der Identifizierung von Webseiten durch die Suchfragen.

CONTENTS

ACKNOWLEDGEMENT

This work would never have been written without the help and support of several people. First of all, I would like to thank Sándor Dominich his excellent supervision I received throughout my work.

Many thanks to Professor Ferenc Friedler, Department of Computer Science, University of Pannonia and Zoltán Birkner, University of Pannonia, Nagykanizsa, for the supportive environment.

In addition, I would like to express my full respect to my colleagues Júlia Góth and Tamás Kiezer for the joint work and the helpful comments in the period of writing the thesis.

A very special thank to István for his support and kindness.

The greatest acknowledgement I reserve for my family for their support to whom I dedicate this dissertation.

CHAPTER 1

INTRODUCTION

In information retrieval (IR) evaluation plays an essential role. Information retrieval system performance may be measured over many different dimensions, but the most important type of evaluation of IR systems is retrieval effectiveness evaluation, that is, how well a given system or algorithm can match, retrieve and rank documents that are the most useful or relevant to the user's information need.

There is a long tradition of experimental work in IR. The pioneering experiment was the Cranfield I in 1960 followed by a more substantial study in 1966. These experiments can claim to be responsible for founding the experimental approach in IR. Retrieval effectiveness evaluation is now usually based on a test reference collection and on standard evaluation measures precision and recall; this is called the Cranfield paradigm. Test collections make it possible for researchers to conduct retrieval tests in laboratories without having to find real users. Such collections allow for comparable results across systems. A number of test collections exist. The most popular standard test collections are ADI, CACM, CISI, MED, REUTERS, TIME, and TREC. These collections vary in size, topic and in the number of queries.

Exhaustive judging is infeasible in case of huge databases, especially when considering the Web. These pose problems for most evaluations, but especially when evaluating the effectiveness of Web search engines. The issues of evaluation of IR on the Web differ from the issues of evaluation of IR. The Web and then the processes of indexing and retrieval of Web pages are very different from those of classical information retrieval systems. This means that the traditional Cranfield type of evaluation is not usually possible in Web environment. The standard measures usually cannot be calculated. The limitations have led to calls for the development of new IR evaluation methods and measures. A detailed literature overview on retrieval effectiveness evaluation can be found in *Chapter 2*.

Traditional information retrieval evaluations and early Web experiments evaluated retrieval effectiveness according to how well methods can find documents that contain relevant text. Recent research suggests, however, that this kind of task is not a typical WWW search task (Broder, 2002). Three WWW-based retrieval tasks can be identified: navigational, informational, and transactional. The navigational task is when the purpose of a query is to reach a particular site that the user has in mind. The

user would like to retrieve this site either because he or she visited it in the past or because the user assumes that such a site exists. One of the most important navigational tasks is the home page finding task. The home page finding problem is one where the user wants to find a particular site and the querznames the site. Home page finding queries tipically specify entities such as people, companies, departments and products.The home page finding task is discussed in *Chapter 3*. The evaluation measure related to home page finding task is the Mean Reciprocal Rank (MRR). The MRR of each individual query is the reciprocal of the rank at which the correct response was returned, or zero if none of the first *N* responses contained a correct answer. The score for a sequence of queries is the mean of the individual query's reciprocal ranks. MRR measures the search engine's capability to find home pages.

In *Chapter 4* I address the home page finding problem from general users' point of view. This viewpoint shows how easily a user can find a home page using search engines. I propose two new measures – Pseudo Precision and Mean Pseudo Rank – to evaluate the effectiveness of home page identification on the Web. The measures are based on the Mathematical Theory of Reliability. The Mathematical Theory of Reliability is the overall scientific discipline that deals with general methods and procedures during the planning, preparation, acceptance and testing of devices. These methods and procedures ensure the maximum effectiveness of devices during use. The Mathematical Theory of Reliability develops methods of evaluating the reliabilities of devices and introduces various quantitative indices for measures of devices performance. The measures and concepts used in the dissertation are presented in *Section 4.1*. The Pseudo Precision and Mean Reciprocal Rank measures were elaborated using the hazard rate function of the Mathematical Theory of Reliability and the Mean Reciprocal Rank measure of retrieval effectiveness evaluation in information retrieval. Pseudo Precision was defined as the proportion of search engines that retrieve the relevant answer, i.e. the target Web page. Mean Pseudo Rank measures how easily a user can reach the target Web page looked for from the hit list. Mean Pseudo Rank considers two factors. The first one is the position, i.e., the rank of the target Web page in the hit list and the second factor considered is the linking structure of the hit list. The score for a group of search engines is the mean of the query's reciprocal rank in the individual search engines. Mean Pseudo Rank measures the query's identification capability.

Based on Pseudo Precision and Mean Pseudo Rank I propose the MICQ (Measure the Identification Capability of Queries) method in *Chapter 5* to measure the capability of search queries to identify the relevant answer using Web search engines.

In *Chapter 6,* the practical applications of the MICQ method are presented. Many people use the Web to obtain information from public institutions and organizations. Because users typically do not know the URL of the desired institution's home page, they use a Web search engine to get there. Thus in the applications it were investigated how easily users can find the home page of several categories of institutions. Institutions' names are usually difficult to recall exactly, thus they are not being used as queries in search engines. Instead, the acronyms of institutions are being used: they are easy to remember and are extensively used in media and by people in everyday life. Therefore, the home page identification capability of acronyms was investigated. This means that the home page finding problem is addressed form the users' point of view. It is evaluated how acronyms can identify its

institution on the Web when the acronym is the search expression. The identification capability of acronyms was evaluated according to the MICQ method. The MICQ method is language independent. Accordingly, the identification capability of several categories of acronyms of Hungarian and Danish institutions was evaluated. The results could give a situation report about how effectively users can find the institutions of a country on the Web.

Finally, *Chapter 7* gives a summary of the results obtained.

CHAPTER 2

INFORMATION RETRIEVAL EFFECTIVENESS EVALUATION

2.1 Evaluation in Information Retrieval

In information retrieval (IR) evaluation plays an essential role. Information retrieval system performance may be measured over many different dimensions, such as economy in the use of computational resources, speed of query processing or user satisfaction with search results. The most important type of evaluation of IR systems is retrieval effectiveness evaluation, that is, how well a given system or algorithm can match, retrieve and rank documents that are most useful or relevant to the user's information need (Mizzaro, 1997).

Retrieval effectiveness evaluation is usually based on a test reference collection and on evaluation measures. This kind of evaluation has more than a 40-year history (Rasmussen, 2002). It evolved from laboratory experimentation now called the Cranfield paradigm. The Cranfield tests were conducted by a group of researchers at the Cranfield College of Aeronautics. Its primary aim was to test the performance of different indexing techniques. The first set of experiments was conducted in 1958-1962. These experiments tested four indexing systems. The results were controversial. The controversy led to a critical examination of the methodology used. Cleverdon devised a second set of experiments with emphasis on rigour and a laboratory model. Cranfield II used 1400 documents and 279 queries. Research papers were used to instantiate queries and the document collection was comprised of the pooled references. Relevance judgments were made by the question providers and augmented by students who screened the entire collection (Spark Jones, 1981). Finally, recall and precision were the evaluation metrics used in the experiments.

Cranfield II thus became a basic model for information retrieval experimentation. This model comprises a document collection, a set of queries and associated relevance judgements by specialists – briefly called test collection –, and measurement usually based on precision and recall. Given a retrieval strategy, the evaluation measure quantifies for each query the similarity between the set of documents retrieved and the set of relevant documents provided by the specialist. Thus, it points the goodness of the retrieval strategy. Test collections allow for standard performance baselines, reproducible results, comparison of retrieval methods in terms of retrieval effectiveness and the potential for collaborative experiments.

Precision and recall are the standard measures for evaluating how well or badly an IR system performs. Let A be the number of retrieved documents in response to query Q, R be the total number of relevant documents to a query Q and G be the number of retrieved and relevant documents.

– *Precision* is defined as the proportion of retrieved documents that are relevant to a query.

$$Precision = \frac{G}{A} \tag{2.1}$$

– *Recall* is defined as the proportion of relevant documents that has been retrieved.

$$Recall = \frac{G}{R} \tag{2.2}$$

Test collections make it possible for researchers to conduct retrieval tests in laboratories without having to find real users. Such collections allow for (to some extent) comparable results across systems. A number of collections exist. The most popular standard test collections are ADI, CACM, CISI, MED, REUTERS, TIME and TREC. These collections vary in size, topic and in the number of queries.

The most frequently used test collection is the TREC. It was initiated in 1990. The purpose of TREC is to encourage research in information retrieval by providing a large test collection and to encourage communication among research groups, etc. (Baeza-Yates & Ribeiro-Neto, 1999). TREC collections are large, with a variety of documents and requests, and have a good range of relevant items. Relevance judgements come from a pooled output of many searches from many different systems. The Text REtrieval Conference (TREC) is now the major forum for laboratory Experiments. NIST – the National Institute of Standards and Technology in the United States – coordinates it. In TREC different types of tests (tracks) are proposed to investigate different IR tasks. As a result, researchers can compare their IR systems on a regular basis. However, this kind of evaluation poses some problems and trigger criticism (e.g. Saracevic, 1995; Tague-Sutcliffe, 1996; Ellis, 1996 and Wu and Sonnenwald, 1999)

2.2 Evaluation of Web Information Retrieval Effectiveness

Early test collections were small enough to permit relevance judgements for every document and every query. Exhaustive judging is infeasible in case of huge databases, such as TREC database, and especially when considering the Web.

As it is well known, the World Wide Web (or briefly Web, WWW) has become one of the most popular and important Internet applications both for users and for information providers, not only for scientists but also for everyone. The World Wide Web dates from the end of the 1980's (Berners-Lee et al., 1994). The extensive use of

the Web and its exponential growth are now well known (Risvik et. al., 2002). Just the amount of data available is estimated to be in order of terabyte. In addition to textual data, other media such as images, audio, video are also available. The Web can be seen as a large, unstructured and inhomogeneous database. These facts trigger the need for efficient tools to manage and retrieve information from this database.

There are three different forms of searching the Web (Baeza-Yates et al., 1999):

- The first is to use search engines that index a portion of the Web documents as a full-text database.
- The second is to use Web directories that classify selected Web documents by subject.
- The third is to search the Web exploiting its hyperlink structure.

More than 80% of Internet users rely on search engines to find the information they need (Dong, 2003).A search engine is a system of program designed to help find information stored on the World Wide Web. The search engine allows one to ask for content meeting specific criteria (typically those containing a given word or phrase) and retrieves a list of references (called hits) that match those criteria. A search engine operates in the following order:

- crawling,
- indexing,
- searching.

Web search engines work by storing information about a large number of Web pages, which they retrieve from the WWW itself. These pages are retrieved by a Web crawler, an automated Web browser that follows every link it sees. The content of each page is then analyzed to determine how it should be indexed (for example, words are extracted from the titles, headings, or special fields called meta tags). Data about Web pages is stored in an index database for use in later queries. Some search engines, such as Google, store all or part of the source page (referred to as a cache) as well as information about the Web pages, whereas some store every word of every page it finds, such as AltaVista. When a user comes to the search engine, makes a query – typically by giving key words – the engine looks up the index, and provides a ranked listing of best–matching Web pages according to its criteria, usually with a short summary containing the document's title and sometimes parts of the text.

The usefulness of a search engine depends on several factors, but mainly on the relevance of the results it gives back. While there may be millions of Web pages that include a particular word or phrase, some pages may be more relevant, popular, or authoritative than others. Most search engines employ methods to rank the results to provide the "best" results first. How a search engine decides which pages are the best matches, and what order the results should be shown in, varies widely from one engine to another. The methods also change over time as Internet usage changes and new techniques evolve.

In Web environment, Information Science currently relies on a methodology for measuring IR effectiveness that is based on the Cranfield paradigm developed in a prior information retrieval environment. On the one hand, information retrieval on the

Web is very different from retrieval in traditional document collections. This difference arises from several factors, e.g., the high degree of dynamism of the Web, its hyper–linked structure, the heterogeneity of document types, etc.. On the other hand, there are problems with applying the classical recall and precision measures to Web IR. Using small test collections it is possible to make relevance judgements for every document and every query. It is infeasible in case of Web. Given a relevance–assessed output, precision can be directly derived, while recall cannot. This is because recall depends not only on what was retrieved, but also on what was not retrieved (what was missed). Thus, recall requires the access to the complete set of documents that was searched. The limitations of recall are discussed in many research papers (Hull, 1993; Chu and Rosenthal, 1996; Ljosland, 1999; Oppenheim et. al., 2000 etc.). In large databases, it is also not possible to assess all documents retrieved as to relevance. In this case precision cannot be, actually, measured either. These pose problems for most evaluations, but especially when evaluating the effectiveness of Web search engines. This means that the traditional Cranfield type of evaluation is not usually possible in Web environment. Thus, it is important to question whether this methodology, which was developed for the batch retrieval era remains valid in Web IR.

Recent research suggests that new or revised evaluative measures are required to assess retrieval effectiveness of Web search engines (e.g., Gwizdka et. al., 1999; Agosti et. al., 2001; Bar-Ilan, 2005; Sufyan-Beg, 2005; Wang et. al., 2006). The limitations of precision and recall have led to calls for the development of new IR evaluation methods and measures.

In practice, the majority of the evaluations of search engines involve only precision. As precision cannot be measured, various numbers of the results are analysed for relevance (typically the first 5, 10, 20), and precision at *N* is measured (e.g., Leighton, 1995; Hawking, 1999; Leighton and Srivastava, 1997, 1999; Ljosland, 1999; Savoy and Picard, 2001). It decreases the amount of manual relevance assessments and focuses on those documents that are typically observed by the user. On the other hand, evaluation is carried out by employing relative recall rather than recall (e.g., Gordon and Pathak, 1999).

Some of the evaluations avoid both recall and precision, and apply alternative methodology for measuring the effectiveness of search engines. MacCall and Cleveland (1999) state that there are inherent problems with applying recall and precision metrics to Web IR. Instead, they propose a quantitative measure called Content–Bearing Click (CBC) Ratio. Its basis is the content–bearing click. It is defined as any hypertext click that is used to retrieve possibly relevant information as opposed to a hypertext click that is used for other reasons. Mizzaro (2001) proposes the Average Distance Measure (ADM) that measures the average distance between the actual relevance of documents (UREs) and their estimates by the IR system (SREs). Joachism's (2002) method is based entirely on clikthrough data that do not require manual relevance judgements unlike traditional methods that require relevance judgements by experts.

CHAPTER 3

THE HOME PAGE FINDING PROBLEM

3.1 Web Information Needs

Traditional information retrieval evaluations and early TREC Web experiments evaluated retrieval effectiveness according to how well methods find documents that contain relevant text. Recent research suggests, however, that this kind of task is not a typical WWW search task. Broder (2002) argues that WWW user information needs are often not of an informational nature and nominates three key WWW–based retrieval tasks:

- **Navigational.** The immediate intent is to reach a particular site or page. The purpose of such queries is to reach a particular site that the user has in mind, either because they visited it in the past or because they assume that such a site exists.
- **Informational.** The intent is to acquire some information assumed to be present on one or more Web pages. The purpose of such queries is to find information assumed to be available on the Web in a static form. No further interaction is predicted, except reading. By static form it is meant that the target document is not created in response to the user query.
- **Transactional.** The intent is to perform some Web–mediated activity. The purpose of such queries is to reach a site where further interaction will happen. This interaction constitutes the transaction defining these queries. Categories for such queries are e.g., shopping, finding various Web-mediated services, etc..

Navigational search, particularly home page finding, is the main motivation of the methodology within this thesis. In the following section, the home page finding retrieval task is discussed in details.

3.2 The Home Page Finding Problem

Evidence derived from query logs suggests that navigational search makes up a significant proportion of the total WWW search requests (Eiren et. al., 2003). The primary aim of a user wanting to obtain specific information is to get to the home page that contains the relevant answer as easily and quickly as possible (Silverstein et. al., 1998). On the other hand, the primary role of a Web page is that it can be easily found by users.

In principle, if a Web site exists, it should be possible for a user to find it. However, manually maintaining a directory of all Web sites is difficult because of Web's size and volatility. For this reason, effective home page finding is an interesting research problem. Most Web sites have a main entry page, sometimes also referred to as a home page. This page usually has introductory information for the site and navigational links to other main pages of the site.

The home page finding problem is one where the user wants to find a particular site and the query names the site. Home page finding queries typically specify entities such as people, companies, departments and products. A searcher who submits an entity name as a query is likely to be pleased to find a home page for that entity at the top of the list of search results, even if they were looking for information. In this way home pages may also provide primary–source information in response to informational and transactional queries (Broder, 1997).

The home page finding problem is different from a subject search where the user's query describes their topic of interest and the list of results should contain as many relevant documents as possible. Home page finding is similar to known item search, in that the user is looking for a particular item (site). However, in known item search the user has seen the item before, whereas home page finding may involve a known or unknown site. In addition, home page finding queries name the required site. Known item search queries might describe the topic of an item, rather than naming it.

For experienced Web users, effective site finding is most important in cases where the required URL is difficult to guess. For users less accustomed to URLs, the ability to enter a name rather than a URL is of even greater importance.

Example 3.1

Let us consider some example queries grouped into two categories. The first category contains queries that may be considered as site finding queries and are as follows:

- Where can I find the Web site of Nokia?
- Where is the Madonna's official home page?
- Where can I find Google?

The next category contains queries that are probably not site finding queries. These queries may be as follows:

- What is Information Retrieval?
- Where can I find airline timetables?
- Where can I find information about the World War II Normandy invasion?

The above examples indicate that different user information needs exist. Asking *What is MTA?* is different from *Where is the MTA home page?* (MTA = Magyar Tudományos Akadémia).

The presence of different information needs types also raises the question of query disambiguation. It seems impossible to determine whether the user is looking for a specific Web site or as many relevant pages as possible on a given topic given an one-word query.

Evaluation measures related to home page finding task are Mean Reciprocal Rank and Success Rate. Both the Mean Reciprocal Rank and Success Rate measures give an indication of how many low value results a user would have to skip before reaching the correct answer (Craswell et. al., 2001), or the first relevant answer (Shah et. al., 2004). The Mean Reciprocal Rank (MRR) measure is commonly used when there is only one correct answer. For each query examined, the rank of the first correct document is recorded. The score for that query is then the reciprocal of the rank at which the document was retrieved. The score for a system as a whole is taken by averaging the reciprocal rank across all queries. The Success Rate measure is often used when measuring effectiveness for exact match queries, such as home page finding and named page finding tasks. Success rate is indicated by S@k, where *k* is the cutoff rank and indicates the percentage of queries for which the correct answer was retrieved in the top *k* ranks (Craswell et. al., 2001b).

These measures may provide important insight as to the utility of a document ranking function. Silverstein et al. (1998) observed from a series of WWW logs that 85% of query sessions never proceed past the first page of results. Further, it has recently been demonstrated that more time is spent by users examining results ranked highly, with less attention paid to results beyond rank five (Upstill, 2005). All results beyond rank five were observed to, on average, be examined for 15% of the time that was spent examining the top result.

There are several papers describing experiments of the evaluation of the site finding capabilities of information retrieval algorithms and search engines. Laboratory testing of retrieval system evaluation follows the Cranfield paradigm (Baeza-Yates et al., 1999). Based on the Cranfield paradigm researchers perform experiments on test collections to compare the relative effectiveness of different retrieval approaches. The Text REtrieval Conference (TREC) is an example of the Cranfield evaluation paradigm. A statement of the purpose of the TREC conference can be found in the TREC Web site (TREC). A TREC workshop consists of a set of tracks, areas of focus in which particular retrieval tasks are defined, for example, Enterprise Track, Video Track, Web Track etc.. Web Track (Web Track) is a track that is featuring search tasks on a document set that is a snapshot of the World Wide Web. Starting in 2001 at TREC–2001 the Web Track (Craswell et al., 2001) includes the home page finding task with 145 homepage finding queries. The systems were compared based on the first correct answer. These evaluations used the following effectiveness measures. One was the Mean Reciprocal Rank of the first correct

answer (set to zero if no correct answer is listed in the top ten). The other was the Success Rate, the proportion of queries for which a correct answer appeared in the engine's top *N* (*N* usually equals 10). There were 43 official runs of the Home Page finding task. The Mean Reciprocal Rank of the first correct answer varied widely across the 43 runs. It ranged from 0.054 to 0.774. The proportion of queries for which a right answer was found in the top 10 results ranged from 13% to 88%. TREC–2002 Web Track (Craswell et al., 2002) included the named page task rather than the home page finding task. In this case the page was searched by name. The answer was only one target page, but not necessarily a home page. TREC–2003 Web Track (Craswell et al., 2003) involved a mixture of home page finding and name page finding tasks. In both cases there was only one target page. The importance of home pages in Web ranking was investigated via both a Topic Distillation task and a Navigational task. In the topic distillation task, systems were expected to return a list of the home pages of sites relevant to each of a series of broad queries. This differed from previous home page experiments in that queries may have multiple correct answers. The navigational task required the systems to return a particular desired Web page as early as possible in the ranking in response to queries. In half of the queries, the target answer was the home page of a site and the query was derived from the name of the site (home page finding) while in the other half, the target answers were not home pages and the queries were derived from the name of the page (named page finding). The two types of query were arbitrarily mixed and not identified. The navigational task results were as follows. Mean Reciprocal Rank varied from 0.067 to 0.727, while Success Rate at 10 varied from 9.3 to 89.3. TREC–2004 Web Track (Craswell et al., 2004) involved a mixed query stream, 75 home page finding queries, 75 named page finding queries and 75 topic distillation queries. The goal was to find ranking approaches that work well over the 225 queries, without access to query type labels. Mean Average Precision, Mean Reciprocal Rank of the first correct answer and Success@n (n = 1, 5, 10, the proportion of queries for which a good answer was at rank n) were used. The averages of the results ranged from 0.025 to 0.546.

In addition to laboratory experiments, real life experiment on the Web also investigates the home page finding capabilities of search engines. In Singhal and Kaszkiel's site finding experiment (Singhal et. al., 2001) the queries were taken from an Excite log and judged as home page finding queries. Craswell et al. (2001a) evaluated the effectiveness of 20 Web search engines on 95 site–finding queries. Each query named an airline with the correct answer being the airlines' official home page URL. Their results showed that the performance varied widely across the 20 engines.

Craswell et al. (2001b) compared the site finding effectiveness of a link–based ranking method and a content–based ranking method. The experiment was based on TREC methodology and the general Web crawl and university crawl were used as a test corpus. The Mean Reciprocal Rank of the first correct answer within the top 10 was 0.228 for the content method and 0.446 for the anchor method.

CHAPTER 4

PSEUDO PRECISION AND MEAN PSEUDO RANK: NEW MEASURES TO EVALUATE THE EFFECTIVENESS OF HOME PAGE IDENTIFICATION CAPABILITY ON THE WEB

Based on Chapter 3 it can be seen that originally the Home Page Finding problem is addressed from the search engines' point of view. The search engines are evaluated and compared. In the Home Page Finding problem the query is the name of the site and the target answer is the home page. The effectiveness of the search engine is evaluated using the Mean Reciprocal Rank measure. For each query the reciprocal rank of the firs correct answer is recorded. The reciprocal ranks are averaged across all queries. This score measures the effectiveness of the search engine. Based on this score search engines can be compared.

In this chapter the Home Page Finding problem is addressed not from an algorithmic (retrieval method) point of view but from a user's viewpoint. In the present Home Page Finding problem the query is an entity and the target answer is the entity's home page. The entity may be a person, institution, etc. It is evaluated how effectively the user can find the target home page.

I elaborated two new measures to evaluate the effectiveness of the home page identification capability on the Web. In *Section 4.2*, these new measures will be presented that I gave in [SKROP 4, SKROP 7]. The measures were derived from the Mathematical Theory of Reliability. However, the measures also preserve some characteristics of classical retrieval performance measures. The following section describes the concepts and measures of Mathematical Theory of Reliability that are used in *Section 4.2*.

4.1 Mathematical Theory of Reliability

Mathematical Theory of Reliability includes theoretical tools – e. g., mathematical models and methods – and also practical tools, whereby the reliability of devices (products, systems, components) can be specified, tested, predicted and demonstrated (Gnedenko et. al., 1969).

The reliability of a device is defined to be the probability of performing its purpose adequately for the time intended under the operating conditions encountered. Adequate performance indicates that failures must be clearly defined. The criteria for what is considered as satisfactory operation must be clearly specified. Reliability measures are related to time. Thus, it is possible to assess the probability of completing a task, which is scheduled to last for a given period. Operating conditions under which the reliability measure is derived also should be stated. Factors affecting operating conditions may have an effect on performance and should be included as part of reliability specifications. When conditions change, different values for reliability will result.

Mathematical and statistical methods can be used for quantifying reliability and for analysing reliability data. Difficulties arise in application of statistical theory to reliability, because the variation is often a function of time or time related factors. Therefore, reliability data from any past situation cannot be used to make forecast of the future behaviour without taking into account non–statistical factors such as design changes or unpredictable events such as service problems.

The simplest, purely inspectors' view of reliability is one in which a product is assessed against a specification or a set of attributes. However, this approach provides no measure of quality over a period. We therefore come to the need for a time–based concept of quality. The inspectors' concept is not time dependent. Either the product passes a given test, or it fails. Contrarily, reliability is usually concerned with failures in the time domain. This distinction marks the difference between traditional quality control and reliability theory.

An attempt to describe mathematically whether a system or device is working properly is a failure distribution. Failure is the partial or total loss of characteristics, which leads to a decrease (partial or total) of functionality. The modes of possible failure for an item in question affect the form of the failure distribution. Furthermore, systems and components can fail in several ways. Thus, the choice of failure distribution based on physical considerations is still nearly impossible.

Example 4.1

This example lists different failure types:

- static failure when a fracture occurs during a single load application;
- instability of a structure caused by strain energy stored in a member;
- chemical corrosion;
- sticking of mechanical assemblies; etc.

A concept that permits us to base the differentiation among distribution functions on physical considerations is the failure rate function $\lambda(t)$ (Barlow et. al., 1965). This is the most important measure of reliability. A failure rate is the average frequency with which a device fails. A device can be an electric bulb, a computer, etc. The failure rate depends on the failure distribution, which describes the probability of failure prior to a specified time. Failure rate is defined as the probability of failure in a finite interval of time, say of length x, given time t.

By this definition the failure rate is

$$\lambda(t) = \frac{R(t) - R(t + \Delta t)}{\Delta t R(t)} \tag{4.1}$$

where R denotes the reliability function, F is the failure distribution function and

$$R(t) = 1 - F(t) \tag{4.2}$$

The empirical value of failure rate is the number of failures that can be expected to take place over a given unit of time. The failure rate is determined as follows (Gnedenko et. al., 1969). Perform experiments with N copies of a device. Let $n(t)$ be the number of surviving devices at time t. Then the failure rate is:

$$\lambda(t) = \frac{R(t) - R(t + \Delta t)}{\Delta t R(t)} \approx \frac{\frac{n(t) - n(t + \Delta(t))}{N}}{\Delta t \frac{n(t)}{N}} = \frac{\Delta n}{\Delta t n(t)} \tag{4.3}$$

where

- Δn : number of failures in $(t, t + \Delta t)$,
- Δt: time period.

One of the primary objectives in system reliability analysis is to obtain a failure rate function of the device.

The failure rate is not always constant. The failure rate of a device may vary with time, such that a single number does not accurately describe the failure rate during all intervals of time. So the hazard function is used to describe the instantaneous failure rate at any point in time, which is usually called the hazard rate (Nash, 2003). By calculating the failure rate for smaller and smaller intervals of time Δt, the interval becomes infinitely small. This results in the hazard function $h(t)$, which is the instantaneous failure rate at any point in time.

$$h(t) = \lim_{\Delta t \to 0} \frac{R(t) - R(t + \Delta t)}{\Delta t R(t)} \tag{4.4}$$

According to *Equation 4.3* the empirical hazard function is as follows:

$$h(t) = \lim_{\Delta t \to 0} \frac{\Delta n}{\Delta t n(t)} \tag{4.5}$$

Practically, in considering the hazard rate of a device, N copies of the device (sample) are tested at a certain point in time. The number of failures in the sample is determined.

Then the hazard rate is as follows:

$$h(t) = \frac{\text{number of failures}}{\text{size of the sample}} \tag{4.6}$$

Example 4.2

Suppose it is desired to estimate the hazard function of a certain device. A test can be performed to estimate its hazard rate. Let the device be a light bulb. Let the sample be ten identical light bulbs. The sample is tested until either they burn out or reach 1000 hours, at which time the test is terminated for that device. The results are as follows:

Light bulb	Hours	Failure
Light bulb 1	1000	No failure
Light bulb 2	1000	No failure
Light bulb 3	467	Failed
Light bulb 4	1000	No failure
Light bulb 5	630	Failed
Light bulb 6	590	Failed
Light bulb 7	1000	No failure
Light bulb 8	285	Failed
Light bulb 9	648	Failed
Light bulb 10	882	Failed

The failure rate is varying with time. In the (0, 1000) interval the failure rate is $\lambda(t) = \frac{6}{1000 \cdot 10} = 0.0006$, while in the (0, 5000) interval $\lambda(t) = \frac{2}{500 \cdot 10} = 0.004$. Thus, the hazard function is used to describe the instantaneous failure rate at any point in time. E.g., the hazard function at 600 hours is $h(600) = \frac{3}{10} = 0.3$

4.2 Pseudo Precision and Mean Pseudo Rank: New Measures to Evaluate the Effectiveness of Home Page Identification Capability on the Web

In the present section I am going to present the new measures I gave in [SKROP 4, SKROP 7] to evaluate the effectiveness of home page identification capability on the Web. The conceptual and notational framework used is given by Mathematical Theory of Reliability and classical retrieval effectiveness evaluation.

The primary aim of Reliability Theory is to determine whether a device performs adequately under predefined operating conditions. The probability of adequate

performance is called reliability. Reliability is determined by measuring the hazard function (see equations 4.6) of the device by testing *N* copies of that device (sample). This means that each device in the sample is tested whether it is working (satisfactory operation), and the proportion of failures is determined.

In Web Information Retrieval, the retrieval effectiveness of information retrieval systems (search engines) is evaluated. A search engine attempts to help a user locate desired information on the Web. The search engine allows users to ask for content meeting specific criteria (typically those containing a given word or phrase) by entering a query and retrieving a ranked list of Web sites that match those criteria. A special case of information retrieval is the Home Page Finding task. In the Home Page Finding task the user's query names an entity (e.g. a company name) and the relevant answer is the home page of the entity. In the original Home Page Finding task, the retrieval effectiveness is evaluated from the search engine's point of view. Namely, the IR method of the particular search engine is investigated. Furthermore, the home page finding retrieval effectiveness of several search engines can be compared.

Let us consider the relevance effectiveness evaluation of the Home Page Finding task from the users 'point of view. In this viewpoint, the basic concepts of the Mathematical Theory of Reliability are used. The following parallel can be drawn between the basic concepts of information retrieval and Reliability Theory. A search engine is a device. The aim is to determine the reliability of this device under specific operating conditions. In the Home Page Finding task we say that a search engine performs adequately if it retrieves the home page the user wants to locate on the Web. Otherwise, the search engine has failure. Reliability is determined by measuring the hazard rate of the search engine by testing a group of search engines. This means that each search engine in the group is tested whether it is working. A search engine is working if it retrieves the relevant answer i.e. the home page the user wants to locate on the Web. In Reliability Theory, the reliability of a device is investigated by taking and testing *N* copies of the device. In this methodology, the Home Page Finding problem is investigated from common users' point of view. The investigation does not consider either the search engine or the IR method of the search engine. The working hypothesis is that different search engines are identical from common users' point of view. Common users do not know how search engines operate. They do not know the IR method of search engines. In this regard, the search engine is a tool that can be used to locate information on the Web. It makes no difference which search engine is chosen by the user. The goal is to find the desired home page. Furthermore, different users are using different search engines. Thus in the present Home Page Finding task the sample is consists of *N* different search engines. However, the search engines can be regarded identical from the viewpoint of general users.

Table 4.1 summarizes the parallel between Reliability theory and information retrieval concepts.

Table 4.1 Parallel between Reliability theory and IR concepts**.**

Reliability Theory Concepts	Information Retrieval Concepts
device	search engine
adequate operation	relevant answer retrieved
failure	relevant answer not retrieved

We say that a search engine is working if it retrieves the relevant answer. Otherwise, it has failure. Consequently, hazard function in IR is defined as the proportion of search engines that do not retrieve the relevant answer.

However, IR is usually interested in effectiveness, namely, precision measures the proportion of relevant answers. Since the hazard function measures the proportion of failures, hence using its complementary a more optimistic measure called Pseudo Precision can be introduced as follows [SKROP 4, SKROP 7]:

$$\text{Pseudo Precision} = 1 - \text{Hazard rate} \tag{4.7}$$

Pseudo Precision, denoted by Π_a, is defined as follows:

$$\Pi_a = \frac{r_a}{N} \tag{4.8}$$

where:

- r_a number of search engines that return the relevant answer when the query is a
- N: number of search engines

Pseudo Precision is the proportion of search engines that retrieve the relevant answer. Pseudo Precision has values between zero and one.

Table 4.2 shows how Pseudo precision can be derived from hazard function.

Table 4.2 The derivation of Pseudo Precision from the Hazard function.

Hazard Function	Pseudo Precision
$h(t) = \frac{\text{number of failures}}{\text{size of the sample}}$	$\Pi_a = \frac{r_a}{N}$
number of failures	r_a: number of search engines that retrieved the relevant answer when the query is *a,* r_a = size of the sample − number of failures
Size of the sample	*N*: number of search engines

In Reliability Theory operating conditions under which the reliability measure is derived also should be stated. Factors affecting operating conditions may have an effect on performance. When conditions change, different values for reliability will result. Taking into account the above consideration of Web IR evaluation with Pseudo Precision, the factor that affects the value of the measure is the parameter *a*, the query. The query names the target Web page the user wants to locate on the Web. Performing the same test (the group of search engines are the same) with different queries may affect different Pseudo Precision values. However, it has to be noted that the value of Pseudo Precision may also be affected by the selected group of search engines. The sample should be selected in such way that it can reflect the search engine usage behaviour of common users. If the sample is selected so, then the measure cam indicate how effectively users can find the desired home page on the Web.

Based on Pseudo Precision and IR measurements (see *Section 2.1* and *Section 2.2*) more articulate measures can be defined.

Here Pseudo Rank is defined. This measure is based on the one hand on Pseudo Rank, on the other hand on Mean Reciprocal Rank (see *Section 3.2*). Mean Pseudo Rank measures how easily users can reach the target Web page from the hit list. It is derived from Pseudo Precision as follows. Pseudo Precision investigates binary operation modes of search engines. A search engine either retrieves the relevant answer or does not retrieve it. Namely, it only considers the presence of the target Web page in the hit list. Pseudo Rank considers two more factors. The first one is the position, i.e., the rank of the target Web page in the hit list. The second factor considered (do not considered in the original reciprocal rank measure) is the linking structure of the hit list. Thus, first the retrieved hits are categorized for example according to the following categories:

- *Category* 1: link to the target Web page. This Web page is desired to be retrieved when the user enters the query.
- *Category* 2: link to a page or site page (i.e., it is not the target page) that contains a site map or a navigational link to the target page that is desired to be retrieved when the user enters the query.

- *Category* 3: irrelevant link. It is neither a link to the desired target page nor a link to a page or site page that contains a site map or a navigational link to the target page.

We say that a search engine is operating adequately if it retrieves *Category* 1 or *Category* 2 hits. Otherwise, it has failure. The parallel between Pseudo Precision and Pseudo Rank is shown in Table 4.3.

Table 4.3 The parallel between measures Pseudo Precision and Pseudo Rank.

Search engine operation	Pseudo Precision	Pseudo Rank
adequate operation	relevant answer retrieved	Category 1 or Category 2 hit is retrieved
failure	relevant answer is not retrieved	only Category 3 hits are retrieved

Based on the above considerations, Pseudo Rank – denoted by PR_{ia} – is calculated by taking into account both the categorization and the rank of the links in the hit list as follows [SKROP 4, SKROP 7]:

$$PR_{ia} = \begin{cases} \dfrac{1}{r_{ia}} & category 1 \text{ link in position } r_{ia} \\ \dfrac{1}{\kappa \times r_{ia}} & category 2 \text{ link in position } r_{ia} \text{ and no } category 1 \text{ link} \\ 0 & \text{no link in } categories \text{ 1 or 2} \end{cases} \quad (4.9)$$

where r_{ia} is the rank of the target Web page for query a in the hit list of search engine i and κ is a penalty factor. The penalty factor κ can be used to penalize the search engine if it retrieves only *category 2* links. In this case $\kappa > 1$. If there is no penalty then κ may be equal to 0.

Pseudo Rank is calculated for one search engine. Thus, this measure can be used to evaluate the effectiveness of a given search engine. However, now the effectiveness of home page identification capability is evaluated from the users' viewpoint. Thus, to get a measure of reliability Pseudo Rank has to be measured over the group of search engines.

Thus, an average Pseudo Rank is defined for query *a*, called *Mean Pseudo Rank*, denoted by MPR_a, as follows:

$$MPR_a = \frac{1}{N}\sum_{i=1}^{N} PR_{ia} \tag{4.10}$$

where *N* is the number of search engines used. The Mean Pseudo rank is calculated by averaging the Pseudo Rank across all search engines.

The derivation of Mean Pseudo Rank from Pseudo Precision is as follows. In Pseudo Precision the number of search engines that return the relevant answer – when the query is *a* – is determined. In Pseudo Rank – denoted by PR_{ia} – both the categorization and the rank of the links in the hit list are taken into account. Thus, the Pseudo Rank is assigned to the relevant answer. The Pseudo Rank values are added across all search engines, i.e., in Pseudo precision $\Pi_a = \frac{r_a}{N}$ the numerator is replaced with the sum of the pseudo ranks as follows $r_a \rightarrow \sum_{i=1}^{N} PR_{ia}$ .and Mean Pseudo Rank is calculated by dividing the sum of the Pseudo Rank values with the number of search engines, i.e., $MPR_a = \frac{1}{N}\sum_{i=1}^{N} PR_{ia}$.

Mean Pseudo Rank (MPR) is different from Mean Reciprocal Rank (MRR) (described in Section 3.2). MRR is calculated for a search engine by averaging the reciprocal rank over all queries, while MPR averages the Pseudo Rank values of search engines in case of a given query and additionally considers the linking structure of the hit list.

Example 4.3

Assume that a user wants to obtain information from the Budapesti Gazdasági Főiskola (BGF) on the Web. He / she does not know the URL of the desired home page, thus the typical scenario is as follows. He or she selects a search engine. The selected search engine now is Google. The user enters the acronym of the institution as a query and examines the first page of the hit list. So our user enters BGF as query to search engine Google, and examines the first page of the retrieved hit list (users typically do not examine more links). The hit list is on Figure 4.1.

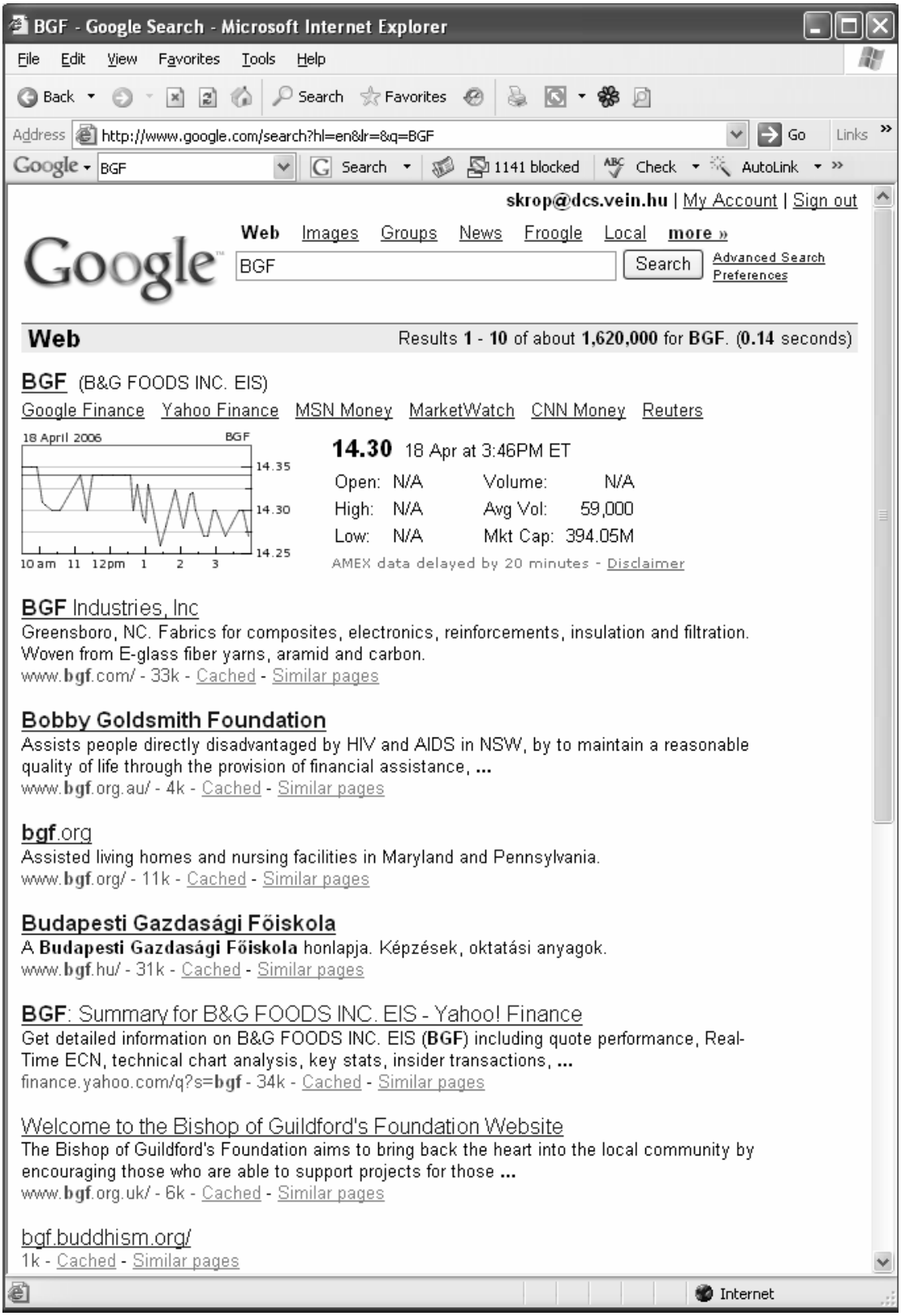

Figure 4.1 The answers retrieved by Google in response to query BGF.

The desired home page is the fifth in the hit list.

The measures are calculated as follows:

- $\Pi_{BGF} = \frac{r_{BGF}}{N} = \frac{1}{1} = 1$ The Pseudo Precision equals one. The only search engine in the sample, i.e., Google retrieved the relevant answer.
- $PR_{Google,BGF} = \frac{1}{r_{Google,BGF}} = \frac{1}{5} = 0.2$ The Pseudo Rank equals 0.2 because the rank of the target web page in the hit list is 5. The Mean Pseudo Rank also equals 0.2 because only one search engine is considered in this example (i.e., $N = 1$).

Summary

[Theses T1]

I proposed the Pseudo Precision and Mean Pseudo Rank measures to evaluate the home page identification capability of queries on the Web.

CHAPTER 5

METHOD TO MEASURE THE HOME PAGE IDENTIFICATION CAPABILITY OF QUERIES ON THE WORLD WIDE WEB

In this chapter I am going to present a method I gave in [SKROP 3, SKROP 6 SKROP 5]. The method can be used to measure the home page identification capability of Web queries in Web search engines. It can be measured how easily a user can find the desired home page using Web search engines The practical motivation of the method is the Home Page Finding problem that is described in *Chapter 3*. The home page finding problem is one where the user wants to find a particular site and the query names the site.

5.1 The MICQ Measurement Method

In this section I am going to present a method to measure the capability of queries to identify home pages on the Web. The method is called MICQ (**M**easure the **I**dentification **C**apability of **Q**ueries). In MICQ the identification capability is the ability of the query to identify the relevant home page in Web search engines. The method was developed based on the Pseudo Precision and Mean Pseudo Rank measures. The measures are described in *Section 4.2*. Pseudo Precision was defined as the proportion of search engines that retrieve the desired home page. Mean Pseudo Rank measures how easily a user can reach the desired home page from the hit list.

The MICQ method has the following steps:

Step 1. Definition of experimental setting:

- Choose database: in this methodology the whole WWW.
- Identify pairs: identify a set of (query, home page) pairs. E.g., (OMSZ, http://www.met.hu/). Each pair represents a query and the target home page. The user enters the query and he / she would like to retrieve the target Web page.
- Choose search engines: select search engines that will be used to evaluate the identification capability of queries.

Step 2. Implementation of experiments:

- Formulate queries.
- Run search engines: for each query being evaluated run the queries for each search engines.
- Examine the results: categorise the retrieved results according to predefined relevance categories.

Step 3. Study of the identification capability of the queries:

- Measure the identification capability: apply some measures to measure the identification capability.
- Create histograms with the results obtained.
- Draw conclusions.

5.2 Implementation of the Method

The first step of MICQ is to determine the experimental setting. MICQ measures the identification capability of queries in Web search engines, thus the database is the Web. One has to identify (query, home page) pairs. For each target Web page one has to determine the query that is meant to identify that page. The URL of the page has to be recorded. It is recommended to create a table that list the queries, the home pages and the URL of the home pages. After that, search engines have to be selected to measure the identification capability. The search engines should be selected so that the sample can reflect the search engine usage behaviour of users.

The next step is the implementation of the experiment. Enter each query for each of the search engines and investigate the results. The links retrieved by search engines are to be assigned to predefined relevance categories. The set of criteria for categorizing links is as follows:

- *Category* 1: link to the target Web page. This is the Web page that is desired to be retrieved when the user enters the query. This link is identified when (query, home page) pairs are defined.
- *Category* 2: link to a page or site page (i.e., it is not the target page) that contains a site map or a navigational link to the target page that is desired to be retrieved when the user enters the query.
- *Category* 3: irrelevant link. It is neither a link to the desired target page nor a link to a page or site page that contains a site map or a navigational link to the target page.

To evaluate the identification capability of queries Pseudo Precision and Mean Pseudo Rank (*Chapter 4 Section 4.2*) can be used.

Example 5.1

Let us suppose that a user wants to find the home page of "*Magyar Tudományos Akadémia*" (Hungarian Academy of Sciences). The Magyar Tudományos Akadémia has the acronym MTA. Because the expression "Magyar Tudományos Akadémia" is long, thus the user uses its acronym as query. The URL of this home page is http://www.mta.hu/.

The identification capability of MTA can be measured as follows. First, create a table as follows:

Query	Name of Home Page	URL of Home Page
MTA	Magyar Tudományos Akadémia	http://www.mta.hu/

Now select search engines to measure the identification capability of MTA. In this example, six search engines are selected: Heuréka, AltaVizsla, Ariadnet, Google, Metacrawler and AltaVista. MTA is entered to each of the six search engines and the hit lists are investigated. The first ten hits retrieved by search engines are assigned to the above defined relevance categories.

For the acronym *MTA* the following rankings were obtained:

Search Engine	Category	Ranking
Heuréka	2	4
AltaVizsla	2	5
Ariadnet	1	7
Google	1	1
Metacrawler	0	0
AltaVista	1	1

To evaluate the identification capability of MTA Pseudo Precision and Mean Pseudo Rank have to be calculated.

Pseudo Precision means the proportion of search engines for which the query identifies the target Web page. The more search engines retrieve the target Web page, the more useful the query is. The Pseudo Precision of a query is directly proportional with the number of search engines that retrieve the target Web page. The Pseudo Precision can be calculated using *Equation 4.8.*

In this example the target Web page was found by Ariadnet, Google and AltaVista. These search engines retrieved *Category 1* hits. The desired home page was not found by Heuréka, AltaVizsla, and Metacrawler. They have not retrieved *Category 1* hit. Thus, the Pseudo Precision Π_{MTA} of *MTA* is 3/6 = 0.5. Three from six search engines have found the home page.

While Pseudo Precision is a measure of the usefulness of a query, a more articulate and combined measure should give an indication of how easy it is for the user to get to the home page looked for from the hit list. Thus for the query the Pseudo Rank is measured. The higher the Pseudo Rank is, the better the acronym is, i.e., the higher the Pseudo Rank is, the easier it is for the user to get to the desired home page. Pseudo Rank and the number of links to be examined are inversely proportional with each other. The Pseudo Rank can be calculated using *Equation 4.9*.

In this example the desired home page was found by Ariadnet (*Category 1* hit retrieved), and was ranked in the 7th position. Thus, the *Pseudo Rank* $PR_{Ariadnet,\ MTA}$ = 1/ 7 = 0.14. By analogy with this $PR_{Google,\ MTA}$ = 1/ 1 = 1 and $PR_{AltaVista,\ ,MTA}$ = 1/ 1 = 1. Heuréka and Altavizsla retrieved *Category 2* hit. In this case according to *Equation 4.9* a penalty factor is used. Now the penalty factor is set to two. Thus, $PR_{Heuréka,\ MTA}$ = 1/ (2 x 4) = 0.125 and $PR_{AltaVizsla\ ,MTA}$ = 1/ (2 x 5) = 0.1. $PR_{Metacrawlert,\ MTA}$ = 0 because this search engine have not retrieved either *Category 1* or *Category 2* hit. In this example, the rank is taken as being the sequence number of the link in the hit list, but it could also be taken as the relevance value – if this is known – of the link given by the search engine. The penalty factor is could be taken as being equal to any other positive integer.

Because we want a measure for the usefulness degree of query regardless of the search engines used, an average of the Pseudo Rank called Mean Pseudo Rank is to be calculated using *Equation 4. 10* as follows.

MPR_{MTA} = (0.14 + 1 + 1 + 0.125 + 0.1 + 0) / 6 = 0.39

Finally, if there are more queries the results can be represented by a histogram, and conclusion can be drawn.

Summary

[Theses T2]

I proposed the MICQ method to measure the home page identification capability of search queries on the Web**.**

CHAPTER 6

STUDY OF THE IDENTIFICATION CAPABILITY OF ACRONYMS ON THE WEB

In this chapter I am going to present the practical applications of the MICQ method I gave in [SKROP 3, SKROP 6, and SKROP 5].

6.1 Background

As it is well known, the Web become one of the most popular and important application both for users and information providers. Web pages can be classified into several category labels, e.g. Yahoo! (Yahoo) organizes Web pages into a hierarchy consisting of thousands of category labels. One important category of information stored in Web pages is the generic category of institutions that includes the Web pages of institutions and organizations of interest to a large mass of users such as state departments or ministries, financial institutions, public transportation companies, libraries, civil organizations, political parties, public health institutions etc..

The primary aim of a user wanting to obtain information from a specific institution is to get to the home page of that institution as easily and quickly as possible. On the other hand, the primary aim of an institution is that its home page be easily found by users.

Many people use the Web to obtain information from public institutions and organizations. A Web user has three alternative ways to reach a Web page on the Web:

- The user knows the page's URL from various sources such as advertisements, newsletters, etc..
- Through a navigational link for the URL, this may be on some other Web pages.
- Through search engine results in response to a query.

Since most users typically do not know the URL of the desired institution's home page, they use a Web search engine to get there. The typical scenario is as follows:

(i) select a Web search engine,
(ii) enter the acronym (or full name) of the institution as a query,
(iii) examine the first page (or two) of the hit list.

Institutions usually have long, multiple words official names. In addition, every institution has its own official acronym that uniquely identifies it.

Users may prefer using acronyms as queries in Web searching for various reasons, such as for example:

- They usually do not know the full names of the institutions exactly.
- Acronyms of institutions' names are commonly and very often used in both media and by people in everyday life.
- The number of single short queries like acronyms submitted to search engines has tripled in four years (Spink et al., 2002).

6.2 Acronyms

The widespread use of acronyms in daily life is a relatively modern phenomenon and the result of growing literacy in the 19th and 20th centuries. In more restricted circumstances, however, they have been in use for thousands of years; both the Roman and Hebrew cultures used them (*Internet Acronym Server*[1]).

An acronym is a pronounceable abbreviation of a compound, name or phrase used as one word, often composed of the initial letters or syllables of the items abbreviated (*Webster's Online Dictionary*[2]). Acronyms are generally formed with capital letters. When first defining an acronym the first letter of each word within the definition is usually capitalized. Depending on the basis of the abbreviated form acronyms can be pronounced as

- a word, e.g., NATO: **N**orth **A**tlantic **T**reaty **O**rganization ,
- a series of the names of the letters, e.g., IBM: **I**nternational **B**usiness **M**achines,
- or some combination of the two, e.g., JPEG: **J**oint **P**hotographic **E**xperts **G**roup.

Sometimes non-initial letters are included in the acronym to make it pronounceable, e.g., Interpol: **Inter**national Criminal **Pol**ice Organization.

There is debate over whether the word acronym can be applied to any set of initials. Some people insist an acronym is only a set of initials, which is pronounceable as a word. Under this view, sets of initials like BBC and IBM are initialisms and not acronyms. However, for many people, the word acronym is used for all such sets of initials regardless of whether they are pronounced as a word or as the names of the letters in sequence.

[1] http://silmaril.ie/cgi-bin/uncgi/acronyms
[2] http://www.websters-online-dictionary.org

Acronyms are not necessarily unique. It is quite common to find polysemous acronyms, with their definitions not even coming from a related domain. An example of this is MTA. The *Acronym Finder*[3] Web site has 91 definitions for MTA (retrieved 24.07.06), ranging from Mail Transport Agent and Metropolitan Transit Authority to Magyar Tudományos Akadémia. In normal texts, non-uniqueness does not pose a problem: usually the meaning is clear from the context of the document. However, ambiguity is likely to be an issue if acronyms are extracted from large, heterogeneous collections. Acronyms are generally three or more characters in length, although two-character acronyms exist (for example AI for Artificial Intelligence). Because of the small number of combinations, two-character acronyms exhibit far greater scope for ambiguity (for instance Artificial Intelligence versus Artificial Insemination) (Yeates et al., 2000).

There are many acronym dictionaries available, both in print and electronically; many are domain specific, while others try to be broader in their approach. Acronym dictionaries available in print can only give a snap-shot of acronyms defined at the time of publishing and may out of date. Electronic acronym dictionaries are available as searchable databases, allowing the user to search for an acronym meaning. Often these databases are out of date, are domain specific or lack sufficient coverage leaving the user without a meaning for their acronym. Acronym dictionaries available online are as follows.

Acronyma[4] provides the users an interface to 472670 acronyms. The *Acronym Finder* is a World Wide Web searchable database of abbreviations and acronyms about computers, technology, telecommunications, and the military. Acronym Finder currently has over 475000 human-edited definitions for acronyms and abbreviations. *Acronym Search*[5] has about 50000 acronyms and abbreviations in many categories, including chat, computer, military, finance, accounting, airports, sports, classified, etc.. *Special Dictionary Acronyms, Abbreviations and Initialisms*[6] contains 583391 acronyms and abbreviations. The *Internet Acronym Server* is collecting acronyms from all over the Internet. The database of *Acronyms and Abbreviations*[7] consists of more than 12000 acronyms. However, the site reports that some of them may be erroneous or garbage. *AbbreviationZ*[8] is a directory and search engine for acronyms, abbreviations and initialisms with 357217 entries. Acronym dictionaries usually rely on users to submit new data in order to keep their acronym definitions current, which can lead to erroneous input and poor quality data if not moderated correctly. Due to the sheer volume of acronyms to consider, acronym databases are incomplete.

The following example will show how effectively one can find the meaning of an acronym using online databases. The above listed seven online acronym databases were used to find the meaning of the acronyms. The experiment was carried out in October 2006. Acronyms are not necessarily unique. Thus, the definition of the

[3] http://www.acronymfinder.com/
[4] http://www.acronyma.com/
[5] http://www.acronymsearch.com/
[6] http://www.special-dictionary.com/acronyms/
[7] http://www.chemie.fu-berlin.de/cgi-bin/acronym
[8] http://www.stands4.com/

acronyms was defined in advance. Three acronyms were used in the experiment. The acronyms were selected from different domains and countries:

- MTA: Magyar Tudományos Akadémia
- ACM: Association for Computing Machinery
- DTM: Danmarks Teknise Universitet

Table 6.1 shows the number of results the individual databases gave back and the rank of the predefined definition. The results are as follows. Two of the seven databases have found the meaning of MTA. In addition, two databases gave back the meaning of DTU, but DTU has better ranking than MTA. The meaning of ACM was found by five databases. The results show that these online databases may not be able to infer a meaning for an acronym given a specific domain.

Table 6.1 The number of results the databases gave back and the rank of the predefined meaning of the acronym.

Database	MTA		ACM		DTU	
	Number of results	Rank	Number of results	Rank	Number of results	Rank
Acronyma	29	0	32	1	5	1
Acronym Finder	92	4	91	2	14	1
Acronym Search	23	0	6	2	2	0
Special Dictionary	77	20	128	108	15	0
Internet Acronym Server	6	0	7	0	3	0
Acronyms and Abbreviations	7	0	4	4	2	0
AbbreviationZ	27	0	53	1	6	0

There are applications that use technologies from Natural Language Processing to identify acronyms automatically in text. In recent years there have been several attempts to create automatic acronym identification systems, such as *Acrophile*[9] (Larkey et al., 2000), Polyfind (Pustejovsky et al., 2001), the Biomedical Abbreviation Server (Chang et al., 2000) and *ARGH*[10] a biomedical acronyms database for specialists.

The need for acronym databases stems from the great number of acronyms present in technical reports, medical journals, newspapers and media. A second use of acronym databases is aiding in construction of a useful search system. Acronyms are

[9] http://ciir.cs.umass.edu/irdemo/acronym/index.html
[10] http://invention.swmed.edu/argh/

synonyms of their definitions; they are interchangeable in their usage. Search engines can use the information contained within acronym databases to act as list of synonyms; in this case it is preferable if the context of the acronym is known, allowing the search engine to substitute the correct acronym definition.

Automatic identification of acronyms allows the construction of large acronym databases. In order for significant acronym databases to be built it will be necessary to construct systems that are able to continually crawl the Web processing documents for acronyms. The main problem in acronym identification is not in identifying possible acronyms but in identifying correct acronym-meaning pairs. Finding the correct definition to an acronym is a challenging task. Acronym identification has been considered several times previously and various different techniques and approaches have been tried (Yeates et al., 2000; Taghva et al., 1998; Wren et al, 2002). However, at present time there may no be system that will infer a meaning for an acronym given a specific domain.

6.3 Motivation

There are many studies on evaluating the effectiveness of Web search engines (Chu et al., 1996; Gordon et al., 1999; Leighton et al., 1999; Oppenheim et al., 2000). In addition acronym as a topic are present in research and applications (see previous section), but in a different way as in this chapter.

Acronyms of institutions' names are commonly and very often used in both media and by people in everyday life. Thus, the aim of this chapter is to investigate the ability of the acronyms of institutions' names to find the home page of their own institutions when being used as queries in Web searching. This kind of identification capability of acronyms is called the usefulness of acronyms.

Several categories of institutions present on the Web are studied, a list of useful acronyms is given, causes of uselessness are presented, and possible remedies are suggested.

Based on the MICQ method described in Section 5.2 five applications have been elaborated to study the usefulness of acronyms on the Web. Different categories of institution acronyms were studied:

- In *Section 6.4* the usefulness of acronyms of Hungarian general institutions was investigated.
- In *Section 6.5* the usefulness of acronyms of Hungarian government offices was investigated and compared with the usefulness of acronyms of Hungarian general institutions.
- In *Section 6.6* the usefulness of acronyms of Hungarian higher educational institutions was investigated and compared with the usefulness of acronyms of Hungarian general institutions.
- In *Section 6.7* the usefulness of acronyms of Danish higher educational institutions was investigated and compared with the usefulness of acronyms of Hungarian higher educational institutions.

– In *Section 6.8* the usefulness of acronyms of Hungarian and Danish parties was investigated and compared.

In each application, the identification capability of acronyms was studied according to the MICQ method. The experimental settings and the implementation of the experiments were as follows.

In every application, institutions were identified that have acronyms and are present with their own website on the Web. The institutions were identified using different Web sites and directories. The specifications of these Web sites can be found in the respective sections. Each institutions home page was visited and lists were compiled containing the full name, home page URL and the acronym of each institution.

Search engines were selected and used to evaluate the usefulness of the acronyms. In different applications different group of search engines were used. The selection of search engines was according to some search engine usage statistics. The experiments were carried out by entering the acronyms to the selected search engines. In order to try to minimize biases (it is well known that biases, both conscious and unconscious, do affect any such test to a certain extent, and this cannot be totally excluded), exactly the same search expressions were entered to the search engines. This means that the searches were based on the exact official form of the acronyms. It was ignored whether search engines regard or disregard capitalization so as not to give advantage any of the search engines.

Only the first ten hits returned, i.e., the first page of hits, were evaluated for every acronym and search engine. The examination of the first ten hits is suggested because on the one hand it would be extremely time consuming to assess each page in the entire retrieved set, on the other hand because users typically do not examine more links. Spink et al. (2002) reported that the trend of viewing fewer pages of search results is going up.

Every link was assigned to exactly one of the categories in Section 4.2. The categories are recalled here:

- *Category* 1: link to the home page of the institution. This Web page is desired to be retrieved when the institution's acronym is used as query.
- *Category* 2: link to a page or site page (i.e., it is not the home page) that contains a site map or a navigational link to the home page.
- *Category* 3: irrelevant link. It is neither a link to the desired home page nor a link to a page or site page that contains a site map or a navigational link to the home page.

6.4 Measuring the Home Page Identification Capability of the Acronyms of Hungarian Institutions

The goal of this application is to evaluate the usefulness of acronyms of Hungarian institutions on the Web, i.e. the ability of the acronyms of Hungarian institutions' names to find the home page of their own institutions when being used as queries in Web searching. The experiment was as follows [SKROP 4].

120 institutions in Hungary that have acronyms and are present with their own Web site on the Web were identified. The institutions were identified using Web sites that list several categories of institutions. The Web sites Startlap, Index, Wahoo, Webmania were used to identify institutions on the Web. A list was compiled containing the full name, home page URL, and the acronym of each institution. The full list is not included in this chapter; Table 6.2 shows a fraction of it. The full list can be found in *Appendix A.1*.

Table 6.2 Full name, home page URL and acronym of institutions in Hungary.

Full Name	Home Page URL	Acronym
Budapesti Közlekedési Vállalat	http://www.bkv.hu/	BKV
Magyar Energia Hivatal	http://www.eh.gov.hu/	MEH
Országos Meteorológiai Szolgálat	http://www.met.hu/	OMSZ
Országos Közoktatási Intézet	http://www.oki.hu/	OKI

Six Web search engines were used to evaluate the usefulness of the acronyms. Table 6.3 presents the selected search engines.

Table 6.3 List of selected search engines: the first three are Hungarian search engines; the next three are general search engines.

Name of the Search Engine	URL of the Search Engine
Heuréka	http://www.heureka.hu
AltaVizsla	http://www.altavizsla.hu
Ariadnet	http://www.ariednet.hu
Google	http://www.google.com
Metacrawler	http://www.metacrawler.com
AltaVista	http://www.altavista.com

Heuréka, AltaVizsla and Ariadnet were selected and used because they are the most frequently used Hungarian search engines (they are hosted and operated in Hungary) in Hungary, which primarily index and search Hungarian Web pages. They are preferred by most Hungarian users, who are lay people and have language difficulties when trying to use search engines in another language. However, three well-known general (not Hungarian) search engines (Google, Metacrawler, AltaVista) were also used because, on the one hand, they are preferred by the computing society, and, on the other hand, non-Hungarian speaking people might want to find out information on Hungarian institutions, (e.g., when they plan to travel to Hungary, or if they live in Hungary.

The experiment was carried out during September – October 2002 by entering each acronym to each of the six search engines, and evaluating the first ten hits using the MICQ method. Thus, some eight thousand hits were examined, because there was *Category 2* and *Category* 3 links in the hit list as well. The hit list of a search engine

for a query can be checked about in two minutes because only the first ten hits returned is evaluated and the URL of the correct answer is known. Thus the evaluation of 120 queries in six search engines takes 120 x 2 x 6 = 1440 minutes = 24 hours.

The results obtained in this application are presented in details in *Appendix B Table B.1*.

Figure 6.1 shows a Pseudo Precision histogram of acronyms for all search engines. It can be seen that the majority of acronyms are useful (Pseudo Precision is greater than 0.5), a few are very useful (Pseudo Precision equals 1) and about 17% can hardly be judged as being useful.

Figure 6.2 shows Pseudo Precision histograms separately for Hungarian and general search engines. It can be seen that the majority of the acronyms have better identification capability in general search engines than in the Hungarian ones. While the average Pseudo Precision is 0.44 in Hungarian search engines, it is much higher, 0.78, in general search engines. This result is perhaps unexpected in that one would have thought that the acronyms should have good identification capability in Hungarian search engines as well. The differences in performance of acronyms may be because each individual search engine uses its own unique algorithm to index and rank Web sites, and the algorithms use various factors to rank pages in their search results. Search engines may provide basics of their indexing and page-ranking policies, however the Hungarian search engines used do not provide the same.

Figures 6.3, 6.4 and 6.5 show the Mean Pseudo Rank histogram for all search engines, for Hungarian search engines, and for general search engines, respectively. It can clearly be seen that, as expected based on Pseudo Precision, the degree of usefulness of about half of the acronyms is much higher in the case of general search engines than Hungarian ones. Average values of the Mean Pseudo Ranks are shown in Table 6.4.

Table 6.4 Average Mean Pseudo Rank of acronyms.

	Average *MPR*
Over all search engines	0.53
Over Hungarian search engines	0.38
Over general search engines	0.68

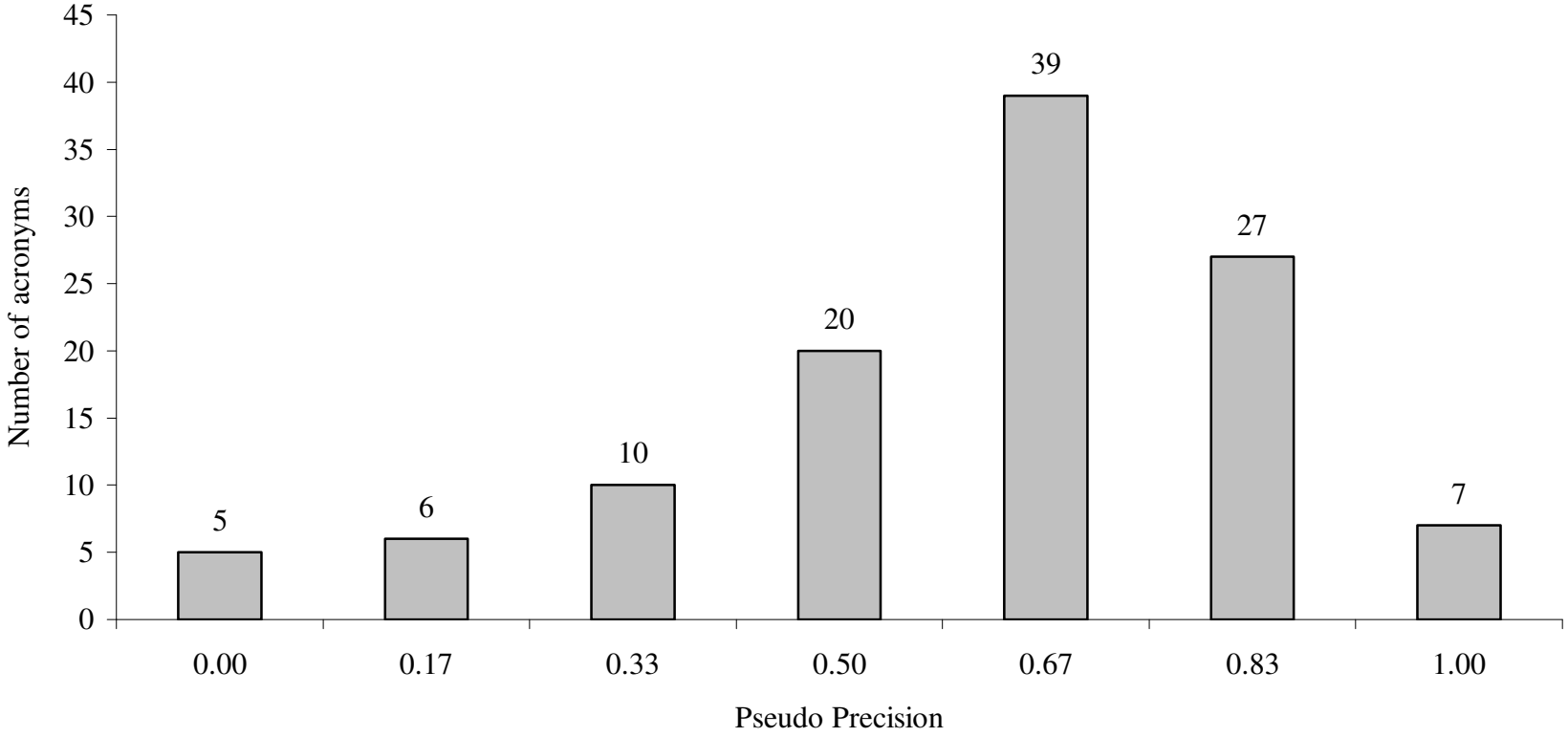

Figure 6.1 Pseudo Precision histogram of acronyms of Hungarian general institutions over all search engines.

On the horizontal axis Pseudo Precision intervals are presented. The intervals represented are half-closed and only the upper endpoints are included.
On the vertical axis, the frequency of Pseudo Precision is shown, i.e., the number of acronyms that fall into the intervals of Pseudo Precision.
36% of acronyms can hardly be judged as being useful (Pseudo Precision is less than 0.5); 64% of the acronyms are useful (Pseudo Precision is greater than 0.5).

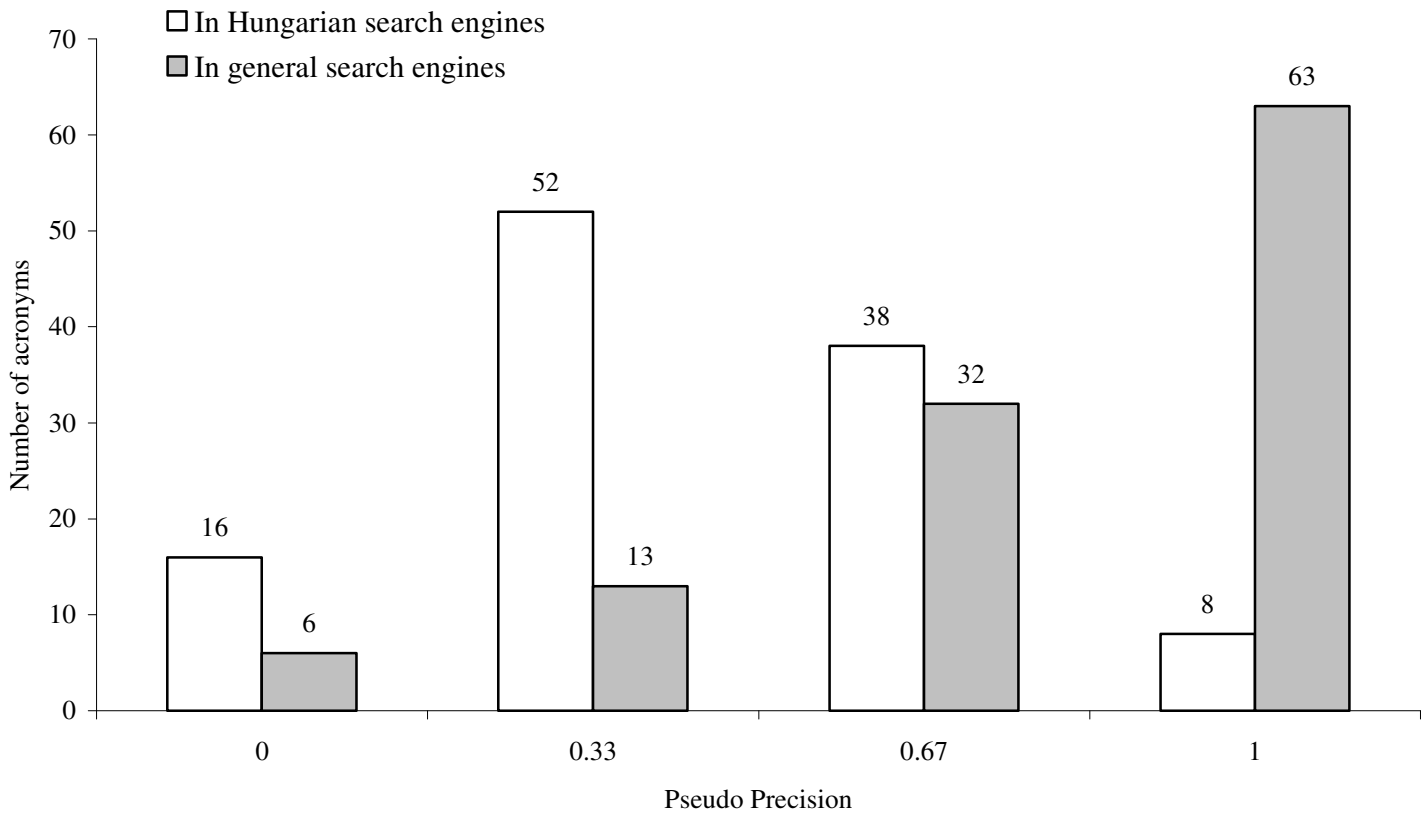

Figure 6.2 Pseudo Precision histogram of acronyms of Hungarian general institutions over Hungarian and general search engines.

On the horizontal axis Pseudo Precision intervals are presented. The intervals represented are half-closed and only the upper endpoints are included.
On the vertical axis, the frequency of Pseudo Precision is shown, i.e., the number of acronyms that fall into the intervals of Pseudo Precision.
60% of acronyms over Hungarian search engines and 17% over general search engines can hardly be judged as being useful (Pseudo Precision is less than 0.5); 40 % of the acronyms over Hungarian search engines and 83 % over general search engines are useful (Pseudo Precision is greater than 0.5).

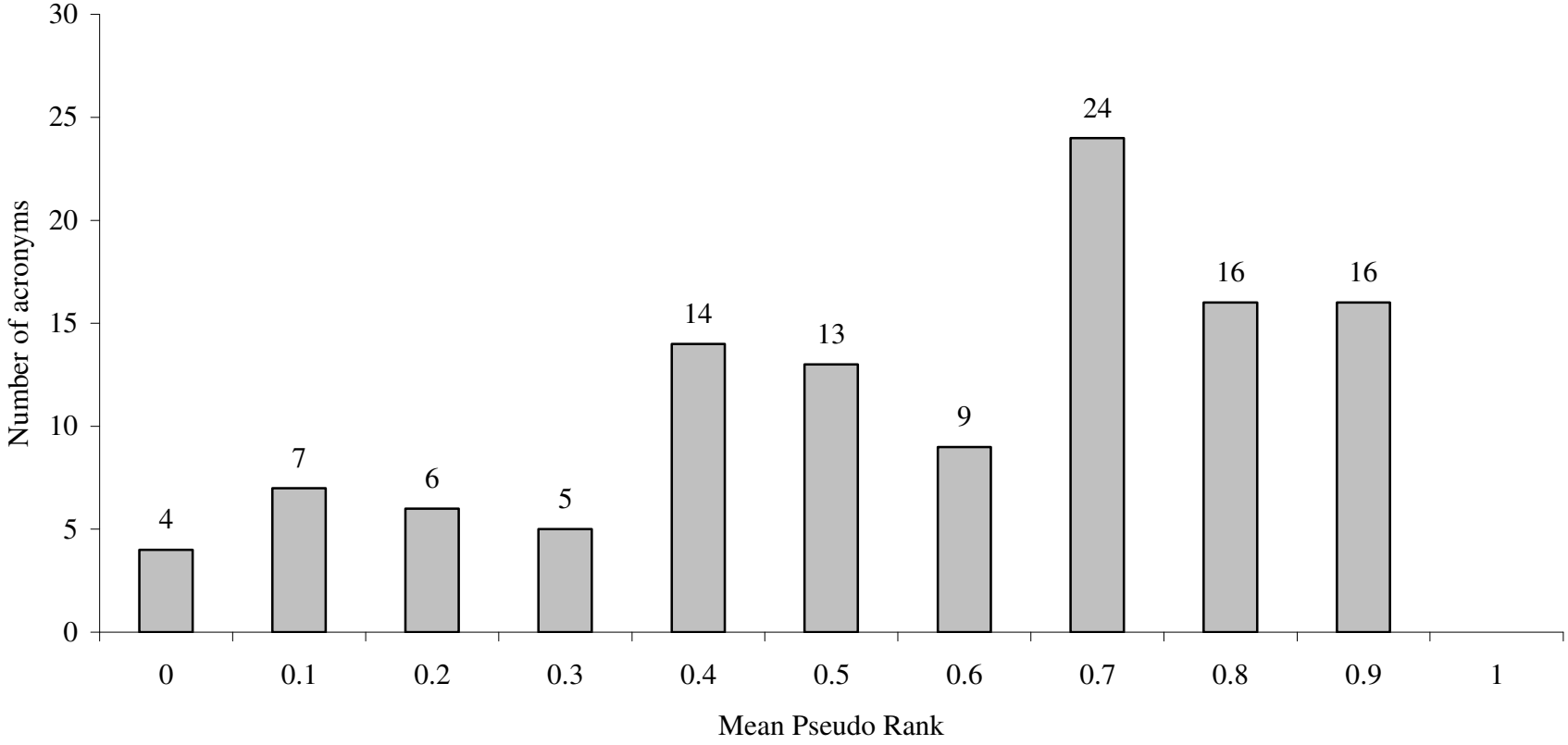

Figure 6.3 Mean Pseudo Rank histogram of acronyms of Hungarian general institutions over all search engines.

On the horizontal axis Mean Pseudo Rank intervals are presented. The intervals represented are half-closed and only the upper endpoints are included. On the vertical axis, the frequency of Mean Pseudo Rank is shown, i.e., the number of acronyms that fall into the intervals of Mean Pseudo Rank. 57% of acronyms have Mean Pseudo Rank less than 0.5; 43% of the acronyms have Mean Pseudo Rank greater than 0.5.

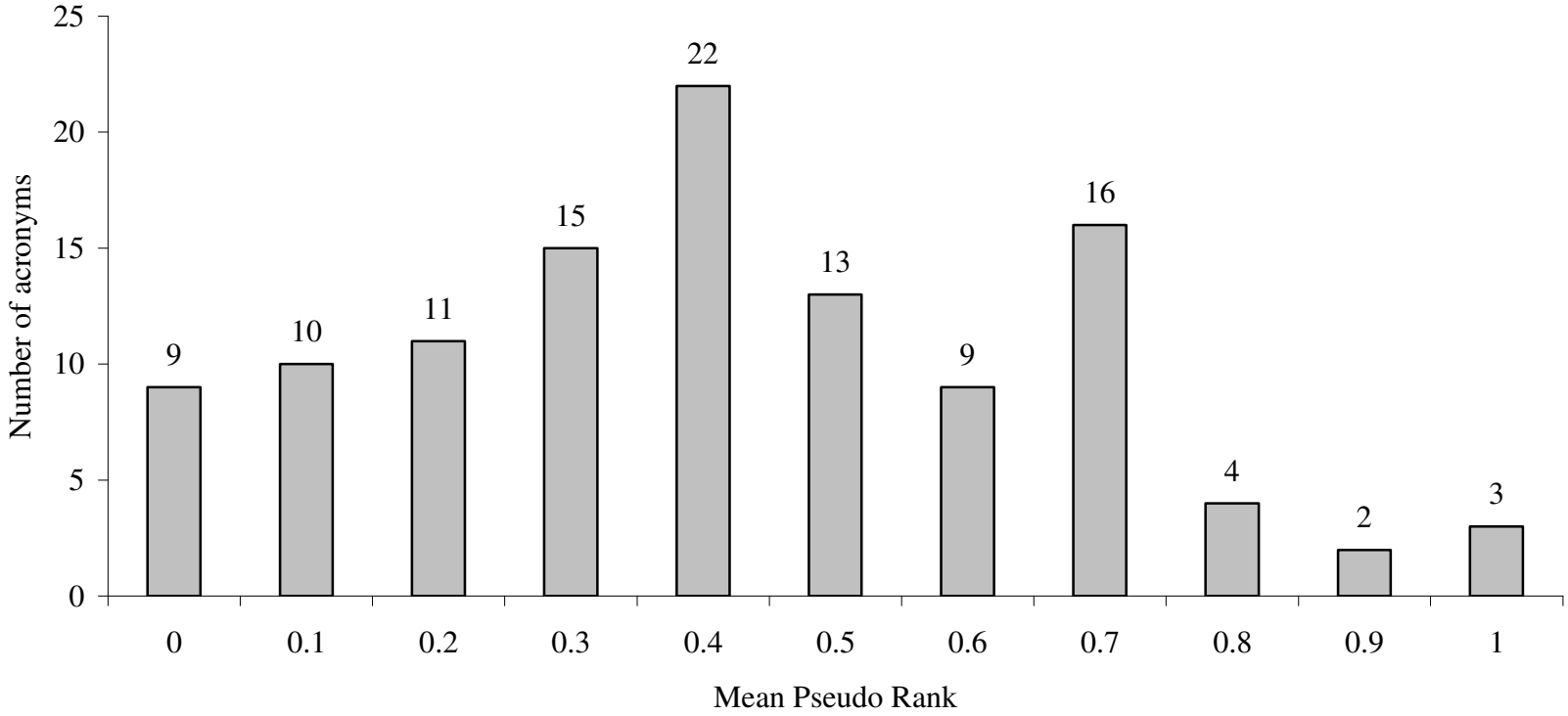

Figure 6.4 Mean Pseudo Rank histogram of acronyms of Hungarian general institutions over Hungarian search engines.

On the horizontal axis Mean Pseudo Rank intervals are presented. The intervals represented are half-closed and only the upper endpoints are included. On the vertical axis, the frequency of Mean Pseudo Rank is shown, i.e., the number of acronyms that fall into the intervals of Mean Pseudo Rank. 70% of acronyms have Mean Pseudo Rank less than 0.5; 30 % of the acronyms have Mean Pseudo Rank greater than 0.5.

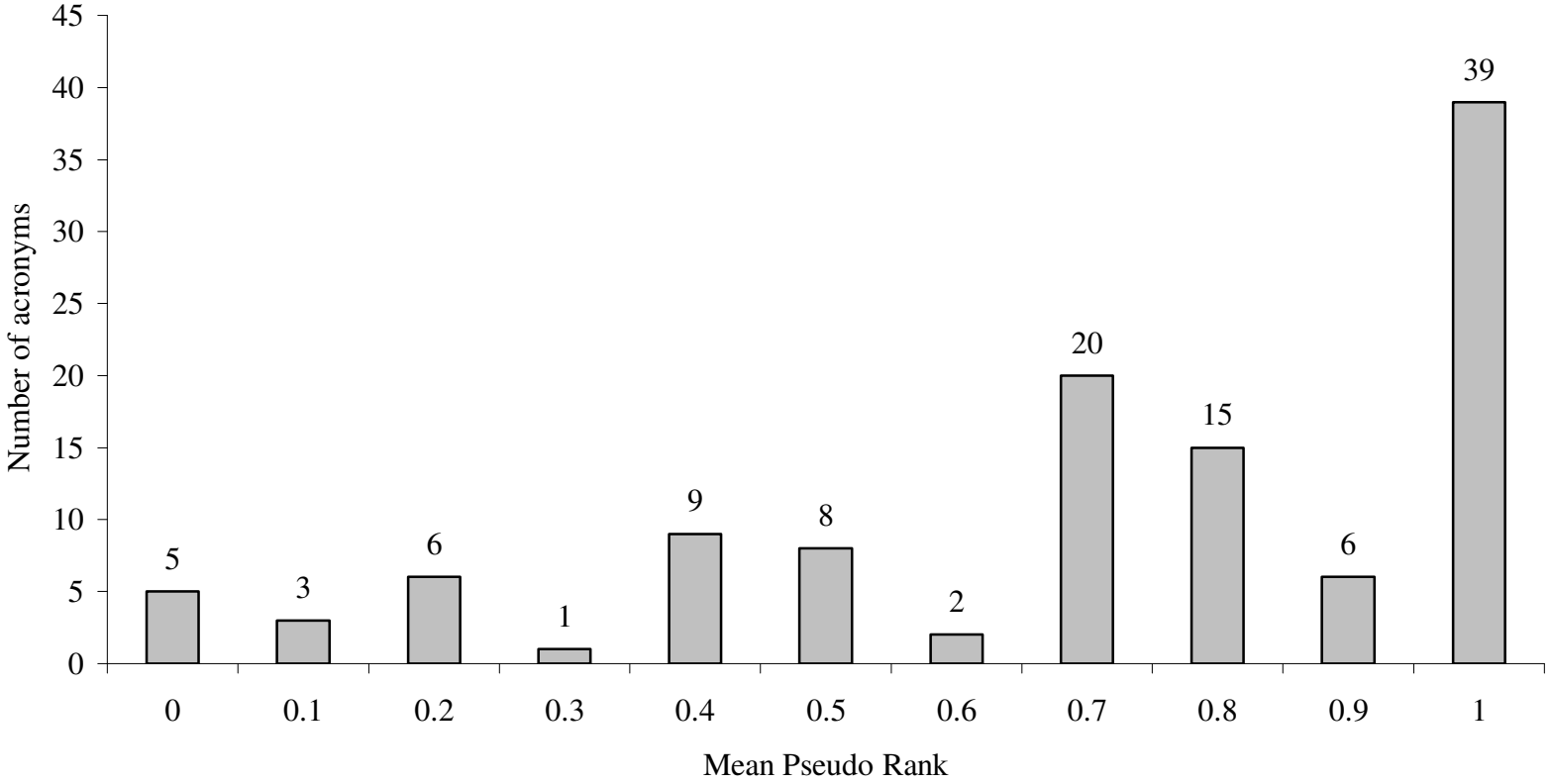

Figure 6.5 Mean Pseudo Rank histogram of acronyms of Hungarian general institutions over general search engines.

On the horizontal axis Mean Pseudo Rank intervals are presented. The intervals represented are half-closed and only the upper endpoints are included. On the vertical axis, the frequency of Mean Pseudo Rank is shown, i.e., the number of acronyms that fall into the intervals of Mean Pseudo Rank.
28% of acronyms have Mean Pseudo Rank less than 0.5; 72% of the acronyms have Mean Pseudo Rank greater than 0.5.

Based on the results, 34 acronyms (28%) identify their institutions in Hungary. Table 6.5 shows the acronyms of these institutions.

Table 6.5 The useful Hungarian acronyms on the Web.

ÁNTSZ	NDP	MKIK	matáv	OEP	MKB	KÉE
FVF	BGF	MKVK	MKGI	OFA	ISM	SOTE
GVH	APEH	MOL	MGYOSZ	OMIKK	TÁRKI	SZTE
MABISZ	DOSZ	MOK	NIIF	KSH	MSZP	ZMNE
MFB	MGYK	MSZT	NIOK	MEHIB	SZDP	

Based on the sample used this low identification capability of acronyms can be seen as a noteworthy situation, this may be due to the following causes.

Poor Web page design of home pages and sites seems to be one of the causes. Apart from content information (using the acronym as content or meta-data) page-related factors (format, placement of the title tag, frequency of keywords etc.) and overall Web page design also contribute to search engine ranking. The usage of title tags, fonts, character sizes, colours as well as the content need to be considerably revised and improved by Webmasters.

Another cause may be that quite a few acronyms have multiple meanings. For example, the acronym *MNB* (which identifies the Hungarian National Bank) also stands for the following institutions: Magyar Nemzeti Bibliográfia, Magyar Nemzeti Bizottság, Moffat, Nichol & Bonney INC., Moody National Bank, Malvern National Bank, which, due to a much better Web page design, are retrieved more easily (ahead of the Hungarian *MNB*).

The results show that the majority of the acronyms are not effective in identifying their institutions. This means that

(i) they fail to fulfil their roles of identifying their institutions,
(ii) Webmasters should seek ways to improve the situation by a more careful design of home pages taking into account the different ways in which different search engines index and rank Web pages,
(iii) the acronyms should be revisited as far as their uniqueness and identification property are regarded (although it is very improbable that, for example, the Hungarian Academy of Sciences (MTA) or Hungarian National Bank (MNB) would even consider changing their acronyms).

In this application the identification capability of the acronyms of Hungarian, institutions were evaluated. The evaluation is meant to represent the Web searching behaviour of users all over the world. Thus, Hungarian and not Hungarian search engines were used in the experiment. The selected institutions represent only a fraction of the Hungarian institutions that are present on the Web. The sample contains financial institutions, higher educational institutions, government offices, etc.. Based on this sample the results show that the identification capability of these acronyms is about 0.5.

MICQ is language independent; therefore, it can be used for carrying out similar experiments in other countries, as well. Considering the results of the experiment, it would be useful to apply this method to investigate the usefulness of several categories of institutions and to repeat the experiment in other countries, where acronyms are being used. The unexpected results raise several questions.

- Are the acronyms of other institutions not useful either?
- Is it (nearly) impossible to find an institution on the Web if only the acronym is known?

These are relevant questions where acronyms of institutions' names are commonly and very often used both in media and by people in everyday life.

Based on the above considerations additional applications have been elaborated to investigate the identification capability of acronyms. These applications concentrate on one institution category. Because the MICQ method is language independent not only Hungarian, but also Danish acronyms were investigated. The applications were as follows.

6.5 Measuring the Home Page Identification Capability of the Acronyms of Hungarian Government Offices

In this section the MICQ method is applied to evaluate the identification capability of acronyms of Hungarian government offices from the viewpoint of Hungarian users. The experiment was as follows.

12 Hungarian government offices that have acronyms and are present with their own Web site on the Web were identified. The government offices were identified using the Magyarorszag.hu Web site. A list was compiled containing the full name, home page URL, and the acronym of each institution. The full list is not included in this section. The full list can be found in *Appendix A.5.* This list represents all government offices in August 2006. Six Web search engines were used to evaluate the usefulness of the acronyms. Table 6.6 presents the selected search engines.

Table 6.6. List of selected search engines.

Name of the Search Engine	URL of the Search Engine
Google.co.hu	http://www.google.co.hu/
Lap.hu	http://lap.hu/
Kurzor	http://www.kurzor.hu/
Vizsla24	http://www.vizsla24.hu/i
Yahoo!	http://www.yahoo.com/
MSN	http://www.search.msn.com

The search engines were selected according to an analysis of Jároli (2006). This analysis shows the distribution of search engines used by Hungarian users. Figure 6.6 presents the snapshot of the distribution of search engines. The picture was taken from the Web site referred above.

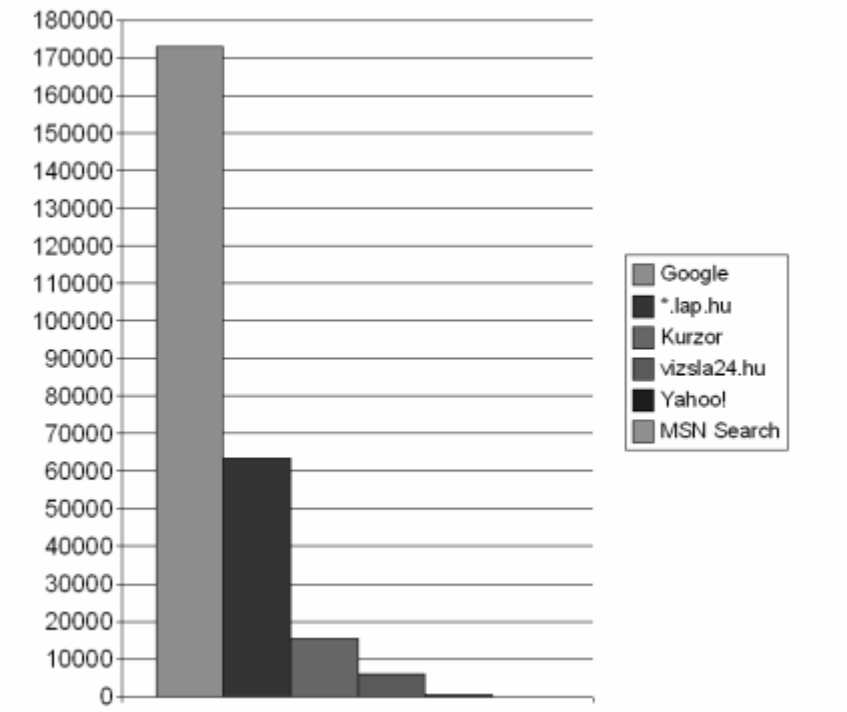

Figure 6.6 Distribution of search engines used by Hungarian users in 2005.
(Source: http://webni.innen.hu/Keres_c5_91oldalak_c3_b6sszehasonl_c3_adt_c3_a1sa2005Augusztus)

The experiment was carried out during September 2006 by entering each acronym to each of the six search engines, and evaluating the first ten hits using the MICQ method. Based on the above search engine usage statistics the Pseudo Precision and Mean Pseudo Rank values were calculated using a weighting scheme. The weighting scheme takes into account the distribution of search engine usage. The weighting scheme was as follows:

Google.co.ho	$w_1 = 0.67$
http://lap.hu/	$w_2 = 0.24$
Kurzor	$w_3 = 0.06$
vizsla24	$w_4 = 0.025$
Yahoo	$w_5 = 0.004$
MSN	$w_6 = 0.0004$

The weighting scheme shows that 67% of Hungarian users are using the Google.co.hu search engine, 24% are using lap.hu, 6% are using Kurzor, 2.5% are using vizsla24, 0.4% are using Yahoo and only 0.004% are using MSN. The results obtained in this application are presented in details in *Appendix B Table B.2*. The calculation of the measures was as follows:

$$\Pi_a = \frac{\sum_{i=1}^{N} v_i}{N} \tag{6.1}$$

where

- N is the number of search engines
- $v_{i=}\begin{cases}0, \text{if search engine } i \text{ have not retrieved the desired homepage}\\ w_i, \text{otherwise}\end{cases}$

$$PR_{ia} = \begin{cases} w_i \times \dfrac{1}{r_{ia}} & category 1 \text{ link in position } r_{ia} \\ w_i \times \dfrac{1}{\kappa \times r_{ia}} & category 2 \text{ link in position } r_{ia} \text{ and no } category 1 \text{ link} \\ 0 & \text{no link in } categories \text{ 1 or 2} \end{cases} \quad (6.2)$$

where r_{ia} is the rank of the target Web page for query a in the hit list of search engine i and κ is a penalty factor.

Figure 6.7 shows the Pseudo Precision of the acronyms. The acronyms are categorised according to usefulness categories. The usefulness categories are defined as follows:

- Not useful: Pseudo Precision equals 0,
- Somewhat useful: 0 < Pseudo Precision ≤ 0.5
- Useful: 0.5 < Pseudo Precision < 1,
- Very useful: Pseudo Precision equals 1.

It can be seen that the majority of acronyms are useful. Nine of the twelve acronyms are useful, and three are very useful. The average value of the Pseudo Precision is 0.74.

Figure 6.8 shows the Mean Pseudo rank of the acronyms. The average value of the Mean Pseudo Rank is 0.55. Only FVM's Mean Pseudo Rank equals one. Other five acronyms have high Mean Pseudo Rank. The remainder five acronyms have low Mean Pseudo Rank. They cannot effectively identify their institutions home page on the Web.

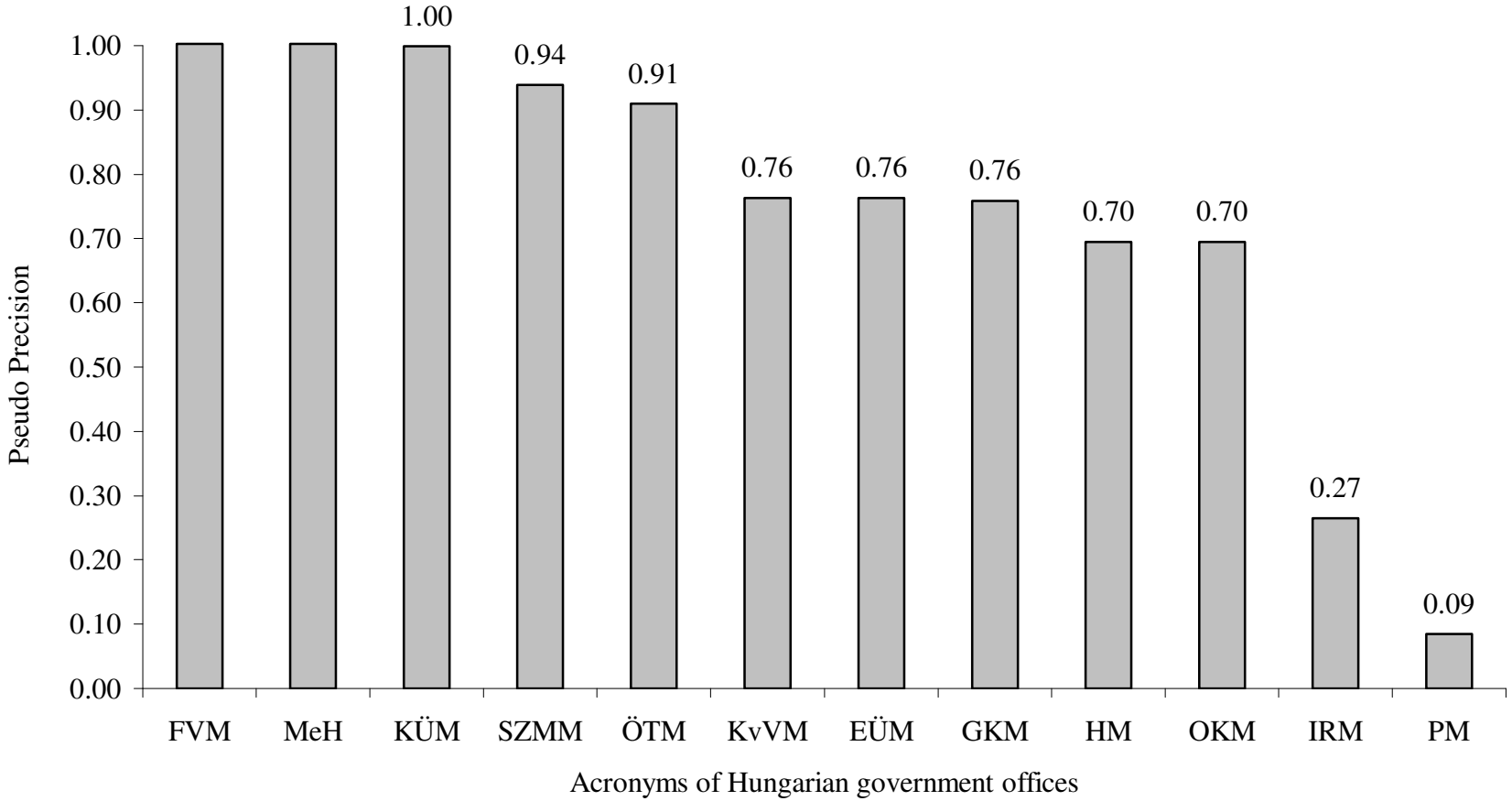

Figure 6.7 Pseudo Precision of the acronyms of Hungarian government offices.

The majority of acronyms are useful. Nine of the twelve acronyms are useful. They Pseudo Precision are greater then 0.5. FVM, MeH and KÜM are very useful. They Pseudo Precision equal 1.

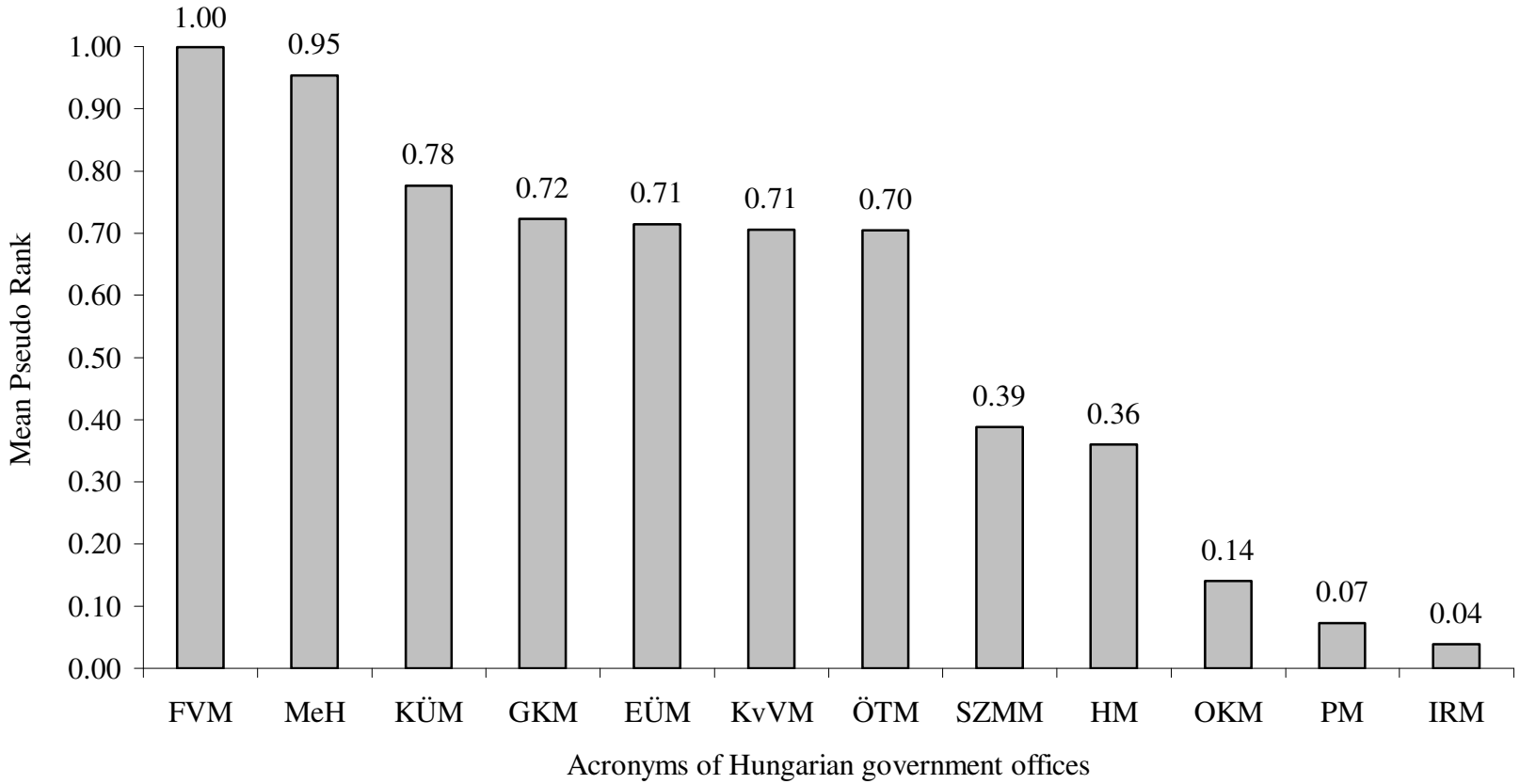

Figure 6.8 Mean Pseudo Rank of the acronyms of Hungarian government offices.

In this application the identification capability of the acronyms of Hungarian government offices were evaluated. The results reflect the Web searching behaviour of Hungarian users.

Table 6.7. Average Pseudo Precision and average Mean Pseudo Rank of the acronyms of Hungarian general institutions and government offices.

Acronyms	Average Pseudo Precision	Average Mean Pseudo Rank
Hungarian institutions	0.61	0.53
Hungarian government offices	0.74	0.55

Table 6.7 compares the identification capability of the Hungarian general institutions and government offices. It can be seen that the average Pseudo Precision of the acronyms of government offices is a little bit greater. However, the average Mean Pseudo Ranks figures agree.

6.6 Measuring the Home Page Identification Capability of the Acronyms of Hungarian Higher Educational Institutions

In this section the MICQ method is applied to the study the usefulness of Hungarian higher educational acronyms (i.e., their ability to identify their own institution on the Web). Two measurements were carried out. In the first measurement, the identification capability of Hungarian higher educational institutions is evaluated using the search engines of *Section 6.4 Table 6.3*[SKROP 3]. The results are compared with the usefulness of general acronyms. The expectation is that the identification capability of the acronyms of higher educational institutions is better than the identification capability of acronyms in general. In the second measurement the identification capability of Hungarian higher educational institutions is evaluated using the search engines of *Section 6.5 Table 6.6*. This experiment reflects the Web searching behaviour of Hungarian users in 2006. The results of *Measurement 2* are compared with the results of *Measurement 1*.

6.6.1 Measurement 1

191 Hungarian higher educational institutions that have acronyms and are present with their own Web site on the Web were identified. The higher educational institutions' home pages were identified using the following Web sites that list higher educational institutions: Egyetemek.lap.hu, Főiskolák.lap.hu, Országos Felsőoktatási Információs Központ. The full list of institution names, acronyms and home page URLs can be found in *Appendix A.2*. This experiment was carried out during March 2004 by entering each acronym to each of the six search engines and evaluating the first ten results according to MICQ. The search engines were the same as in *Section 6.4 Table 6.3.* In this experiment some twelve thousand hits were examined because there were *Category* 2 and *Category* 3 links in the hit list as well. The hit list of a search engine for a query can be checked about in two minutes because only the first

ten hits returned is evaluated and the URL of the correct answer is known. Thus the evaluation of 191 queries in 6 search engines takes 191 x 2 x 6 = 2292 minutes = 38.2 hours.

The results obtained in this application 2 are presented in details in *Appendix B Table B.3*. The Pseudo Precision results of acronyms of general institutions are recalled from *Section 6.4* and are presented here for comparison. However, it is worthy of note that the group of general institutions contains some higher educational institutions. The overlap is 13 %. (The general institution group contains 15 higher educational institutions). This overlap may not influence the comparison. For example, the average Pseudo Precision of general institutions with the higher educational institutions is 0.61 over all search engines and 0.62 omitting higher educational institutions. In this experiment the Pseudo Precision of higher educational institutions' acronyms were measured.

Figure 6.9 shows a Pseudo Precision histogram of acronyms of higher educational institutions for all search engines. It can be seen that the majority of acronyms are not so useful (74% of the acronyms have Pseudo Precision less than 0.5), only 6% are very useful (Pseudo Precision equals 1) and about 38% cannot be judged as being useful (Pseudo Precision equals 0).

Table 6.8 compares the distribution of general and higher educational acronyms over usefulness categories. The usefulness categories are the same as in *Section 6.5*.

Table 6.8 Percentage of general and higher educational acronyms over usefulness categories.

	General acronyms	**Higher educational acronyms**
Not useful	4%	38%
Somewhat useful	32%	36%
Useful	58%	20%
Very useful	6%	6%

Only 6% of the acronyms are very useful in case of both general and higher educational acronyms. This means that only 6% of the desired home pages were found by all the six search engines when the acronym of the institutions was used as query. The proportion of not useful acronyms differs significantly. 38% of the home pages of higher educational institutions could not be found using their acronyms as queries in Web searching. This rate is only 4% in case of general acronyms.

Figure 6.10 shows Pseudo Precision histograms of acronyms of higher educational institutions separately for Hungarian and general search engines. 25% of the acronyms are very useful in general search engines and 10% in Hungarian search engines. The average Pseudo Precision is 0.3 in Hungarian search engines and 0.36 in general search engines. The Pseudo Precision regarding general search engines decreased by 40% relative to the general acronyms.

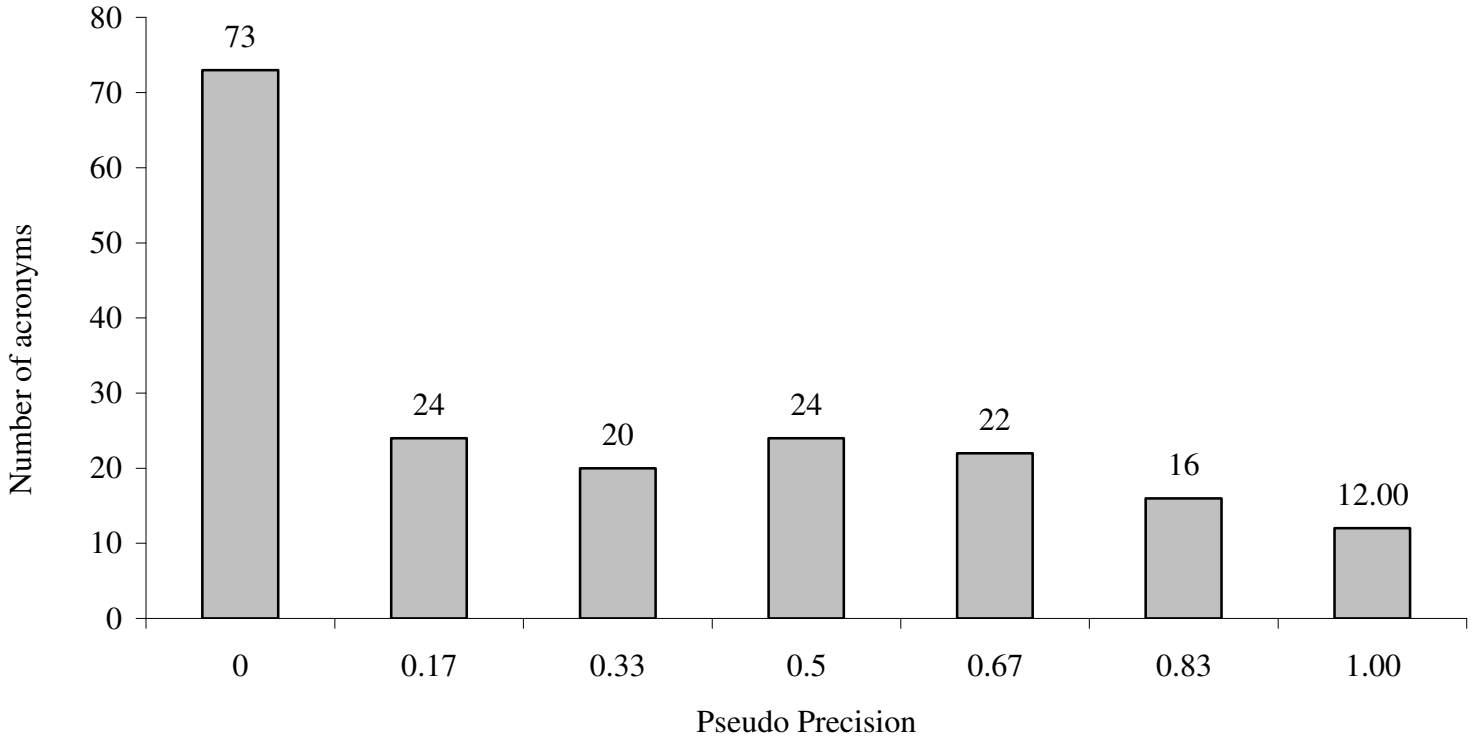

Figure 6.9 Pseudo Precision histogram of acronyms of Hungarian higher educational institutions over all search engines.

On the horizontal axis Pseudo Precision intervals are presented. The intervals represented are half-closed and only the upper endpoints are included.
On the vertical axis, the frequency of Pseudo Precision is shown, i.e., the number of acronyms that fall into the intervals of Pseudo Precision.
74% of acronyms can hardly be judged as being useful (Pseudo Precision is less than 0.5); 26% of the acronyms are useful (Pseudo Precision is greater than 0.5)

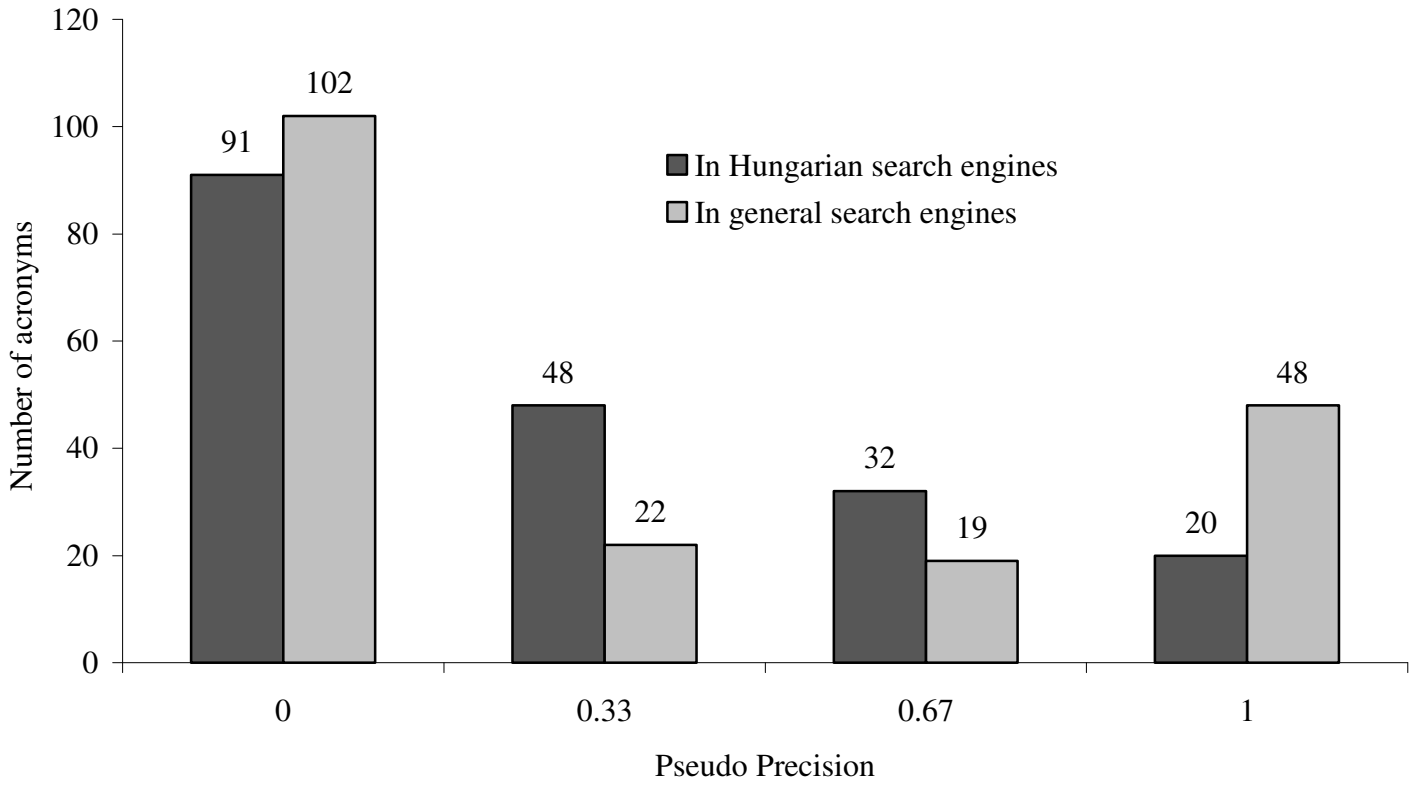

Figure 6.10 Pseudo Precision histogram of acronyms of Hungarian higher educational institutions over Hungarian and general search engines.

On the horizontal axis Pseudo Precision intervals are presented. The intervals represented are half-closed and only the upper endpoints are included.
On the vertical axis, the frequency of Pseudo Precision is shown, i.e., the number of acronyms that fall into the intervals of Pseudo Precision.
65% of acronyms over Hungarian search engines and 56% over general search engines can hardly be judged as being useful (Pseudo Precision is less than 0.5); 35 % of the acronyms over Hungarian search engines and 35 % over general search engines are useful (Pseudo Precision is greater than 0.5).

Table 6.9 compares the distribution of general and higher educational acronyms over the usefulness categories separately for Hungarian and general search engines. It can be seen that both general and higher educational acronyms have better identification capability in general search engines than in Hungarian ones.

Table 6.9 Percentage of general and higher educational acronyms over usefulness categories separately for Hungarian and general search engines.

	General acronyms		**Higher educational acronyms**	
	Hungarian search engines	**General search engines**	**Hungarian search engines**	**General search engines**
Not useful	14%	5%	48%	53%
Somewhat useful	46%	12%	17%	12%
Useful	33%	28%	25%	10%
Very useful	7%	55%	10%	25%

This result is perhaps unexpected. On the one hand, one would have thought that the acronyms of higher educational institutions identify similarly or better the institutions than those of general institutions. It is assumed that at higher educational institutions there are well-skilled staffs (professionals) responsible for the management of the Web site of the institution. Hence, it is also assumed that these home pages have better Web page design. Consequently, these home pages should be easier found by Web search engines when the acronyms of higher educational institutions are used as queries. They are assumed to be more useful than acronyms in general.

On the other hand, one would have thought that the acronyms should identify the institutions well in Hungarian search engines as well or even better. The used Hungarian search engines index Hungarian Web pages, while general search engines index the "whole" Web. The differences in performance of acronyms may be because each individual search engine uses its own unique algorithm to index and rank Web sites, and the algorithms use various factors to rank pages in their search results. Search engines may provide basics of their indexing and page-ranking policies, however, the Hungarian search engines used do not provide the same.

The mean Pseudo Rank results of acronyms of general institutions are recalled from *Section 6.4* and are presented for comparison. The Mean Pseudo Rank histograms of acronyms of higher educational institutions are obtained in this experiment.

Figure 6.11 shows the Mean Pseudo Rank histogram of acronyms of higher educational institutions for all search engines. Table 6.10 shows the number and percentage of general and higher educational acronyms at different *MPR* intervals over all search engines respectively. The intervals represented in Table 6.10 are half-closed and only the upper endpoints are included.

Table 6.10 The number and percentage of general and higher educational acronyms at different *MRR* intervals over all search engines. The intervals represented are half-closed and only the upper endpoints are included.

MPR	Number of general acronyms	Percentage [%]	Number of higher educational acronyms	Percentage [%]
0	4	4	14	7
0 – 0.1	7	6	34	18
0.1 – 0.2	6	5	33	17
0.2 – 0.3	5	4	25	13
0.3 – 0.4	14	12	24	13
0.4 – 0.5	13	12	20	11
0.5 – 0.6	9	8	14	7
0.6 – 0.7	24	21	10	5
0.7 – 0.8	16	14	8	4
0.8 – 0.9	16	14	5	3
0.9 –1	0	0	4	2

It can clearly be seen that, as expected based on Pseudo Precision, the degree of usefulness of about half of the general acronyms is much higher than that of higher educational ones in case of over all search engines. Interestingly enough, none of the home page of general institutions was found and listed in the first position in the hit list by all the six search engines. However, four of the home pages of higher educational institutions were listed as first by all the six search engines.

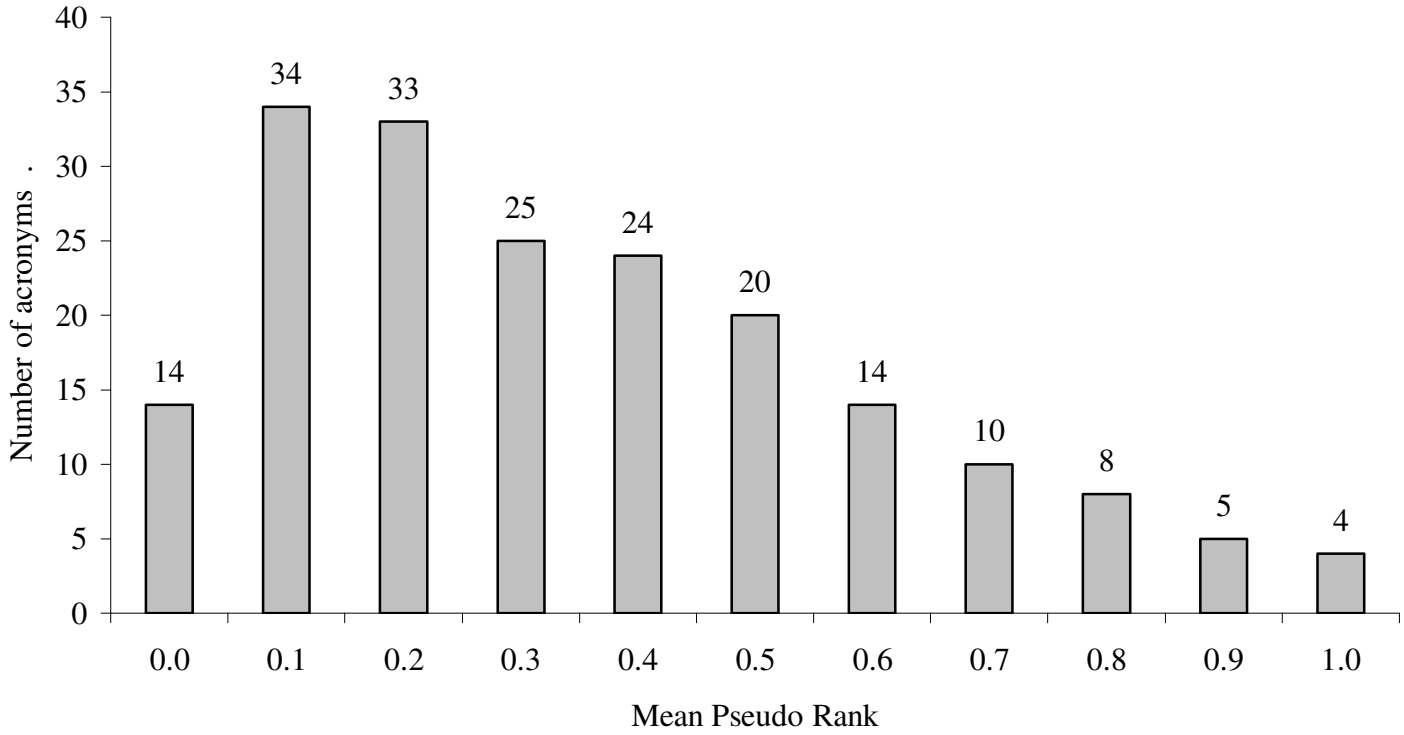

Figure 6.11 Mean Pseudo Rank histogram of acronyms of higher educational institutions over all search engines.

On the horizontal axis Mean Pseudo Rank intervals are presented. The intervals represented are half-closed and only the upper endpoints are included. On the vertical axis, the frequency of Mean Pseudo Rank is shown, i.e., the number of acronyms that fall into the intervals of Mean Pseudo Rank. 79% of acronyms have Mean Pseudo Rank less than 0.5; 21% of the acronyms have Mean Pseudo Rank greater than 0.5.

Figure 6.12 shows the mean Pseudo Rank histogram of acronyms of higher educational institutions for Hungarian search engines. Table 6.11 shows the number and percentage of general and higher educational acronyms at different *MPR* intervals for Hungarian search engines, respectively. The intervals represented in Table 6.11 are half-closed and only the upper endpoints are included.

Table 6.11 The number and percentage of general and higher educational acronyms at different *MPR* intervals for Hungarian search engines. The intervals represented are half-closed and only the upper endpoints are included.

MPR	Number of general acronyms	Percentage [%]	Number of higher educational acronyms	Percentage [%]
0	9	8	18	9
0 – 0.1	10	9	32	17
0.1 – 0.2	11	10	37	19
0.2 – 0.3	15	13	23	12
0.3 – 0.4	22	19	25	13
0.4 – 0.5	13	11	27	14
0.5 – 0.6	9	8	10	5
0.6 – 0.7	16	14	8	4
0.7 – 0.8	4	3	3	2
0.8 – 0.9	2	2	7	4
0.9 –1	3	3	1	1

The results show that about 42% of home pages of general institutions and 31% of home pages of higher educational institutions can easily be found (*MPR* is greater than 0.5) by users using their acronym as Web query in Hungarian search engines. General acronyms have better identification capability than higher educational acronyms in Hungarian search engines.

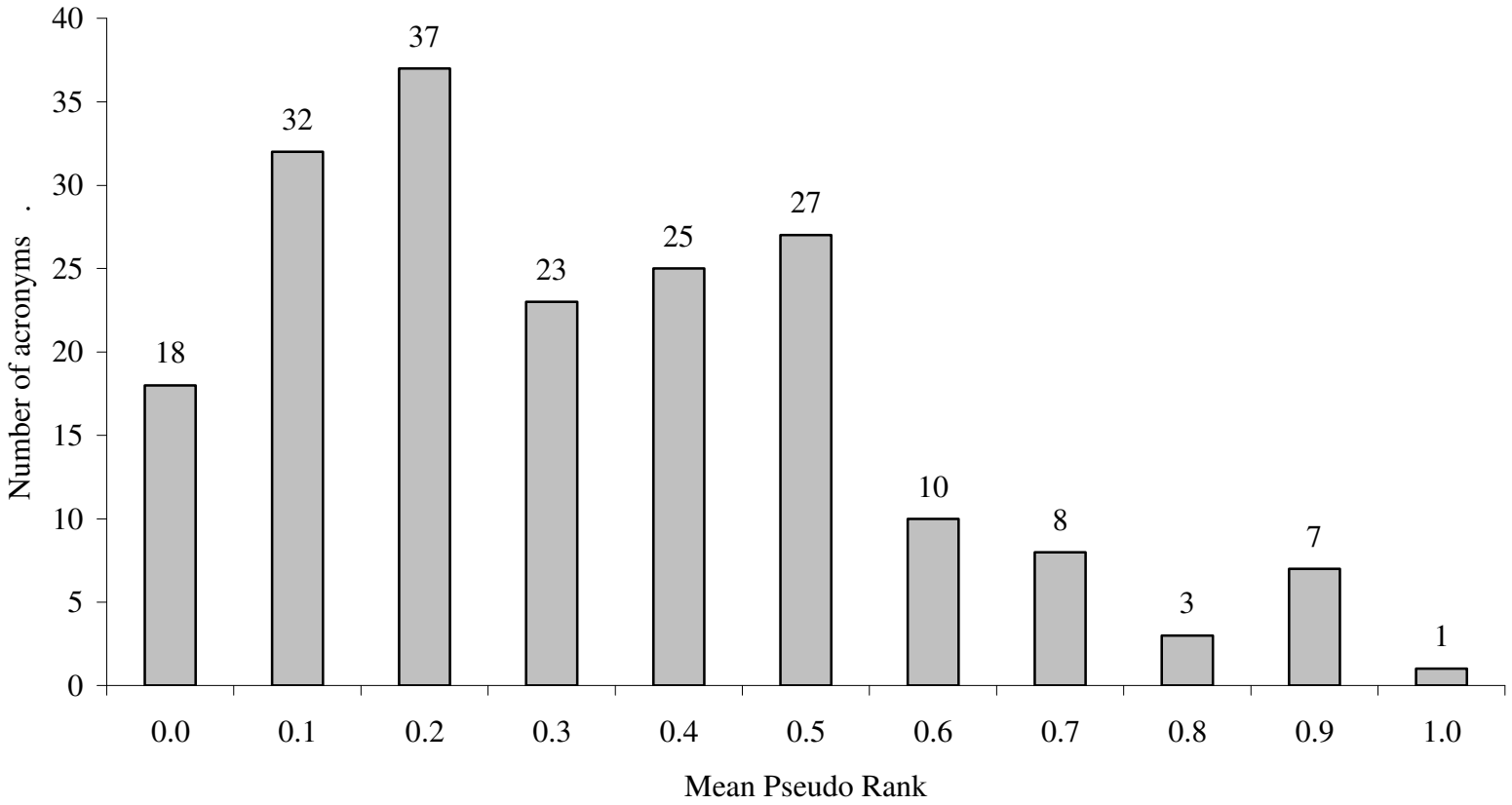

Figure 6.12 Mean Pseudo Rank histogram of acronyms of higher educational institutions over Hungarian search engines.

On the horizontal axis Mean Pseudo Rank intervals are presented. The intervals represented are half-closed and only the upper endpoints are included. On the vertical axis, the frequency of Mean Pseudo Rank is shown, i.e., the number of acronyms that fall into the intervals of Mean Pseudo Rank. 84% of acronyms have Mean Pseudo Rank less than 0.5; 16% of the acronyms have Mean Pseudo Rank greater than 0.5.

Figure 6.13 shows the Mean Pseudo Rank histogram of acronyms of higher educational institutions for general search engines. Table 6.12 shows the number and percentage of general and higher educational acronyms at different *MPR* intervals for general search engines respectively. The intervals represented in Table 6.12 are half-closed and only the upper endpoints are included.

Table 6.12 The number and percentage of general and higher educational acronyms at different *MPR* intervals for general search engines. The intervals represented are half-closed and only the upper endpoints are included.

MPR	Number of general acronyms	Percentage [%]	Number of higher educational acronyms	Percentage [%]
0	5	4	44	23
0 – 0.1	3	3	24	13
0.1 – 0.2	6	5	21	11
0.2 – 0.3	1	1	19	10
0.3 – 0.4	9	8	19	10
0.4 – 0.5	8	7	14	7
0.5 – 0.6	2	2	5	3
0.6 – 0.7	20	18	10	5
0.7 – 0.8	15	13	9	5
0.8 – 0.9	6	5	6	3
0.9 –1	39	34	20	10

The results show that about 79% of home pages of general institutions and only 33% of home pages of higher educational institutions can easily be found (MPR is greater than 0.5) by users using their acronym as Web query by general search engines. General acronyms have better identification capability than higher educational acronyms in general search engines.

It can clearly be seen that, as expected on the basis of Pseudo Precision, the degree of usefulness of about half of the acronyms is low, however it is higher in general search engines than in Hungarian ones and also higher in case of general institutions than in higher educational ones. The desired institution cannot be found in 23% in general search engines and in 10% in Hungarian search engines at all using the acronyms. Only 0.5 % of the home pages of institutions can easily be found by Hungarian search engines and 10% by general search engines.

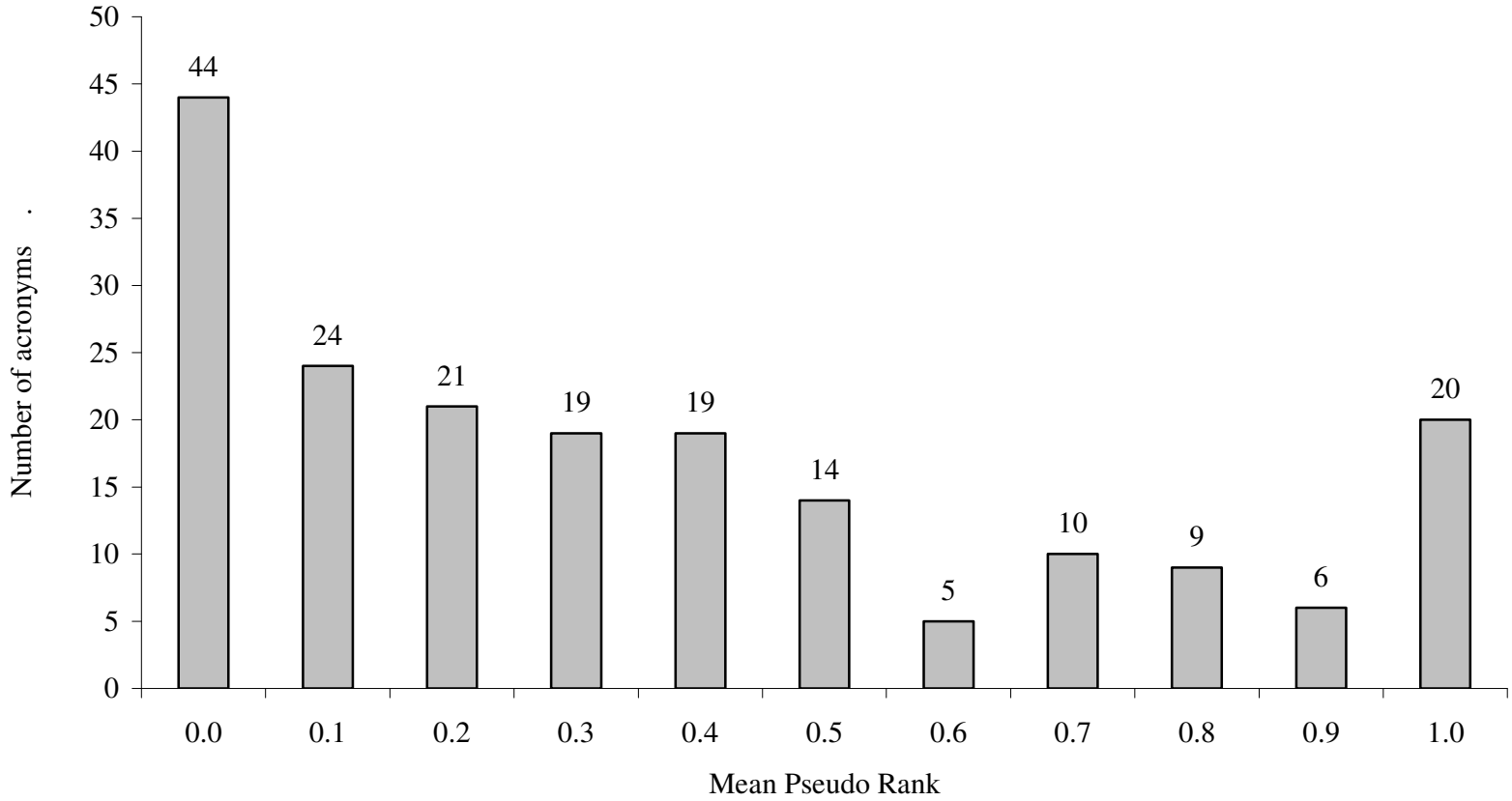

Figure 6.13 Mean Pseudo Rank histogram of acronyms of higher educational institutions over general search engines.

On the horizontal axis Mean Pseudo Rank intervals are presented. The intervals represented are half-closed and only the upper endpoints are included. On the vertical axis, the frequency of Mean Pseudo Rank is shown, i.e., the number of acronyms that fall into the intervals of Mean Pseudo Rank. 74% of acronyms have Mean Pseudo Rank less than 0.5; 26% of the acronyms have Mean Pseudo Rank greater than 0.5.

Average values of the Mean Pseudo Ranks are shown in Table 6.13. It can be seen that the value of Mean Pseudo Rank of higher educational institutions' acronyms is less than that of general acronyms by 23% in all search engines, by 9% in Hungarian search engines and by 35% in general search engines.

Table 6.13 Average Mean Pseudo Ranks.

	Average *MPR*	
	General institutions	**Higher educational institutions**
Over all search engines	0.53	0.3
Over Hungarian search engines	0.38	0.29
Over general search engines	0.68	0.33

This poor identification capability of higher educational acronyms can be seen as a noteworthy situation, this may be due to the following causes.

Poor Web page design of home pages and sites may again be one of the causes. Another cause may be that quite a few acronyms have multiple meanings. It is true when regarding general acronyms. For example, the acronym MNB (which identifies the Hungarian National Bank) also stands for the following institutions:

- Magyar Nemzeti Bibliográfia
- Magyar Nemzeti Bizottság
- Moffat, Nichol & Bonney INC.
- Moody National Bank
- Malvern National Bank

which, due to a much better Web page design, are retrieved more easily (ahead of the Hungarian MNB).

However, most Hungarian higher educational institutions have a unique acronym. This is because the acronyms for e.g. university faculties are composed of the acronym of the university name and of the acronym of the faculty name. Examples are presented in Table 6.14.

Table 6.14 Acronyms of Hungarian higher educational institutions (examples).

University	**Faculty**	**Acronym**
Veszprémi Egyetem	Műszaki Informatikai Kar	VE MIK
Szent István Egyetem	Gépészmérnöki Kar	SZIE GÉK
Szegedi Tudományegyetem	Bölcsészettudományi Kar	SZTE BTK

The usefulness of the acronyms of Hungarian higher educational institutions to identify institutions was evaluated on the Web using Hungarian as well as general search engines. The results show that the majority of the acronyms are not effective in identifying their institutions.

The results are surprising to the effect that they refute the expectation that acronyms of higher educational institutions are more useful than acronyms of general institutions. It was assumed that higher educational institutions have well designed home pages because they usually employ professionals.

6.6.2 Measurement 2

In this measurement, the identification capability of the 191 Hungarian higher educational institutions is evaluated using the search engines of *Section 6.5 Table 6.6*. The results of this experiment are compared with the results of *Measurement 1* (*Section 6.6.1*). This experiment was carried out during September 2006 by entering each acronym to each of the six search engines of *Section 6.5*. The results were evaluated using the MICQ method. Based on the search engine usage statistic the Pseudo Precision and Mean Pseudo Rank values were calculated using the weighting scheme of *Section 6.5*. The weighting scheme takes into account the distribution of search engine usage in Hungary in 2005.

The results obtained in this application are presented in details in *Appendix B Table B.4.*

Figure 6.14 shows the Pseudo Precision histogram of acronyms of Hungarian higher educational institutions in 2006. The histogram obtained on the 2004 sample (*Section 6.6.1*) is used as a comparison.

Table 6.15 compares the distribution of higher educational acronyms over usefulness categories. The usefulness categories are the same as previously.

Table 6.15 Percentage of general and higher educational acronyms over usefulness categories.

	Measurement 1	**Measurement 2**
Not useful	38%	19%
Somewhat useful	36%	13%
Useful	20%	47%
Very useful	6%	22%

In the 2004 measurement only 6% of the acronyms (institutions) can be found effectively with search engines. In the 2006 measurement 22% of the acronyms identify its institution. In Experiment 1 38% of the acronyms do not identify its institution while in Experiment 2 the percentage of acronyms that do not identify its institution is 19%.

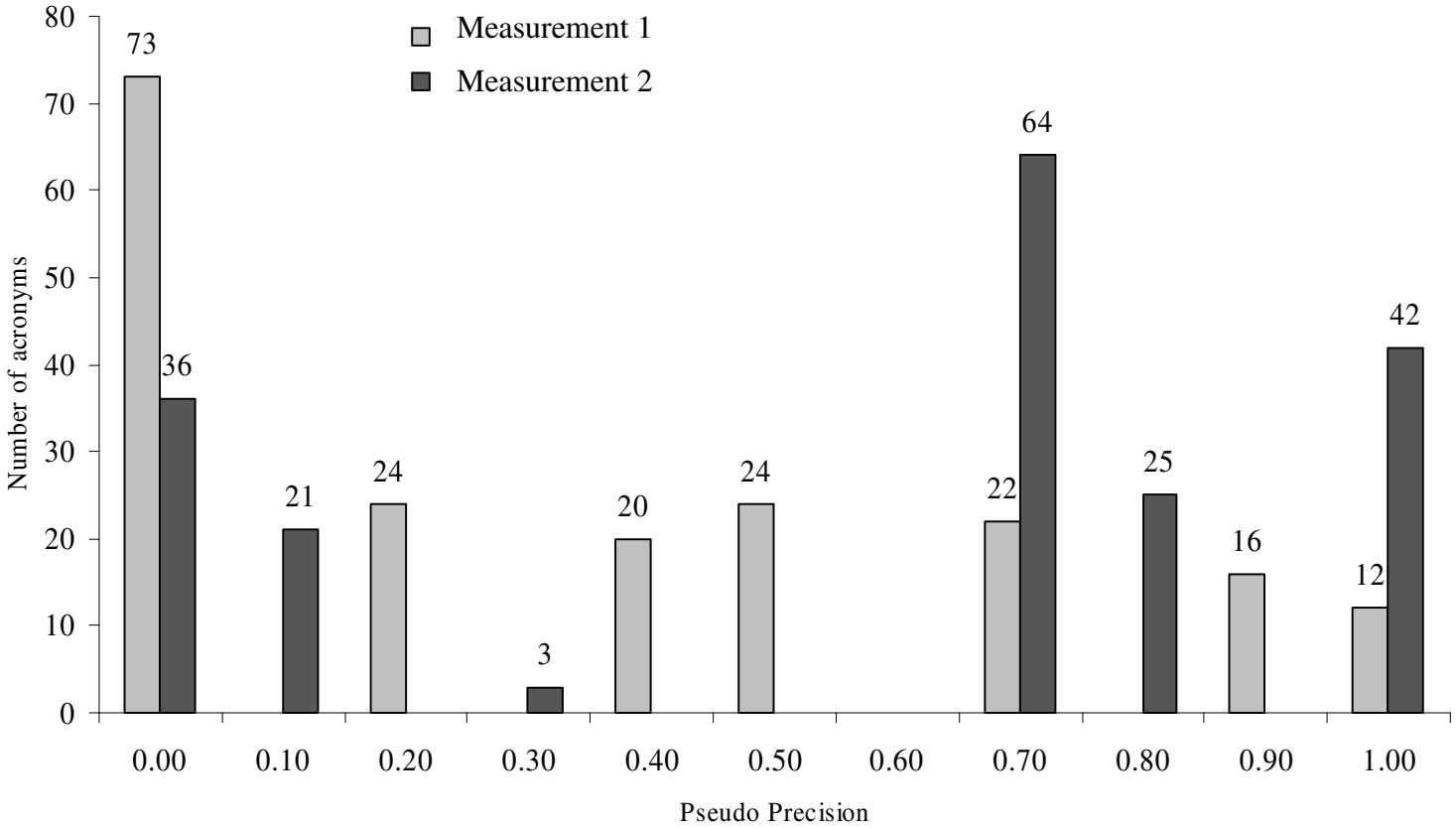

Figure 6.14 Pseudo Precision histogram of acronyms of Hungarian higher educational institutions in 2004 (*Measurement 1*) and in 2006 (*Measurement 2*).

On the horizontal axis Pseudo Precision intervals are presented. The intervals represented are half-closed and only the upper endpoints are included.
On the vertical axis, the frequency of Pseudo Precision is shown, i.e., the number of acronyms that fall into the intervals of Pseudo Precision.
It can be seen that more acronyms have high Pseudo Precision in *Measurement 2* than in *Measurement 1*.

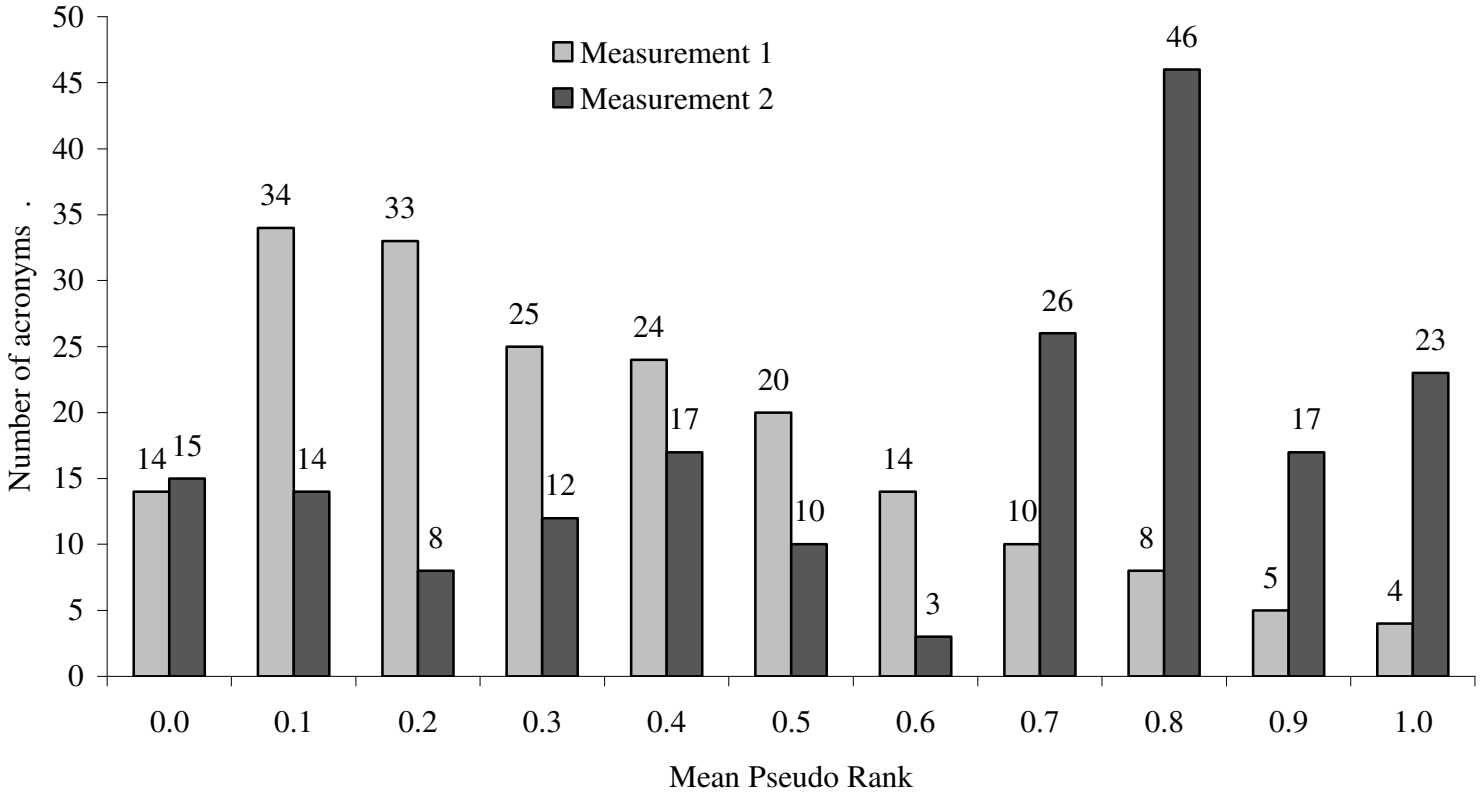

Figure 6.15 Mean Pseudo Rank histogram of acronyms of higher educational institutions in 2004 (*Measurement 1*) and in 2006 (*Measurement 2*).

On the horizontal axis Mean Pseudo Rank intervals are presented. The intervals represented are half-closed and only the upper endpoints are included. On the vertical axis, the frequency of Mean Pseudo Rank is shown, i.e., the number of acronyms that fall into the intervals of Mean Pseudo Rank. It can be seen that more acronyms have high Mean Pseudo Rank in *Measurement 2* than in *Measurement 1*.

Table 6.16 shows the averages of Pseudo Precisions. It can be seen that in *Measurement 2* the Average Pseudo Precision is higher by 22% than in *Measurement 1*.

Table 6.16 Average Pseudo Precision of Hungarian Higher educational institutions.

	Average Pseudo Precision
Measurement 1	0.33
Measurement 2	0.55

Figure 6.15 shows the Mean Pseudo Rank histogram of acronyms of Hungarian higher educational institutions in 2006. The histogram obtained on the 2004 sample (*Section 6.6.1*) is used as a comparison. Table 6.17 shows the number and percentage of higher educational acronyms at different Mean Pseudo Rank intervals.

Table 6.17 The number and percentage of Hungarian higher educational acronyms at different *MPR* intervals. The intervals represented are half-closed and only the upper endpoints are included.

MPR	**Measurement 1**		**Measurement 2**	
	Number of acronyms	**Percentage [%]**	**Number of acronyms**	**Percentage [%]**
0	14	7	15	8
0 – 0.1	34	18	14	7
0.1 – 0.2	33	17	8	4
0.2 – 0.3	25	13	12	6
0.3 – 0.4	24	13	17	9
0.4 – 0.5	20	11	10	5
0.5 – 0.6	14	7	3	2
0.6 – 0.7	10	5	26	14
0.7 – 0.8	8	4	46	24
0.8 – 0.9	5	3	17	9
0.9 –1	4	2	23	12

The results show that in *Measurement 1* about 32% of the acronyms can identify the institutions. In *Measurement 2* about 56% of the acronyms can identify the institutions. Their Mean Pseudo Rank is at least 0.5. In *Measurement 1* there are 14 acronyms that are not able to identify its institution at al. In *Measurement 2* this number is 15. Table 6.18 shows the averages of Mean Pseudo Ranks. It can be seen that in *Measurement 2* the Average Pseudo Precision is higher by 25% than in *Measurement 1*.

Table 6.18 Average Mean Pseudo Rank of Hungarian higher educational institutions.

	Average Mean Pseudo Rank
Measurement 1	0.3
Measurement 2	0.55

In this application the identification capability of Hungarian higher educational institutions were investigated. In *Measurement 1* the most frequently used Hungarian and general (not Hungarian) search engines were used. This experiment is intended to show how users in general can found Hungarian higher educational institutions on the Web. In *Measurement 2* it was evaluated how Hungarian users can find these home pages on the Web. In this experiment the search engines were selected according to the available search engine usage statistics. The results showed that Hungarian users could more effectively find these home pages in 2006 than general ("all over the world") users in 2004.

6.7 Measuring the Home Page Identification Capability of the Acronyms of Danish Higher Educational Institutions

The MICQ method to evaluate the usefulness of acronyms is language independent. It can be used to carry out similar experiments in other countries, as well, where acronyms are being used. Thus, the goal of the present application is to evaluate the usefulness of acronyms of Danish higher educational institutions on the Web, i.e. the ability of the acronyms of Danish higher educational institutions' names to find the home page of their own institutions when being used as queries in Web searching.

Two measurements were carried out. In the first measurement, the identification capability of the acronyms of Danish higher educational institutions was evaluated in January 2006. In the second measurement, the identification capability of the acronyms of Danish higher educational institutions was evaluated in September 2006. In the two experiment the search engines were different. In the first experiment Danish and general search engines were used. In the second experiment the search engines were selected according to a Danish search engine usage statistic. The results of the measurements are compared. The results of *Measurement 1* are compared with the results of *Section 6.6*. The usefulness of Danish higher educational acronyms is compared with the usefulness of Hungarian higher educational acronyms. One would expect that Danish and Hungarian acronyms are equally useful. The results of *Measurement 1* are also compared with the results of *Measurement 2*.

6.7.1 Measurement 1

29 higher educational institutions in Denmark that have acronyms and are present with their own Web site on the Web were identified. The Danish higher educational Web sites were identified using the following sites: Braintrack University Index and Norden. Both sites list universities in Denmark. After visiting each site a list was compiled manually containing the full name, home page URL and the acronym of each institution. The full list is not included in this section; Table 6.19 shows a fraction of it. The full list can be found in *Appendix A.3*.

Table 6.19 Full name, home page URL and acronym of higher educational institutions in Denmark.

Full Name	Home Page URL	Acronym
Danmarks Farmaceutiske Universitet	http://www.dfuni.dk/	DFU
Danmarks Journalisthøjskole	http://www.djh.dk/	DJH
IT-Universitetet i København	http://www1.itu.dk/	ITU
Aalborg Universitet	http://ekstern.aau.dk/	AAU

Seven Web search engines were used to evaluate the usefulness of the Danish higher educational acronyms. Table 6.20 presents the selected search engines.

Table 6.20 List of selected search engines: the first three are Danish search engines; the next four are general search engines.

Name of the Search Engine	URL of the Search Engine
soegning	http://www.soegning.dk/
Jubii	http://www.jubii.dk/
OFIR	http://ofir.dk/
I^2R Meta	http://www.dcs.vein.hu/CIR/i2rmeta/i2rmeta.cgi
Google	http://www.google.com
Metacrawler	http://www.metacrawler.com
AltaVista	http://www.altavista.com

The first three search engines are Danish. They primarily index and search Danish Web pages. They were selected according to Web sites that list Danish search engines. These sites are as follows: Danish Search Engines and Directories, Search Engines Worldwide, Search Engine Colossus / Denmark, Denmark Internet Usage Stats and Danish Search Engine. The other four are general search engines. This experiment is intended to evaluate how users (all over the world) can find Danish higher educational home pages using the institution's acronym as query. The experiments (searches) were carried out during January 2006 by entering each acronym to each of the seven search engines.

The results obtained in this application are presented in details in *Appendix B Table B.5*.

Figure 6.16 shows a Pseudo Precision histogram for Danish higher educational acronyms for all search engines. It can be seen that 10% of the acronyms are very useful (Pseudo Precision equals 1), 3% cannot be judged as being useful (Pseudo Precision equals 0), 42 % of the acronyms are not so useful (Pseudo Precision is less than 0.5) and 45% of the acronyms are useful.

Table 6.21 compares the distribution of Danish and Hungarian higher educational acronyms over usefulness categories.

Table 6.21 Percentage of Danish and Hungarian higher educational acronyms over usefulness categories.

	Danish acronyms	**Hungarian acronyms**
Not useful	3%	38%
Somewhat useful	42%	36%
Useful	45%	20%
Very useful	10%	6%

The results show that Danish acronyms have better identification capability than Hungarian acronyms considering Pseudo Precision over all search engines. 55% of Danish and 26% of Hungarian acronyms can be considered as useful.

Figure 6.17 shows Pseudo Precision histograms separately for Danish and general search engines. It can be seen that the majority of the acronyms have better identification capabilities in Danish search engines than in general ones. While the average Pseudo Precision is 0.74 in Danish search engines, it is much smaller, 0.44, in general search engines.

Table 6.22 compares the distribution of Danish and Hungarian higher educational acronyms over usefulness categories separately for Hungarian, Danish and general search engines, respectively.

Table 6.22 Percentage of Danish and Hungarian higher educational acronyms over usefulness categories separately for Hungarian or Danish and general search engines.

	Danish acronyms		**Hungarian acronyms**	
	Danish search engines	**General search engines**	**Hungarian search engines**	**General search engines**
Not useful	4%	34%	48%	53%
Somewhat useful	14%	28%	17%	12%
Useful	41%	14%	25%	10%
Very useful	41%	24%	10%	25%

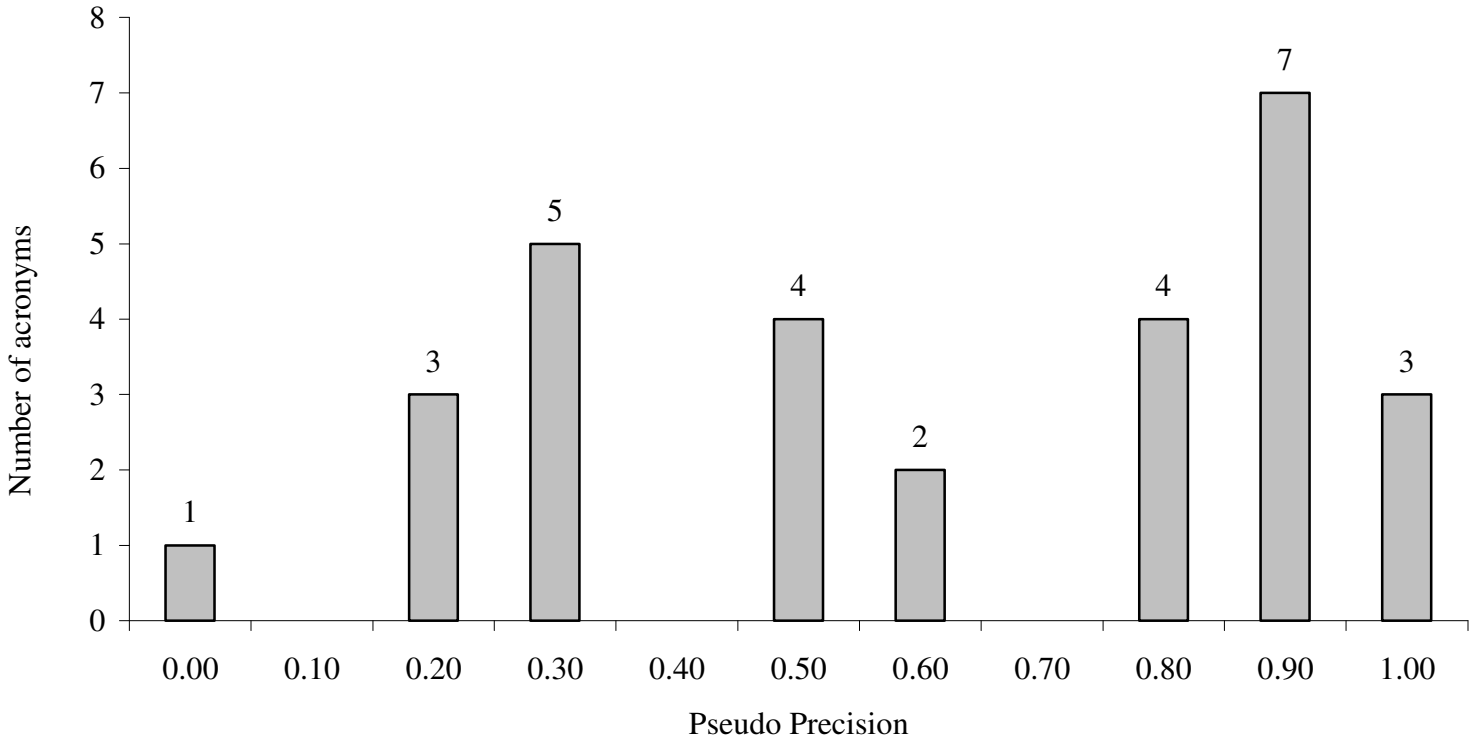

Figure 6.16 Pseudo Precision histogram of acronyms of Danish higher educational institutions over all search engines.

On the horizontal axis Pseudo Precision intervals are presented. The intervals represented are half-closed and only the upper endpoints are included.
On the vertical axis, the frequency of Pseudo Precision is shown, i.e., the number of acronyms that fall into the intervals of Pseudo Precision.
45% of acronyms can hardly be judged as being useful (Pseudo Precision is less than 0.5); 45% of the acronyms are useful (Pseudo Precision is greater than 0.5).

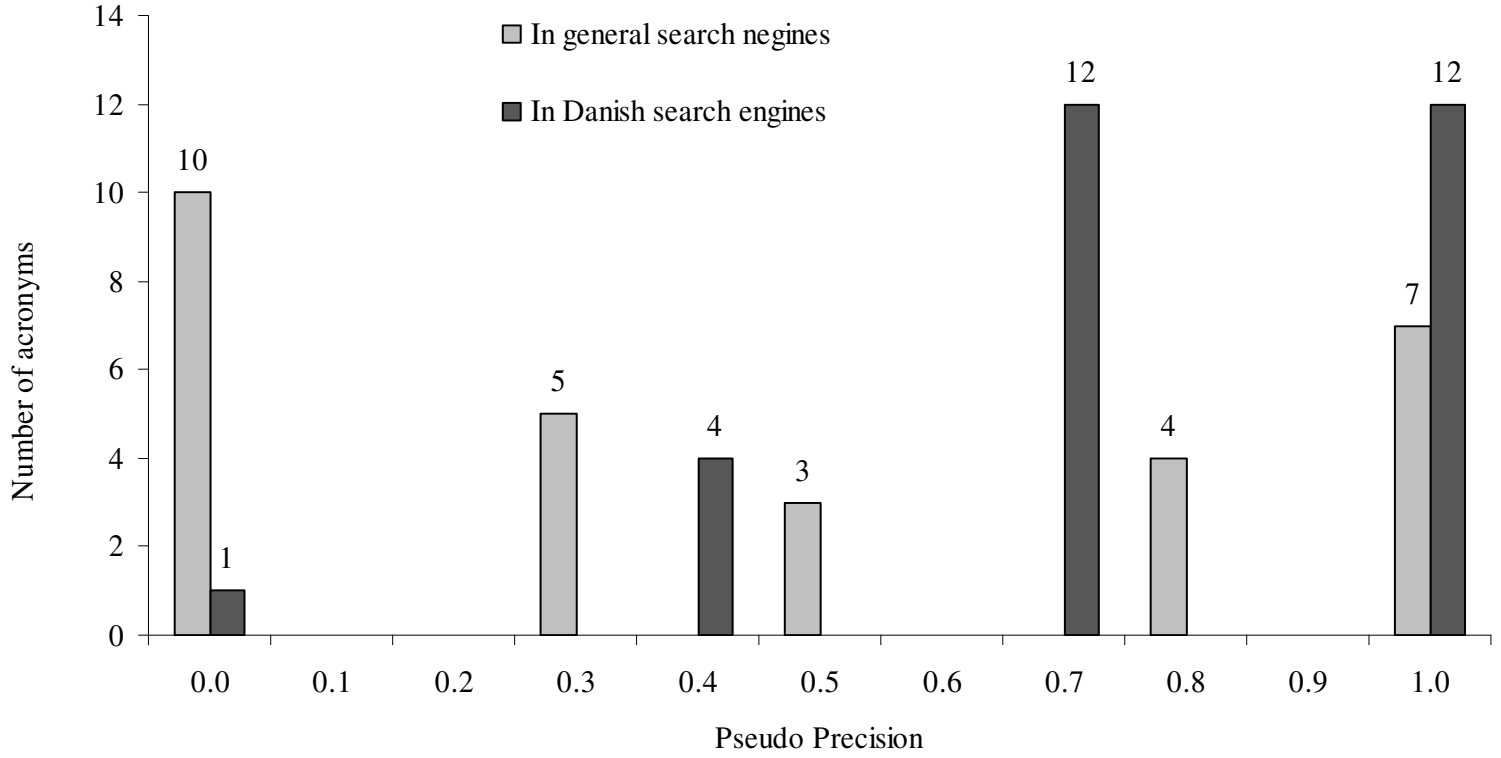

Figure 6.17 Pseudo Precision histogram of acronyms of Danish higher educational institutions over Danish and general search engines.

On the horizontal axis Pseudo Precision intervals are presented. The intervals represented are half-closed and only the upper endpoints are included.
On the vertical axis, the frequency of Pseudo Precision is shown, i.e., the number of acronyms that fall into the intervals of Pseudo Precision.
18% of acronyms over Danish search engines and 62% over general search engines can hardly be judged as being useful (Pseudo Precision is less than 0.5); 82 % of the acronyms over Danish search engines and 38 % over general search engines are useful (Pseudo Precision is greater than 0.5).

The results show that Danish acronyms have better identification capability than Hungarian acronyms, especially in case of Danish search engines. 82% of Danish acronyms are useful (Pseudo Precision is greater than 0.5) in Danish search engines and 38% in general search engines. 35–35% of Hungarian acronyms are useful in Hungarian and general search engines, respectively.

Figures 6.18, 6.19 and 6.20 show the Mean Pseudo Rank histograms of Danish higher educational acronyms for all search engines, for Danish search engines, and for general search engines, respectively. Average values of the Mean Pseudo Ranks are shown in Table 6.23. It can be seen that acronyms in Danish search engines outperform the acronyms in general search engines by 30%.

Table 6.23. Average Mean Pseudo Ranks of Danish higher educational acronyms.

	Average *MPR*
Over all search engines	0.36
Over Danish search engines	0.53
Over general search engines	0.23

Table 6.24 shows the number and percentage of Danish and Hungarian higher educational acronyms at different *MPR* intervals over all search engines, respectively.

Table 6.24 The number and percentage of Danish and Hungarian higher educational acronyms at different *MRR* intervals over all search engines. The intervals represented are half-closed and only the upper endpoints are included.

MPR	**Number of Danish higher educational acronyms**	**Percentage [%]**	**Number of Hungarian higher educational acronyms**	**Percentage [%]**
0	0	0	14	7
0 – 0.1	5	17	34	18
0.1 – 0.2	3	11	33	17
0.2 – 0.3	4	14	25	13
0.3 – 0.4	8	28	24	13
0.4 – 0.5	2	7	20	11
0.5 – 0.6	2	7	14	7
0.6 – 0.7	2	7	10	5
0.7 – 0.8	2	7	8	4
0.8 – 0.9	1	4	5	3
0.9 –1	0	0	4	2

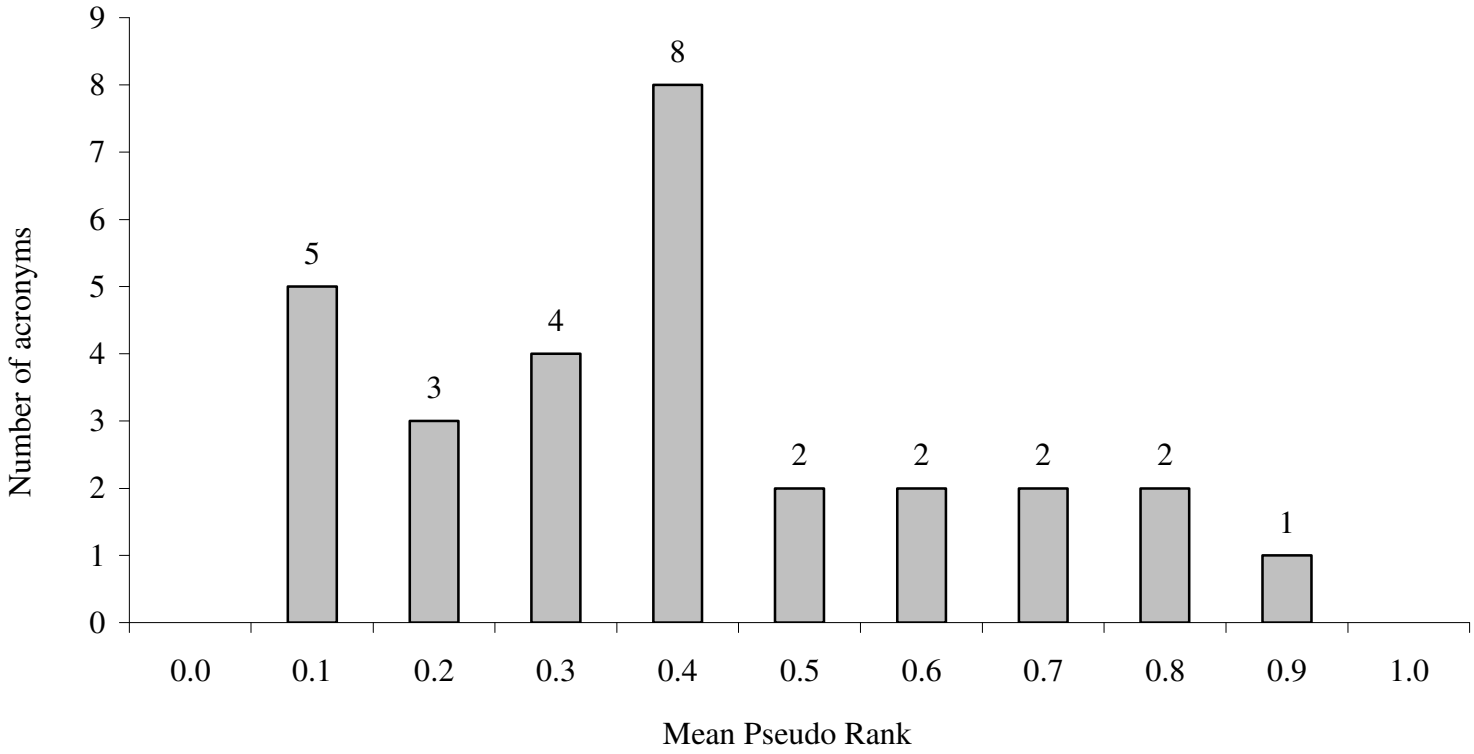

Figure 6.18 Mean Pseudo Rank histogram of acronyms of Danish higher educational institutions over all search engines.

On the horizontal axis Mean Pseudo Rank intervals are presented. The intervals represented are half-closed and only the upper endpoints are included. On the vertical axis, the frequency of Mean Pseudo Rank is shown, i.e., the number of acronyms that fall into the intervals of Mean Pseudo Rank. 68% of acronyms have Mean Pseudo Rank less than 0.5; 32% of the acronyms have Mean Pseudo Rank greater than 0.5.

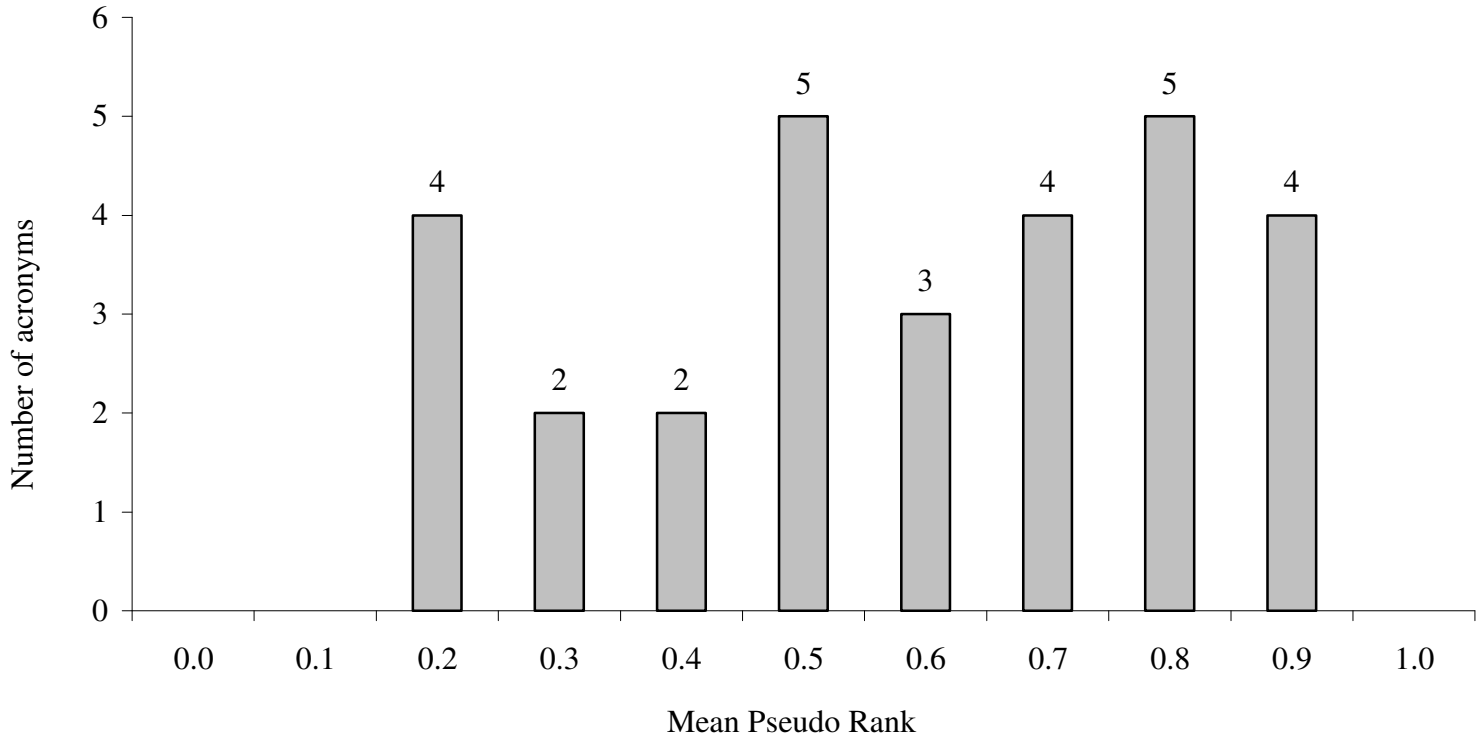

Figure 6.19 Mean Pseudo Rank histogram of acronyms of Danish higher educational institutions over Danish search engines.

On the horizontal axis Mean Pseudo Rank intervals are presented. The intervals represented are half-closed and only the upper endpoints are included. On the vertical axis, the frequency of Mean Pseudo Rank is shown, i.e., the number of acronyms that fall into the intervals of Mean Pseudo Rank. 28% of acronyms have Mean Pseudo Rank less than 0.5; 72 % of the acronyms have Mean Pseudo Rank greater than 0.5.

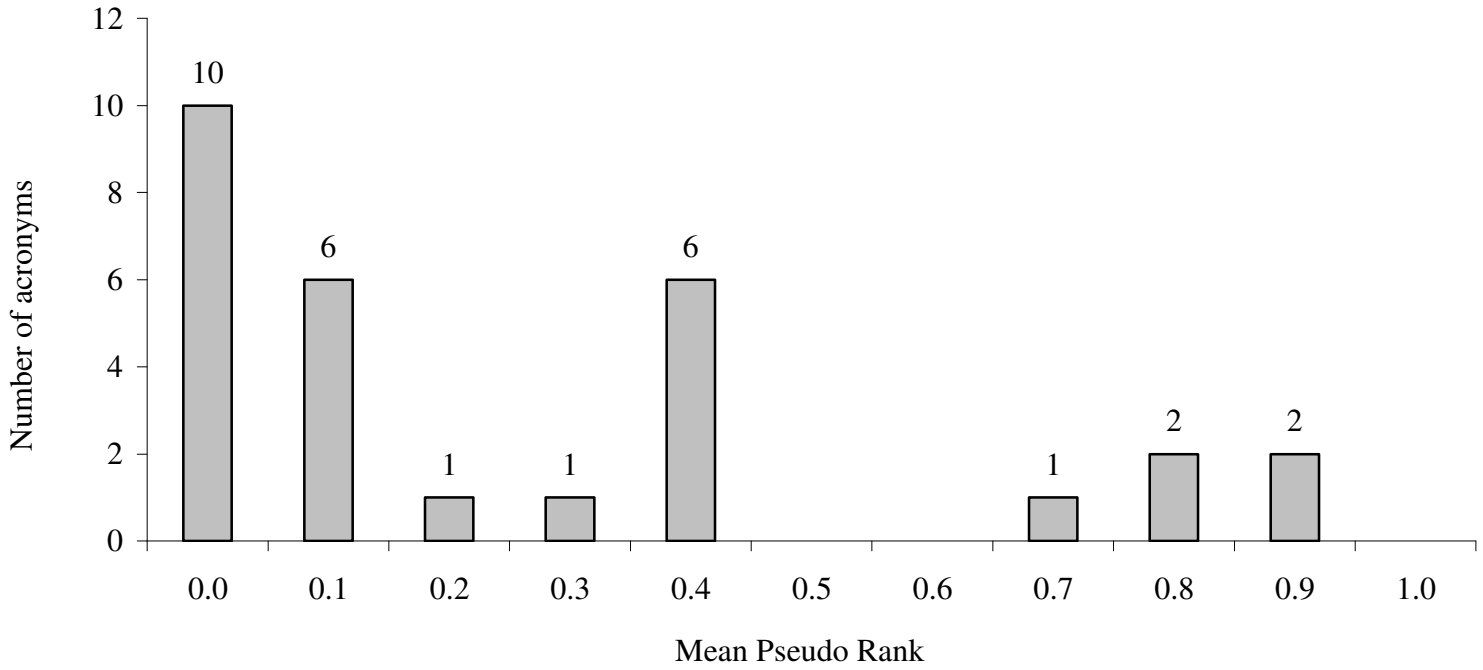

Figure 6.20 Mean Pseudo Rank histogram of acronyms of Danish higher educational institutions over general search engines.

On the horizontal axis Mean Pseudo Rank intervals are presented. The intervals represented are half-closed and only the upper endpoints are included. On the vertical axis, the frequency of Mean Pseudo Rank is shown, i.e., the number of acronyms that fall into the intervals of Mean Pseudo Rank. 82% of acronyms have Mean Pseudo Rank less than 0.5; 17% of the acronyms have Mean Pseudo Rank greater than 0.5.

The results show that about 32% of home pages of Danish higher educational institutions and 32% of home pages of Hungarian higher educational institutions can easily be found (*MPR* is greater than 0.5) by users using their acronym as Web query in Danish and Hungarian search engines, respectively. Danish acronyms have similar identification capability as Hungarian acronyms over all search engines. Interestingly enough, none of the home pages of Danish higher educational institutions was found and listed in the first position by all the seven search engines.

Table 6.25 shows the number and percentage of Danish and Hungarian higher educational acronyms at different *MPR* intervals for Danish and Hungarian search engines, respectively.

Table 6.25 The number and percentage of Danish and Hungarian higher educational acronyms at different *MPR* intervals for Danish and Hungarian search engines. The intervals represented are half-closed and only the upper endpoints are included.

MPR	Number of Danish higher educational acronyms	Percentage [%]	Number of Hungarian higher educational acronyms	Percentage [%]
0	0	0	18	9
0 – 0.1	0	0	32	17
0.1 – 0.2	4	14	37	19
0.2 – 0.3	2	7	23	12
0.3 – 0.4	2	7	25	13
0.4 – 0.5	5	17	27	14
0.5 – 0.6	3	10	10	5
0.6 – 0.7	4	14	8	4
0.7 – 0.8	5	17	3	2
0.8 – 0.9	4	14	7	4
0.9 –1	0	0	1	1

The results show that about 72% of home pages of Danish higher educational institutions and 30% of home pages of Hungarian higher educational institutions can easily be found (*MPR* is greater than 0.5) by users using their acronym as Web query in Danish and Hungarian search engines, respectively. Danish acronyms have better identification capability in Danish search engines than Hungarian acronyms in Hungarian search engines. Interestingly enough, again none of the home pages of Danish higher educational institutions was found and listed in the first position by the Danish search engines.

Table 6.26 shows the number and percentage of Danish and Hungarian higher educational acronyms at different *MPR* intervals for general search engines, respectively.

Table 6.26 The number and percentage of Danish and Hungarian higher educational acronyms at different *MPR* intervals for general search engines.

MPR	Number of Danish higher educational acronyms	Percentage [%]	Number of Hungarian higher educational acronyms	Percentage [%]
0	10	35	44	23
0 – 0.1	6	21	24	13
0.1 – 0.2	1	3	21	11
0.2 – 0.3	1	3	19	10
0.3 – 0.4	6	21	19	10
0.4 – 0.5	0	0	14	7
0.5 – 0.6	0	0	5	3
0.6 – 0.7	1	3	10	5
0.7 – 0.8	2	7	9	5
0.8 – 0.9	2	7	6	3
0.9 –1	0	0	20	10

The results show that about 17% of home pages of Danish higher educational institutions and 33% of home pages of Hungarian higher educational institutions can easily be found (MPR is greater than 0.5) by users using their acronym as Web query by general search engines. It is interesting, that Hungarian higher educational acronyms have better identification capability than Danish higher educational acronyms in general search engines. Interestingly enough, again, none of the home pages of Danish higher educational institutions was found and listed in the first position by the general search engines. Average values of the Mean Pseudo Ranks of Danish and Hungarian higher educational institutions are shown in Table 6.27.

Table 6.27 Average Mean Pseudo Rank of Hungarian and Danish acronyms.

	Average *MPR*	
	Danish Higher educational institutions	**Hungarian Higher educational institutions**
All search engines	0.36	0.3
Danish / Hungarian search engines	0.53	0.29
General search engines	0.36	0.33

It can be seen that Danish higher educational acronyms have better identification capability than Hungarian higher educational acronyms. It is interesting that they are nearly equally useful over all search engines. However, Danish higher educational

acronyms perform better in Danish search engines and worse in general search engines while Hungarian higher educational acronyms perform better in general and worse in Hungarian search engines.

The usefulness of the acronyms of Danish higher educational institutions to identify institutions was evaluated on the Web using Danish as well as general search engines. The usefulness of Hungarian higher educational acronyms was used as a comparison. The results show that the majority of the acronyms are not effective in identifying their institutions.

The results partially support the expectation that acronyms of Danish and Hungarian higher educational institutions are equally useful. Considering Mean Pseudo Rank, both groups of acronyms have similar identification capability over all search engines and general search engines. Danish higher educational acronyms have better identification capability than Hungarian ones when country specific search engines are used. This means that Danish users can the most effectively find the desired Danish home page when the acronym is used as query. Furthermore, Hungarian users can find the least effectively the desired Hungarian home page using Hungarian search engines. However, in the best case (MPR of Danish higher educational acronyms in Danish search engines) the value of Mean Pseudo rank is only 0.53.

6.7.2 Measurement 2

In this experiment the identification capability of the 29 Danish higher educational institutions is evaluated. It is evaluated how Danish users can find the home page of Danish higher educational institutions when the acronyms of the institutions are used as queries. The results are compared with the results of *Measurement 1*. In *Measurement 1* it was investigated how general users (all over the world) can find the Danish higher educational home pages. In this measurement, six Web search engines were used to evaluate the usefulness of the acronyms. Table 6.28 presents the selected search engines.

Table 6.28. List of selected search engines.

Name of the Search Engine	URL of the Search Engine
Google.dk	http://www.google.dk/
Google.com	http://www.google.com/
Eniro	http://www.eniro.dk/
MSN.dk	http://dk.msn.com/
Jubii	http://www.jubii.dk/
Yahoo.dk	http://dk.yahoo.com/

The search engines were selected according to the Search Engine Landscape in Denmark. This landscape shows the distribution of search engines used by Danish users. The experiment was carried out during September 2006 by entering each acronym to each of the six search engines, and evaluating the first ten hits using the

MICQ method. Based on the search engine landscape the Pseudo Precision and Mean Pseudo Rank values were calculated using a weighting scheme. The weighting scheme takes into account the distribution of search engine usage. The weighting scheme was as follows:

Google.dk	$w_1 = 0.67$
Google.com	$w_2 = 0.11$
Eniro	$w_3 = 0.09$
MSN.dk	$w_4 = 0.07$
Jubii	$w_5 = 0.04$
Yahoo.dk	$w_6 = 0.02$

The weighting scheme shows that 67% of Danish users are using Google.dk, 11% are using Google.com, etc. The Pseudo Precision and meanPseudo Rank measures were calculated using *Equation 6.1* and *Equation 6.2*.

The results obtained in this application are presented in details in *Appendix B Table B.6*.

Figure 6.21 shows the Pseudo Precision histograms of Danish higher educational acronyms in *Measurement 1* and *Measurement 2*. In *Measurement 2* 59% of the acronyms are very useful (Pseudo Precision equals 1), 3% cannot be judged as being useful (Pseudo Precision equals 0), 6 % of the acronyms are not so useful (Pseudo Precision is less than 0.5) and 32% of the acronyms are useful.

Table 6.29 compares the distribution of Danish higher educational acronyms over usefulness categories.

Table 6.29 Percentage of Danish higher educational acronyms over usefulness categories.

	Measurement 1	**Measurement 2**
Not useful	3%	3%
Somewhat useful	41%	6%
Useful	46%	32%
Very useful	10%	59%

It can be seen that 3% of the acronyms are not useful – cannot identify its institution – in *Measurement 1* and in *Measurement 2*. In *Measurement 2* 59% of the acronyms can effectively identify its institution, while only 10% in *Measurement 1*. Table 6.30 shows the averages of Pseudo Precisions. It can be seen that in *Measurement 2* the Average Pseudo Precision is higher by 29% than in *Measurement 1*.

Table 6.30 Average Pseudo Precision of Hungarian Higher educational institutions.

	Average Pseudo Precision
Measurement 1	0.57
Measurement 2	0.86

Figure 6.22 shows the Mean Pseudo Rank histogram of acronyms of Danish higher educational institutions in September 2006 (*Measurement 2*). The histogram obtained on the January 2006 sample (*Measurement 1*) is used as a comparison.

Table 6.31 shows the number and percentage of higher educational acronyms at different Mean Pseudo Rank intervals.

Table 6.31 The number and percentage of Danish higher educational acronyms at different *MRR* intervals. The intervals represented are half-closed and only the upper endpoints are included.

MPR	**Measurement 1**		**Measurement 2**	
	Number of Danish higher educational acronyms	**Percentage [%]**	**Number of Danish higher educational acronyms**	**Percentage [%]**
0	0	0	1	4
0 – 0.1	5	16	1	4
0.1 – 0.2	3	10	1	4
0.2 – 0.3	4	14	1	3
0.3 – 0.4	8	28	1	3
0.4 – 0.5	2	7	1	3
0.5 – 0.6	2	7	0	0
0.6 – 0.7	2	7	0	0
0.7 – 0.8	2	7	3	10
0.8 – 0.9	1	4	15	52
0.9 –1	0	0	5	17

In *Measurement 1* none of the acronyms (institutions) can be found effectively with search engines. In *Measurement 2* 17% of the acronyms identify its institution. In *Measurement 1* there is no acronym that does not identify its institution and in *Measurement 2* there is only one acronym that can not identify its institution. The results show that in *Measurement 1* about 32% and in *Measurement 2* about 82 % of the acronyms can identify its institution. Their Mean Pseudo Rank is at least 0.5.

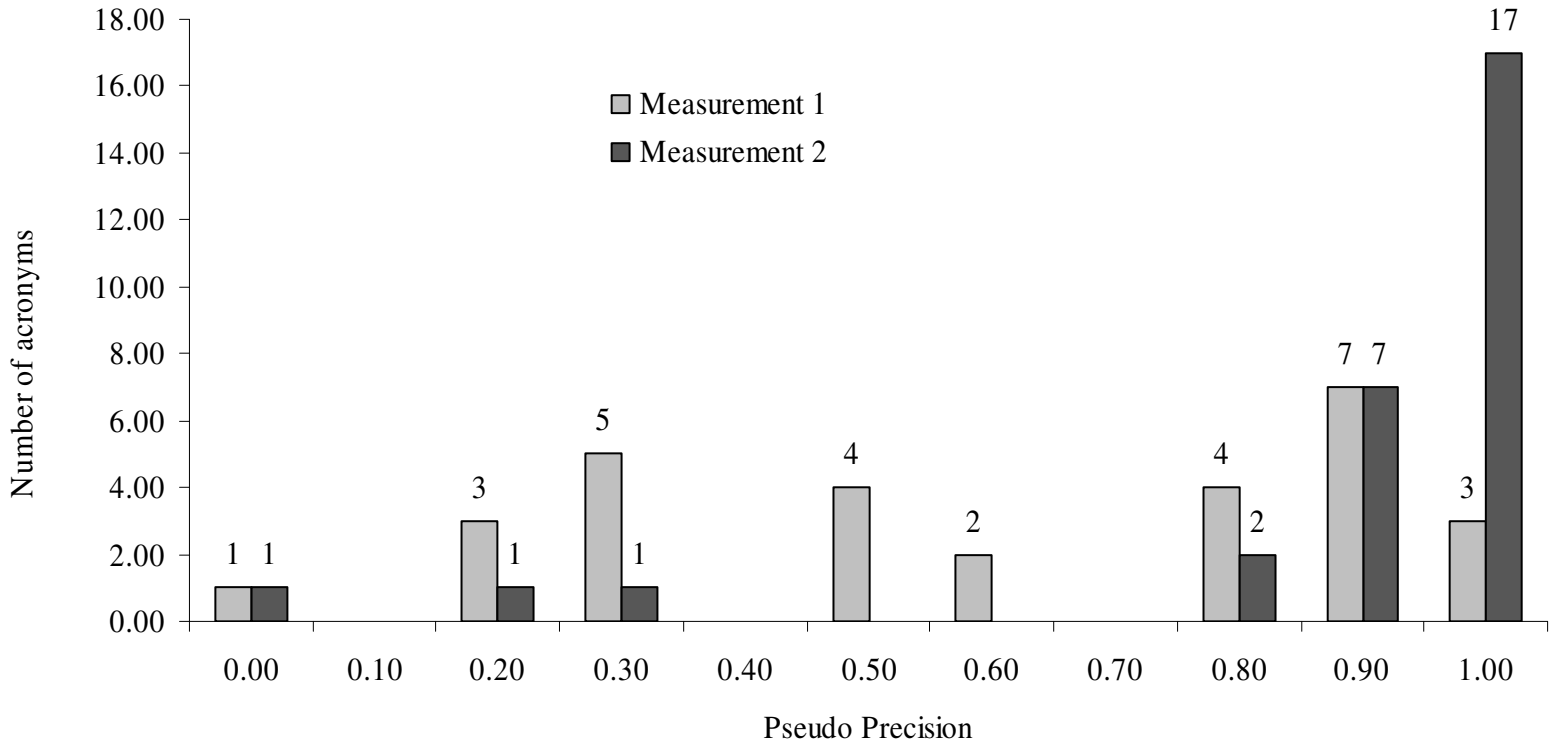

Figure 6.21 Pseudo Precision histograms of acronyms of Danish higher educational institutions in January 2006 (*Measurement 1*) and in September 2006 (*Measurement 2*).

On the horizontal axis Pseudo Precision intervals are presented. The intervals represented are half-closed and only the upper endpoints are included. On the vertical axis, the frequency of Pseudo Precision is shown, i.e., the number of acronyms that fall into the intervals of Pseudo Precision. It can be seen that more acronyms have high Pseudo Precision in *Measurement 2* than in *Measurement 1*.

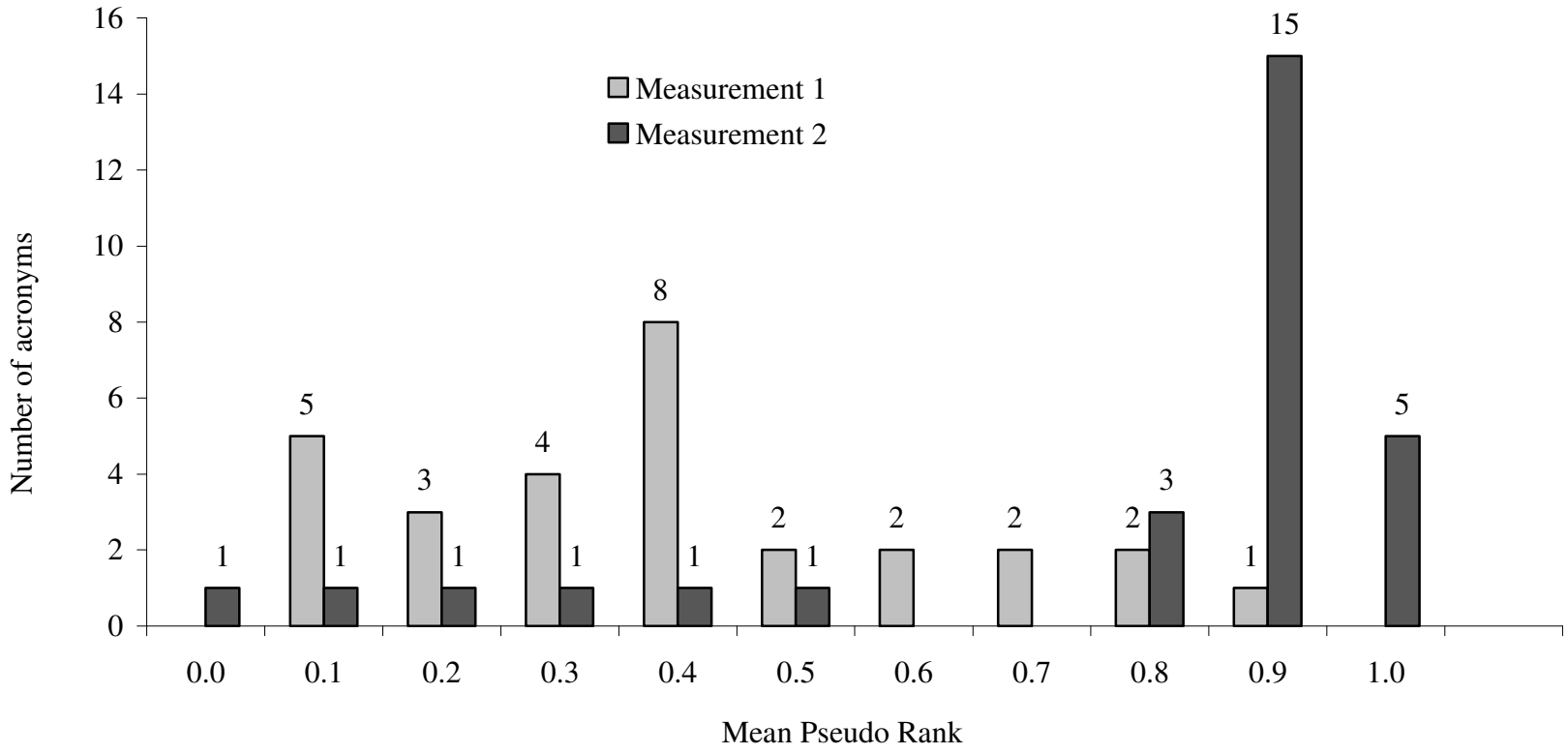

Figure 6.22 Mean Pseudo Rank histograms of acronyms of Danish higher educational institutions in January 2006 (*Measurement 1*) and in September 2006 (*Measurement 2*).

On the horizontal axis Mean Pseudo Rank intervals are presented. The intervals represented are half-closed and only the upper endpoints are included. On the vertical axis, the frequency of Mean Pseudo Rank is shown, i.e., the number of acronyms that fall into the intervals of Mean Pseudo Rank. It can be seen that more acronyms have high Mean Pseudo Rank in *Measurement 2* than in *Measurement 1*.

Table 6.32 shows the averages of Mean Pseudo Ranks. It can be seen that in Experiment 2 the Average Pseudo Precision is higher by 38% than in Experiment 1.

Table 6.32 Average Mean Pseudo Rank of Danish higher educational institutions.

	Average Mean Pseudo Rank
Measurement 1	0.36
Measurement 2	0.74

In this application the identification capability of Danish higher educational institutions were investigated. In *Measurement 1* the most frequently used Danish and general (not Danish) search engines were used. This experiment was intended to show how general Web users (all over the world) could find Danish higher educational institutions on the Web. In *Measurement 2* it was evaluated how Danish users could find these home pages on the Web. In this experiment the search engines were selected according to the available Danish search engine usage statistics. The results showed that Danish users could more effectively find these home pages.

6.8 Measuring the Home Page Identification Capability of the Acronyms of Hungarian and Danish Parties

In this section the MICQ method is applied to the study and to compare the usefulness of acronyms of Hungarian and Danish parties. The identification capability of Hungarian parties' acronyms is evaluated from the viewpoint of Hungarian users. The identification capability of Danish acronyms is evaluated from Danish users' viewpoint

15 Hungarian and 8 Danish parties that have acronyms and are present with their own Web site on the Web were identified. The parties' home pages were identified using the Párt.lap.hu Website. The site was visited in September 2006. The full list of parties, acronyms and home page URLs can be found in *Appendix A.4.* This experiment was carried out during September 2006.

Six Web search engines were used to evaluate the usefulness of Hungarian acronyms. The search engines were selected according to the Hungarian search engine usage statistics (see *Section 6.5 Table 6.6*). In addition, six Web search engines were used to evaluate the identification capability of Danish acronyms. These search engines were selected according to the Danish search engine usage statistics (see *Section 6.7.2 Table 6.28*). The identification capability of acronyms was evaluated by entering each acronym to each of the search engines and evaluating the first ten results according to MICQ. Based on the search engine usage statistics in Hungary and Denmark the Pseudo Precision and Mean Pseudo Rank values were calculated using weighting schemes. The weighting scheme takes into account the distribution of search engine usage. The weighting schemes can be found in *section 6.5 and Section 6.7.2 The* Pseudo Precision and Mean Pseudo Rank were calculated according to *Equation 6.1* and *Equation 6.2.*

The results obtained in this application are presented in details in *Appendix B. Table B.7* shows the results of Hungarian acronyms; *Table B.8* shows the results of Danish acronyms.

Figure 6.23 shows the Pseudo Precision of acronyms of Hungarian parties. It can be seen that 3 acronyms – FIDESZ, MSZP, NDP– are very useful, the next 9 acronyms (60%) are useful and two acronyms can not be judged as being useful. MNYP and ZP do not identify its party in the selected search engines.

Figure 6.24 shows the Pseudo Precision of Acronyms of Danish parties. It can be seen that none acronym is very useful (Pseudo Precision equals 1). Five acronyms are useful (Pseudo Precision is greater than 0.5), the other three acronyms are somewhat useful in the selected search engines.

Table 6.33 shows the averages of Pseudo Precisions. It can be seen that the Average Pseudo Precision of the acronyms of Hungarian parties is higher by 15% than of Danish parties.

Table 6.33 Average Pseudo Precision of Hungarian and Danish parties.

	Average Pseudo Precision
Acronyms of Hungarian parties	0.72
Acronyms Danish parties	0.57

Figure 6.24 shows the Mean Pseudo rank of the acronyms of Hungarian parties. It can be seen that there are two acronyms – MSZP and FIDESZ – that can be used effectively to identify its party on the Web. 60% of the acronyms can also identity its party on the Web. MNYP and ZP cannot identify their institutions' home page on the Web.

Figure 6.25 shows the Mean Pseudo Rank of the acronyms of Danish parties. It can be seen that SF, DF and FrP identifies its party on the Web. KPD, SD and CD cannot be used to identify its party on the Web.

Table 6.34 shows the averages of Mean Pseudo Ranks. It can be seen that the average MPR of Hungarian acronyms is higher by 29% than the average MPR of Danish acronyms.

Table 6.34 Average Mean Pseudo Rank of the acronyms of Hungarian and Danish parties.

	Average Mean Pseudo Rank
Acronyms of Hungarian parties	0.68
Acronyms of Danish parties	0.39

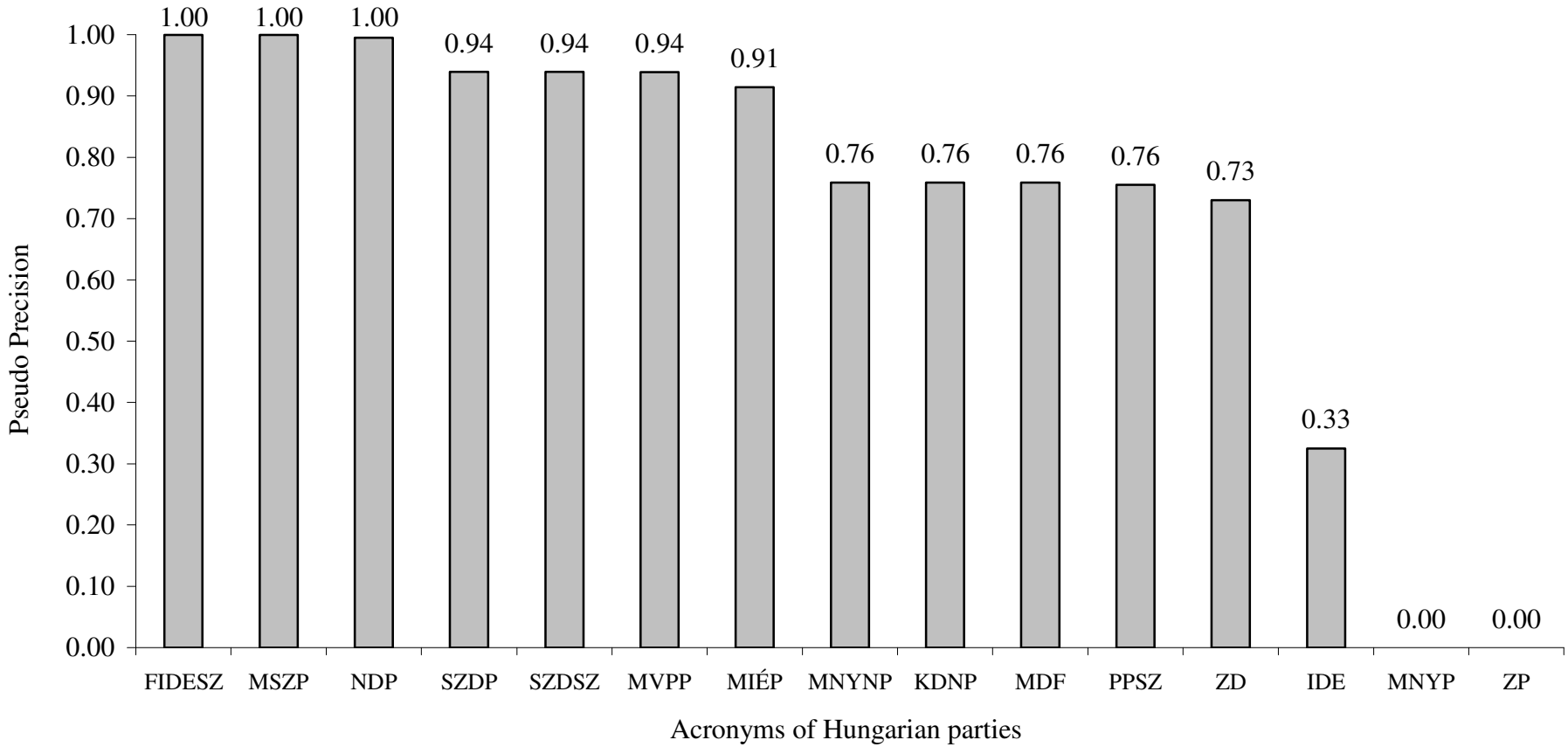

Figure 6.23 Pseudo Precision of the acronyms of Hungarian parties.

The majority of acronyms can identify its party on the Web. Twelve of the acronyms are useful. They Pseudo Precision are greater then 0.5. FIDESZ, MSZP and NDP are very useful. They Pseudo Precision equal 1. MNYP and ZP cannot identify its party on the Web.

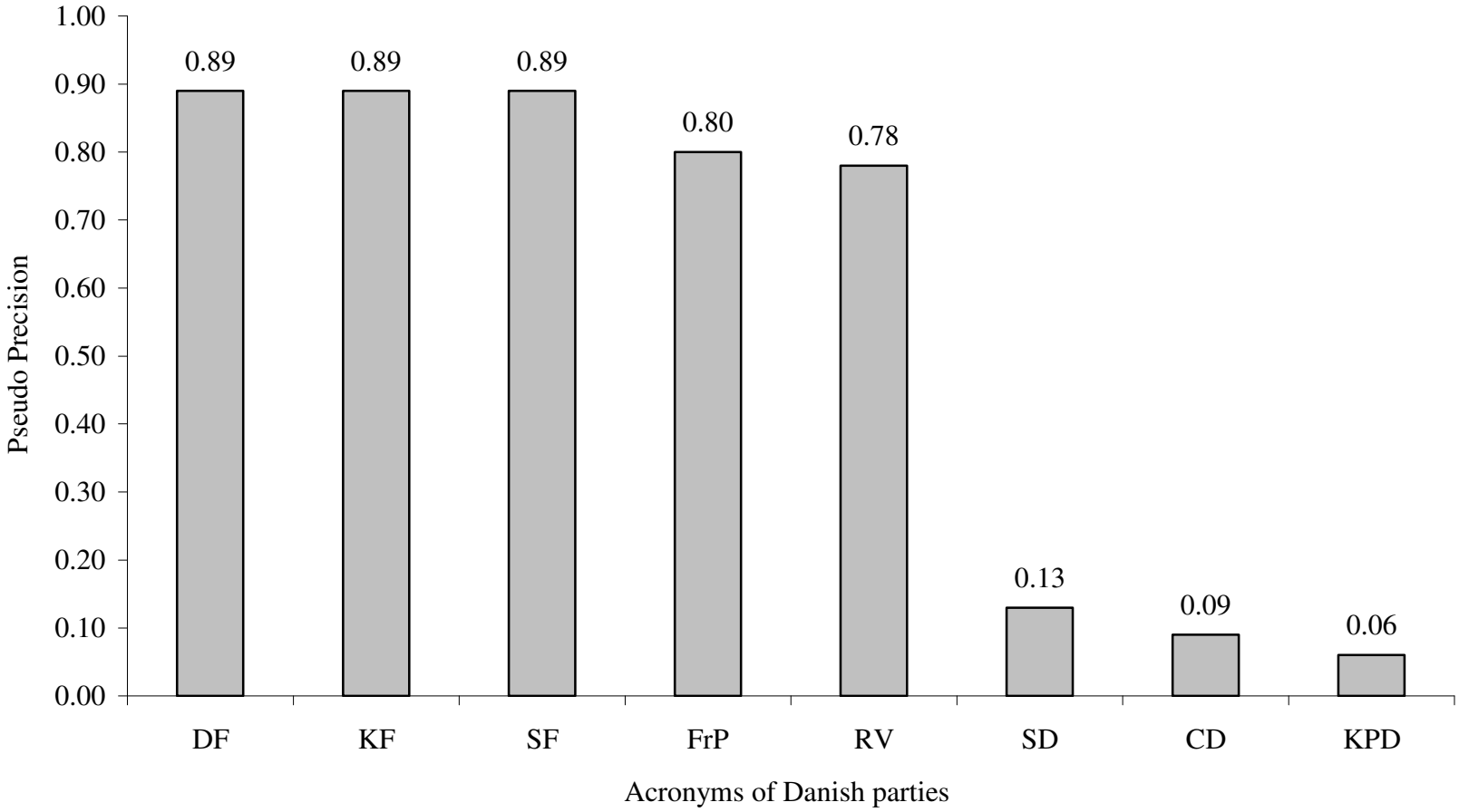

Figure 6.24 Pseudo Precision of the acronyms of Danish parties.

The majority of acronyms can identify its party on the Web. Five of the eight acronyms (63%) are useful. They Pseudo Precision are greater then 0.5. SD, CD and KPD cannot be judged as being useful.

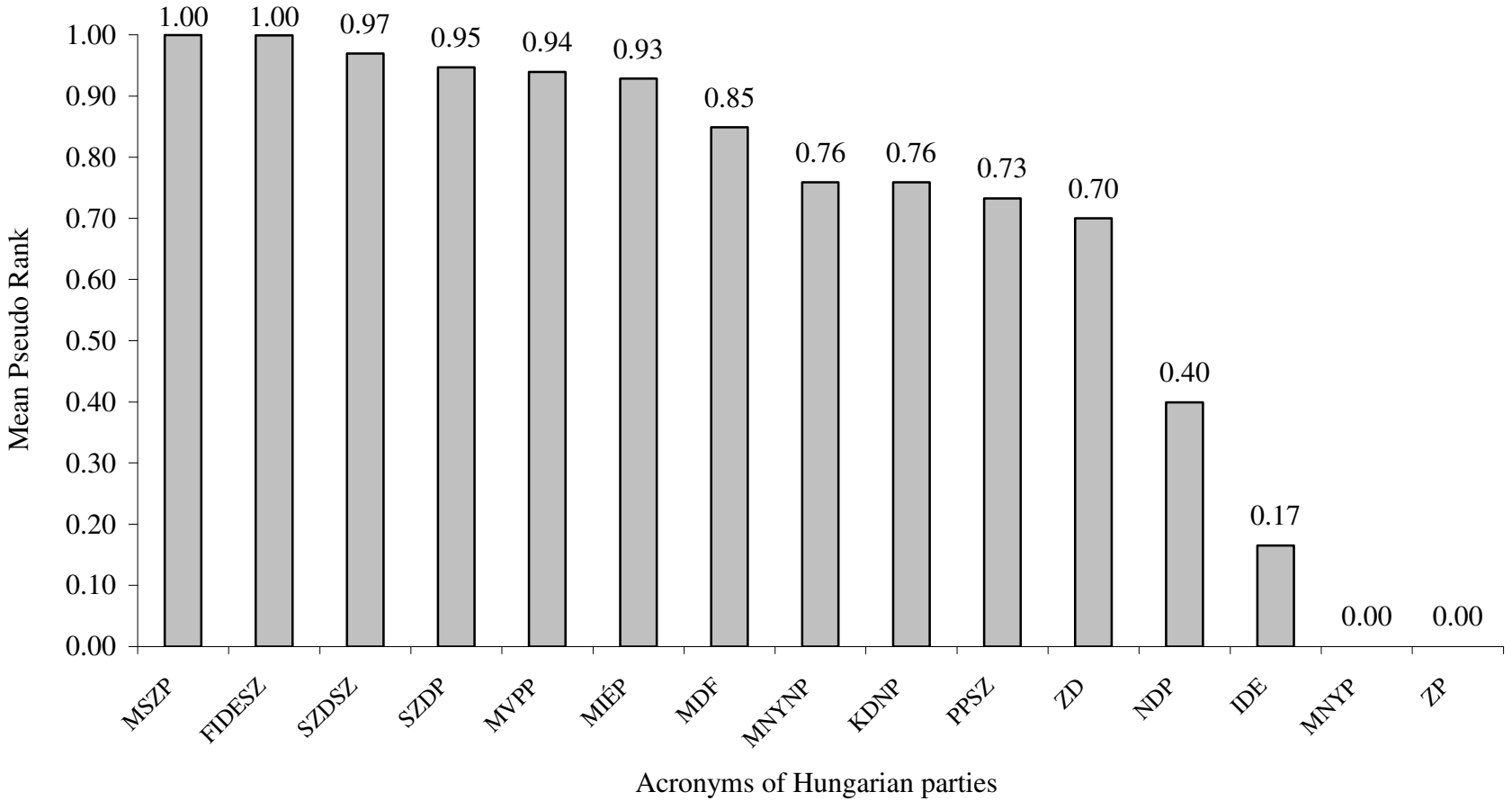

Figure 6.24 Mean Pseudo Rank of the acronyms of Hungarian parties.

It can be seen that the majority of acronyms (73%) has high identification capability (MPR is greater than 0.5). MNYP and ZP cannot identify its party on the Web using the selected search engines.

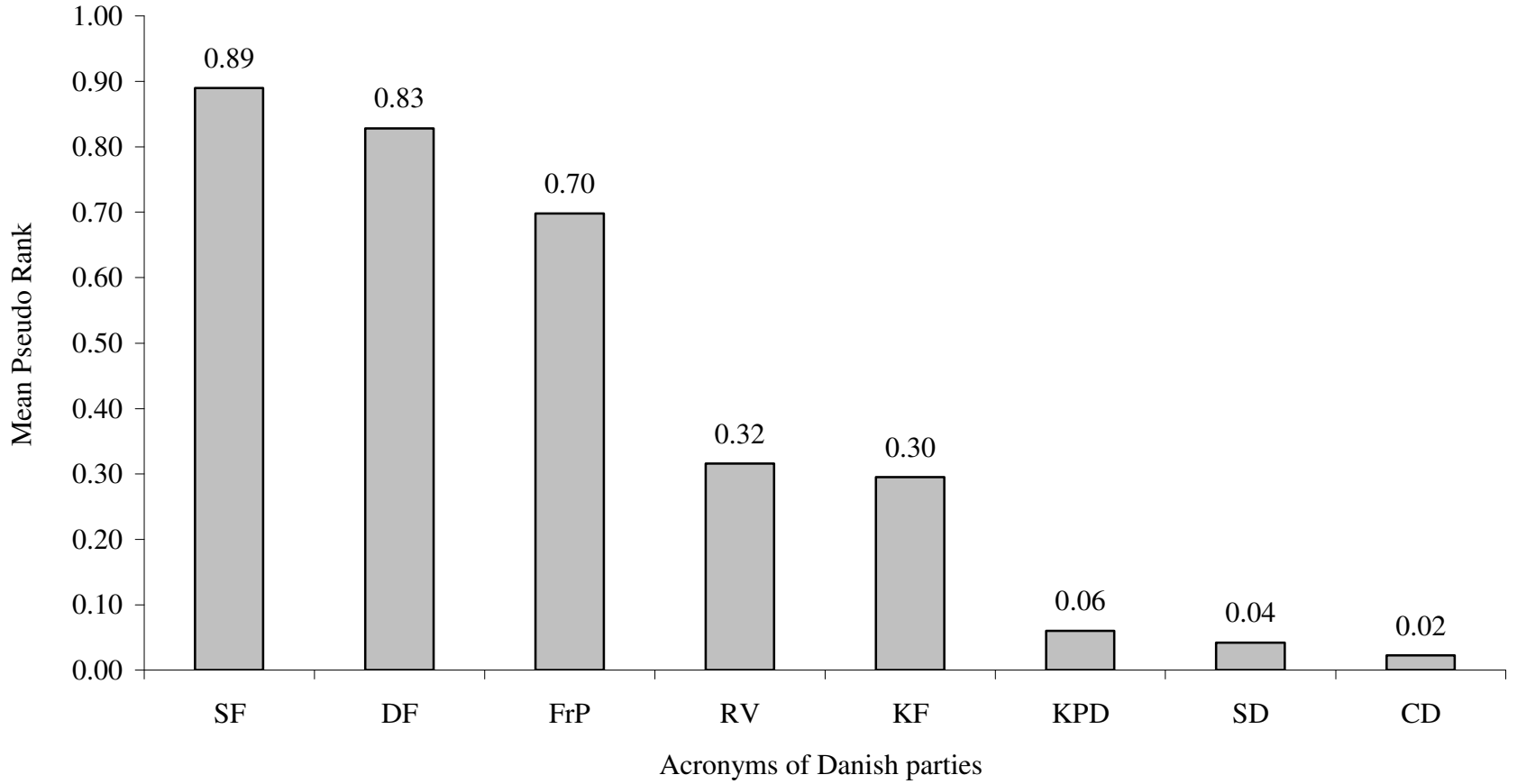

Figure 6.25 Mean Pseudo Rank of the acronyms of Danish parties.

It can be seen that the majority of acronyms (63%) are not effective in identifying its party on the Web. Only SF, DF and FrP can be effectively used to identify its party on the Web.

In this application the identification capability of the acronyms of Hungarian and Danish parties were investigated. The identification capability of Hungarian acronyms was investigated from Hungarian users' point of view. The identification capability of Danish acronyms was investigated from Danish users' point of view. Based on the results Hungarian users can easier find the home pages of Hungarian parties than Danish users can find the home pages of Danish parties.

6.9 Conclusions

In this chapter the applications of the MICQ method were presented. The identification capability of different institution categories was investigated. Table 6.35 shows the summary table statistics.

Table 6.35 Summary table statistics of the identification capability of acronyms.

Institution category	Average Pseudo Precision		Average Mean Pseudo Rank	
	General users	Hungarian / Danish users	General users	Hungarian / Danish users
Hungarian General	0.61	0.44	0.53	0.38
Hungarian Government Offices	–	0.74	–	0.55
Hungarian Higher Educational	0.33	0.55	0.3	0.55
Danish Higher Educational	0.57	0.86	0.36	0.74
Hungarian Parties	–	0.72	–	0.68
Danish Parties	–	0.57	–	0.39

Based on the results a situation report can be given about how effectively users can find the institutions of a country on the Web.

Summary

[Theses T3]

Using the MICQ method it was showed that:

a) the home page identification capability of the acronyms of Hungarian government offices: average Pseudo Precision is 0.74, average Mean Pseudo Rank is 0.55.

b) the home page identification capability of the acronyms of higher educational institutions in Hungary: average Pseudo Precision is 0.55, average Mean Pseudo Rank is 0.55.

c) the home page identification capability of the acronyms of higher educational institutions in Denmark: average Pseudo Precision is 0.86, average Mean Pseudo Rank is 0.74.

d) the home page identification capability of the acronyms of Hungarian parties: average Pseudo Precision is 0.72, average Mean Pseudo Rank is 0.68.

e) the home page identification capability of the acronyms of Danish parties: average Pseudo Precision is 0.57, average Mean Pseudo Rank is 0.39.

CHAPTER 7

CONCLUSIONS

The main contributions and the theses of the dissertation – both in English and in Hungarian – are presented in this chapter. Then, the publications related to the dissertation are listed.

7.1 Theses

Theses T1

I proposed the Pseudo Precision and Mean Pseudo Rank measures to evaluate the home page identification capability of queries on the Web. [Chapter 4] [SKROP 4, SKROP 7]

Theses T2

I proposed the MICQ method to measure the home page identification capability of search queries on the Web. [Chapter 5] [SKROP 4, SKROP 3, SKROP 7, SKROP 6]

Theses T3

Using the MICQ method it was showed that:

a) the home page identification capability of the acronyms of Hungarian government offices: average Pseudo Precision is 0.74, average Mean Pseudo Rank is 0.55. [Section 6.5]

b) the home page identification capability of the acronyms of higher educational institutions in Hungary: average Pseudo Precision is 0.55, average Mean Pseudo Rank is 0.55. [Section 6.6.2] [SKROP 3, SKROP 5, SKROP 6]

c) the home page identification capability of the acronyms of higher educational institutions in Denmark: average Pseudo Precision is 0.86, average Mean Pseudo Rank is 0.74. [Section 6.7.2]

d) the home page identification capability of the acronyms of Hungarian parties: average Pseudo Precision is 0.72, average Mean Pseudo Rank is 0.68. [Section 6.8]

e) the home page identification capability of the acronyms of Danish parties: average Pseudo Precision is 0.57, average Mean Pseudo Rank is 0.39. [Section 6.8]

7.2 Tézisek

Az értekezés új tudományos eredményei az alábbiakban foglalhatók össze:

Tézis T1

Megadtam a Pszeudo-pontosság és az Átlag Pszeudo-rang mértékeket a keresőkérdéseknek a honlapazonosító képességének mérésére a Web-en. [Chapter 4] [SKROP 4, SKROP 7]

Tézis T2

Megadtam a MICQ eljárást keresőkérdések honlapazonosító képességének mérésére a Web–en. [Chapter 5] [SKROP 4, SKROP 3, SKROP 7, SKROP 6]

Tézis T3

A MICQ eljárás alkalmazásával megmutattam, hogy

a) a magyar minisztérium–betűszavak honlapazonosító képességének Átlag Pszeudo–pontossága 0,74, Átlag Pszeudo–rangja 0,55. [Section 6.5]

b) a magyar felsőoktatási intézménynév–betűszavak honlapazonosító képességének Átlag Pszeudo–pontossága 0,55, Átlag Pszeudo–rangja 0,55. [Section 6.6.2] [SKROP 3, SKROP 5, SKROP 6]

c) a dán felsőoktatási intézménynév–betűszavak honlapazonosító képességének Átlag Pszeudo–pontossága 0,86, Átlag Pszeudo–rangja 0,74. [Section 6.7.2]

d) a magyar politikai pártbetűszavak honlap–azonosító képességének Átlag Pszeudo–pontossága 0,72, Átlag Pszeudo–rangja 0,68. [Section 6.8]

e) a dán politikai pártbetűszavak Átlag Pszeudo–pontossága 0,57, Átlag Pszeudo–rangja 0,39. [Section 6.8]

7.3 Publications

[SKROP 1] DOMINICH, S., SKROP, A., and TUZA, ZS. (2006). Formal Theory of Connectionist Web Retrieval. In: Crestani, F. et al. (eds.) *Soft Computing in Web Information Retrieval: Models and Applications*, Springer Verlag, pp. 161-194. ISBN: 3-540-31588-8.

[SKROP 2] DOMINICH, S., and SKROP, A. (2005). PageRank and Interaction Information Retrieval. *Journal of the American Society for Information Science and Technology*. vol. 56, no. 1, pp: 63-69. John Wiley & Sons , ISSN 1532-2882, IF = 2.086.

[SKROP 3] SKROP, A. and DOMINICH, S. (2004). Measuring the Identification Capability of Acronyms on the World Wide Web: a Comparative Study. *Journal of Web Engineering*. Vol. 3. No. 3 & 4, pp. 200-215. Rinton Press**,** ISSN 1540-9589.

[SKROP 4] DOMINICH, S., GÓTH, J., and SKROP, A. (2003). A Study of the Usefulness of Institutions' Acronyms as Web Queries.*: In: Sebastiani, F. (Ed.), Advances in Information Retrieval, Lecture Notes in Computer Science, LNCS 2633.* Springer Verlag, pp. 580-587. ISSN 0302-9743, IF=0.513.

[SKROP 5] SKROP, A. (2004). Do Acronyms of Higher Educational Institutions in Hungary Identify Their Own Institutions on the Web? *Proceedings of WWW2004 Workshop on Measuring Web Search Effectiveness: The User Perspective*, New York, NY, USA, May 17–22.

[SKROP 6] SKROP, A. (2004). Do acronyms of institutions in Hungary identify their own institutions on the Web?. *Proceedings of the 2nd PhD Mini-Symposium,* University of Veszprem, Veszprem, Hungary, June 21. pp. 75–77. ISBN 9639495573.

[SKROP 7] SKROP, A. (2003). A Study of the Usefulness of Institutions' Acronyms as Web Queries. *Proceedings of the 1st PhD Mini-Symposium,* University of Veszprem, Veszprem, Hungary, June 16. pp. 58–60. ISBN 9639495271.

[SKROP 8] GÓTH, J., and SKROP, A. (2005). Varying Retrieval Categoricity Using Hyperbolic Geometry. *Information Retrieval*. Vol. 8. No. 2. pp. 265-283. IF = 1.231

[SKROP 9] SKROP, A. (2002). Measuring Relevance Effectiveness in Information Retrieval. *Conference of PhD Students in Math, Physics and Informatics.* University of Veszprem, Veszprem, Hungary, May 30. (in Hungarian)

[SKROP 10] HORVÁTH, M. és SKROP, A. (2002). Searching the Hungarian Web. *Conference of INFO Savaria 2002*, Szombathely, Hungary, April 4-6. (in Hungarian)

[SKROP 11] SKROP, A. (2001). Content Exploration by Automatic Classification. *International Conference on Cultural Innovation and Information Retrieval*, National Széchenyi Library, Budapest, Hungary, September 14. (in Hungarian)

[SKROP 12] DOMINICH, S., HORVÁTH, M., and SKROP, A. (2001). Evaluations of Relevance Effectiveness in Interaction Information Retrieval. *Proceedings of the 23rd European Colloquium on Information Retrieval Research*, Springer Verlag, eWic, British Computer Society Information Retrieval Specialist Group, GMD IPSI, Darmstadt, Germany, April 4-6, pp. 208-221.

REFERENCES

AbbreviationZ. Available: http://www.stands4.com/

Acronym Finder. Available: http://acronymfinder.com

Acronym Search. Available: http://www.acronymsearch.com/

Acronyma. Available: http://www.acronyma.com/

Acronyms and Abbreviations. Available: http://www.chemie.fu-berlin.de/cgi-bin/acronym

Acrophile. Available: http://ciir.cs.umass.edu/irdemo/acronym/index.html

Agosti, M. and Melucci, M. (2001). Information Retrieval on the Web. In M. Agosti, F. Crestani, G. Pasi (Eds.),Lecture Notes in Computer Science ,Volume 1980 / 2001 Lectures on Information Retrieval: Third European Summer-School, ESSIR 2000, Varenna, Italy, September 11-15, 2000. Revised Lectures, pp. 242-285, Springer Berlin / Heidelberg ISSN 0302-9743.

ARGH. Available: http://invention.swmed.edu/argh/

Baeza-Yates, R. and Ribeiro-Neto, B. (1999). *Modern Information retrieval*. ACM Press New York, Addison-Wesley.

Bar-Ilan, J. (2005). Comparing rankings of search results on the Web. *Information Processing & Management*, Volume 41, Issue 6, December 2005, pp. 1511-1519.

Barlow, R.E. and Porschan, F. (1965.) *Mathematical Theory of Reliability*, New York: John Wiley and Sons.

Berners-Lee, T., Cailliau, R., Luotonen, A., Nielsen, H. F. And Secret, A. (1994). The World-Wide Web. *Communication of the ACM,* 37(8). pp. 76-82.

Braintrack University Index. Available: http://www.braintrack.com.

Broder, A. (1997). On the Resemblance and Containment of Documents. *Proceedings of Compression and Complexity of Sequences (SEQUENCES'97*), pp. 21--29. IEEE Computer Society, 1998..

Broder, A. (2002). A taxonomy of Web search. *ACM SIGIR Forum 36, 2,* pp. 3-10.

Chang, J.T., Schutze, H., and Altman, R.B. (2002). Creating an online dictionary of abbreviations from MEDLINE. *Journal of the American Medical Informatics Association*, 9(6). pp. 612--620, 2002.

Chu, H. and Rosenthal, M. (1996). Search Engines for the World Wide Web: A Comparative Study and Evaluation Methodology. *Proceedings of the American Society for Information Science Annual Meeting*, 33, pp: 127–135.

Craswell, N. and Hawking, D. (2001). Overview of the TREC-2001 Web Track. In Vorhees, E. and Harman, D. K. (eds.). *Proceedings of the 10th Text Retrieval Conference (TREC-2001)*, pp. 61-68. NIST special publication 500-250.

Craswell, N. and Hawking, D. (2002). Overview of the TREC-2002 Web Track. *Proceedings of TREC-2002*, Gaithersburg MD, November 2002. NIST special publication 500-251.

Craswell, N. and Hawking, D. (2004). Overview of the TREC-2004 Web Track. *Proceedings of The Thirteenth Text Retrieval Conference (TREC 2004).* NIST Special Publication: SP 500-261

Craswell, N., Hawking, D. and Griffiths, K. (2001a). Which search engine is best at finding airline site home pages? *Technical report*, CSIRO mathematical and informationsciences.
Available: http://research.microsoft.com/users/nickcr/pubs/craswell_tr01.pdf

Craswell, N., Hawking, D. and Robertson, S. (2001b). Effective Site Finding using Link Anchor Information. *Proceedings of the 24th Annual International ACM SIGIR Conference.* New Orleans, ACM Press, New York. pp. 250–257.

Craswell, N., Hawking, Wilkinson, D. R. and Wu, M. (2003). Overview of the TREC-2003 Web Track. In *NIST Special Publication 500-255: The Twelfth Text REtrieval Conference (TREC 2003)*, pages 78-92.

Dodgson, C. L. (1867). An Elementary Treatise on Determinants, with Their Application to Simultaneous Linear Equations and Algebraical Geometry. London: Macmillan, 1867.

Dong, X. (2003). Searching Information and Evaluation of Internet: A Chinese Academic User Survey. *The International Information & Library Review,* Volume 35, Issues 2-4, June-December 2003, Pages 163-187.

Egyetemek.lap.hu. Available: http://egyetem.lap.hu/

Eiron, N. And McCurley, K. S. (2003). Untangling Compound Documents on the Web. *Proceedings of the Fourteenth ACM Conference on Hypertext and Hypermedia*, pp. 85-94, 2003.

Ellis, D. (1996). The dilemma of measurement in information retrieval research. *Journal of the American Society for Information Science*, 47, 23-36.

Főiskolák.lap.hu. Available: http://foiskola.lap.hu/

Gnedenko, B. V., Belyayev, Y. K., and Solovyev, A. D. (1969). Mathematical Methods of Reliability Theory, Academic Press, New York.

Google, Inc., http://www.google.com

Gordon, M., and Pathak, P. (1999). Finding information on the World Wide Web: the retrieval effectiveness of search engines. *Information Processing and Management*, 35, pp: 141–180.

Gulli, A. and Signorini, A. (2005). The indexable Web is more than 11.5 billion pages. *Proceedings of 14th International World Wide Web Conference*, pages 902–903, Chiba, Japan.

Gwizdka, J. and Chignell, M. (1999). Towards Information Retrieval Measures for Evaluation of Web Search Engines. *Unpublished manuscript.* Available at http://www.imedia.mie.utoronto.ca/~jacekg/pubs.html

Hawking, D. et al. (1999). Results and challenges in Web search evaluation. *Proceedings of WWW 8.* pp. 243–252, Amsterdam: Elsevier.

Hull, D. (1993). Using statistical testing in the evaluation of retrieval experiments. *Proceedings of ACM–SIGIR `93.* pp. 329–338, New York: ACM.

Index. Available: http://katalog.index.hu

Jároli, J. (2006a). Keresőoldalak összehasonlítása. Available: http://webni.innen.hu/Keres_c5_91oldalak_c3_b6sszehasonl_c3_adt_c3_a1sa2005Augusztus

Jároli, J. (2006b). Which Hungarian Search Engines are Used the Most. Available: http://totheweb.blogspot.com/2006_01_01_totheweb_archive.html

Joachims, T. (2002). Evaluating retrieval performance using clickthrough data. *Proceeding of the ACM SIGIR `02 Workshop on Mathematical/Formal Methods in Information Retrieval.* Tampere, Finland, August 15, 2002.

Kleiberg, J. M. (1999). Authoritative Sources in a Hyperlinked Environment. *Journal of the ACM*, vol. 46, no. 5, pp: 604-632.

Larkey, L.S., Ogilvie, P., Price, M.A., and Tamilio, B. (2000). Acrophile: An Automated Acronym Extractor and Server. *Digital Libraries – ACM Conference on Digital Libraries*, San Antonio, TX, June 2-7, ACM Press, pp: 205-214.

Leighton, H. V. (1995). Performance of four World Wide Web (WWW) index services: Infoseek, Lycos, WebCrawler, and WWWWorm [On-line]. Available: http://www.winona.msus.edu/is-f/library-f/Webind.htm

Leighton, H. V. and Srivastava, J. (1997). Precision Among WWW Search Services (Search Engines): Alta Vista, Excite, Hotbot, Infoseek, Lycos. http://www.winona.msus.edu/library/Webind2/Webind2.htm

Leighton, H. V., and Srivastava, J. (1999). First Twenty Precision among World Wide Web Search Services (Search Engines). *Journal of the American Society for Information Science,* **50**(10): 870-881.

Ljosland, M. (1999). Evaluation of Web search engines and the search for better ranking algorithms. *Proceeding of the ACM SIGIR '99 Workshop on Evaluation of Web Document Retrieval*. Berkeley, CA, USA - August 19, 1999.

MacCall, S. L. and Cleveland, A. D. (1999). Relevace–based quantitative measure for Internet information retrieval evaluation. *Proceedings of the American Society for Information Science Annual Meeting*, 36, pp.763–768.

Magyarország.hu. Available: http://www.magyarorszag.hu/

Martin, W. T., Reissner, E. (1961). *Elementary Differential Equations*. Addison-Wesley, Reading-Massachusetts, U.S.A.

Mizzaro, S. (1997). Relevance: The Whole History. *JASIS 48, 9*. pp. 810-832.

Mizzaro, S. (2001). A new measure of retrieval effectiveness (or : what's wrong with precision and recall). *In T. Ojala (Ed.), International Workshop on Information Retrieval* (IR 2001). pp. 43-52, Infotech Oulu, Oulu, Finnland.

Nash, F. R. (2003). *Estimating Device Reliability: Assessment of Credibility*. Kluwer Academic Publishers, ISBN 0-7923-9304-X.

Norden. Available: http://www.abo.fi/norden/uniinfo/uniinfdk.htm.

Oppenheim, C., Morris, A., and McKnight, C. (2000). The evaluation of WWW search engines. *Journal of Documentation*, 56(2), pp. 190-211.

Országos Felsőoktatási Információs Központ http://www.felvi.hu/

Párt.lap.hu. Available: http://part.lap.hu/

Poole, D. G. (1995).The Stochastic Group. *American Mathematical Monthly* **102**, pp. 798-801.

Pustejovsky, J., Castano, J., Cochran, B., Kotecki, M., Morrell, M. (2001). Automatic extraction of acronym-meaning pairs from MEDLINE databases. *Medinfo* 2001, 10. pp. 371-375.

Rasmussen, E. (2002). Evaluation in Information Retrieval. *Proceedings of the ISMIR 2002 Conference Panel I.: Music Information Retrieval Evaluation Frameworks*, Paris, France, October 17 2002, pp. 45-49.

Risvik K. M., Michelsen R. (2002). Search engines and Web dynamics. *Computer Networks*, 39 (3), pp. 289-302.

Ruch, P., and Baud, R. (2002). Evaluating and reducing the effect of data corruption when applying bag of words approaches to medical records. *International Journal of Medical Informatics*. 67 (1-3) pp.75-83.

Saracevic, T. (1995). Evaluation of evaluation in information retrieval. *Proceedings of the 18th ACM–SIGIR international conference on research and development in information retrieval.* pp. 138–146. WA, Seattle: Association of Computing Machinery.

Savoy, J. and Picard, J. (2001). Retrieval effectiveness on the Web. *Information Processing and Management*, 37, pp. 543–569.

Shah, C. and Croft, W. B. (2004). Evaluating High Accuracy Retrieval Techniques. Proceedings of ACM SIGIR'04 (Sheffield, United Kingdom, 2004), pp. 2–9.

Silverstein, C., Henzinger, M., Marais, H. and Moricz, M. (1998). Analysis of a Very Large AltaVista Query Log. *Technical Report* 1998-014*, Digital Systems Research Center.*

Singhal, A. And Kaszkiel, M. A. (2001). A Case Study in Web Search using TREC algorithms. *Proceedings of WWW10,* Hong Kong, pp. 708-716.

Sparck Jones, K. (1981) The Cranfield tests. In: *Information Retrieval Experiment.* (K. Sparck Jones, ed.). London: Butterworth.

Special Dictionary Acronyms, Abbreviations and Initialisms. Available: http://www.special-dictionary.com/acronyms/

Spink, A., Jansen, B. J., Wolfram, D. and Saracevic, T. (2002): From E-Sex to E-Commerce: Web Search Changes. *IEEE Computer* 35(3): 107-109.

Startlap. Available: http://www.lap.hu/temakor.html

Sufyan Beg, M. M. (2005). A subjective measure of web search quality. *Information Sciences*, Volume 169, Issues 3-4, 1 February 2005, pp. 365-381.

Taghva, K. and Gilbreth, J. (1999). Recognizing acronyms and their definitions. *International journal on Document Analysis and Recognition*. pp. 191-198.

Tague–Sutcliffe, J. M. (1996). Some perspectives on the evaluation of information retrieval systems. *Journal of the American Society for Information Science*, 47, 1-3.

The Internet Acronym Server. Available: http://silmaril.ie/cgi-bin/uncgi/acronyms

The Search Engine Landscape in Denmark. Available: http://www.multilingual-search.com/the-danish-search-landscape/04/01/2006/en/

Top search engines in Hungary? http://gugli.wordpress.com/2006/07/27/top-search-engines-in-hungary/

TREC: Text REtrieval Conference (TREC) Home Page. Available: http://trec.nist.gov

TREC-2003 Web Track Guidelines. Available: http://es.csiro.au/TRECWeb/guidelines_2003.html.

Upstill, T. G. (2005). Document ranking using web evidence. *PhD Theses.* Available: http://www.webir.org/resources/phd/Upstill_2005.pdf

Wahoo. Available: http://www.wahoo.hu/index

Wang, Y. D. and Forgionne, G. (2006). A decision-theoretic approach to the evaluation of information retrieval systems. *Information Processing & Management*, Volume 42, Issue 4, July 2006, pp. 863-874.

Web Track. Available: http://es.cmis.csiro.au/TRECWeb

Webmania. Available: http://www.Webmania.hu

Wren, J. D. and Garner, H. R. (2002). Heuristics for identification of acronym-definition patterns within text: Towards an automated construction of comprehensive acronym-definition dictionaries. *Methods of Information in Medicine* 2002; 41(5). pp. 426-34.

Wu, M. M. & Sonnenwald, D. H. (1999.) Reflections on information retrieval evaluation. *Proceedings of the 1999 EBTI, ECAI, SEER & PNC Joint Meeting* (pp. 63-81), Academia Sinica, Taipei.

Yahoo. Available: http://www.yahoo.com/

Yeates, S., Bainbridge, D., and Witten, I. H. (2000). Using Compression to Identify Acronyms in Text. *Proceedings of the Conference on Data Compression*, March 28 - 30, 2000. DCC. IEEE Computer Society, Washington, DC, pp. 582.

APPENDIX A

Appendix A contains the lists that were compiled during the Applications. The lists are presented in table format. The tables contain the acronym of the desired institution, the full name of the institution and the URL of the home page. The appendix is organised as follows:

A.1 Acronyms and home pages of Hungarian General Institutions

A.2 Acronyms and home pages of Hungarian Higher Educational Institutions

A.3 Acronyms and home pages of Danish Higher Educational Institutions

A.4 Acronyms and home pages of Hungarian and Danish parties

A.5 Acronyms and home pages of Hungarian government offices

A.1 Acronyms and Home Pages of Hungarian General Institutions

Hungarian General Institutions		
Acronym	**Full Name**	**Home Page URL**
APEH	Adó- és Pénzügyi Ellenőrzési Hivatal	http://www.apeh.hu/
ÁNTSZ	Állami Népegészségügyi és Tisztiorvosi Szolgálat	http://www.antsz.hu/
ÁPV	Állami Privatizációs és Vagyonkezelő Rt.	http://www.apvrt.hu/
BKIK	Budapesti Kereskedelmi és Iparkamara	http://www.bkik.hu/index.shtml
BKV	Budapesti Közlekedési Vállalat	http://www.bkv.hu/
BVK	Budapesti Vállalkozásfejlesztési Közalapítványt	http://www.bvk.hu/
DOSZ	Doktoranduszok Országos Szövetsége	http://www.phd.hu/
FVF	Fogyasztóvédelmi Főfelügyelőség	http://www.fvf.hu/
GVH	Gazdasági Versenyhivatal	http://www.gvh.hu/
HíF	Hírközlési Felügyelet	http://www.hif.hu/
KOMA	Közoktatási Modernizációs Közalapítvány	http://www.koma.hu/
MÁV	Magyar Államvasutak	http://www.mav.hu/
MBH	Magyar Bányászati Hivatal	http://www.mbh.hu/
MEH	Magyar Energia Hivatal	http://www.eh.gov.hu/
MÉK	Magyar Építész Kamara	http://www.mek.hu/
MGYK	Magyar Gyógyszerész Kamara	http://www.mgyk.hu/
MAHART	Magyar Hajózási Rt.	http://www.mahart.hu/
MKIK	Magyar Kereskedelmi és Iparkamara	http://www.mkik.hu/
MKVK	Magyar Könyvvizsgálói Kamara	http://www.mkvk.hu/
AMC	Magyar Közösségi Agrármarketing Centrum Közhasznú Társaság	http://www.amc.hu/
MALÉV	Magyar Légitársaság	http://www.malev.hu/

Hungarian General Institutions		
Acronym	**Full Name**	**Home Page URL**
MOL	Magyar Olaj- és Gázipari Rt.	http://www.mol.hu/site/magyar/index.php
MOKK	Magyar Országos Közjegyzői Kamara	http://www.mokk.hu/
MOK	Magyar Orvosi Kamara	http://www.mok.hu/
MÖB	Magyar Ösztöndíj Bizottság	http://www.scholarship.hu/, http://www.mob.om.hu/
MSZH	Magyar Szabadalmi Hivatal	http://www.hpo.hu/
MSZT	Magyar Szabványügyi Testület	http://www.mszt.hu/
matáv	Magyar Távközlési Részvénytársaság	http://www.matav.hu/
MTA	Magyar Tudományos Akadémia	http://www.mta.hu
MVA	Magyar Vállalkozásfejlesztési Alapítvány	http://www.mva.hu/
MVK	Magyar Vöröskereszt	http://www.voroskereszt.hu/
MKGI	Miniszterelnökség Közbeszerzési és Gazdasági Igazgatósága	http://www.kozbeszerzes.gov.hu/
MGYOSZ	Munkaadók és Gyáriparosok Országos Szövetsége	http://www.mgyosz.hu/fomenufl.html
MBF	Műszaki Biztonsági Főfelügyelet	http://www.mbf.hu/
NBH	Nemzetbiztonsági Hivatal	http://www.nbh.hu/
NIIF	Nemzeti Információs Infrastruktúra Fejlesztési Program	http://www.iif.hu/
NKA	Nemzeti Kulturális Alapprogram	http://www.nka.hu/
NIOK	Nonprofit Információs és Oktató Központ Alapítvány	http://www.niok.hu/
OEP	Országos Egészségbiztosítási Pénztár	http://www.oep.hu/
OFA	Országos Foglalkoztatási Közalapítvány	http://www.ofa.hu/
OKI	Országos Közoktatási Intézet	http://www.oki.hu/
OMH	Országos Mérésügyi Hivatal	http://www.omh.hu/
OMSZ	Országos Meteorológiai Szolgálat	http://www.met.hu/
OMIKK	Országos Műszaki Információs Központ és Könyvtár	http://www.omikk.hu/
ONYF	Országos Nyugdíjbiztosítási Főigazgatóság	http://www.onyf.hu/onyfuj/jsp/Intro.jsp?mode=flash
ORTT	Országos Rádió és Televízió Testület	http://www.ortt.hu/
OTKA	Országos Tudományos Kutatási Alapprogramok	http://www.otka.hu/
SZF	Szerencsejáték Felügyelet	http://www.szf.hu/
VOSZ	Vállalkozók és Munkáltatók Országos Szövetsége	http://vosz.cyber.hu/index.html
FVM	Földművelésügyi és Vidékfejlesztési Minisztérium	http://www.fvm.hu/
GM	Gazdasági és Közlekedési Minisztérium	http://www.gm.hu/
HM	Honvédelmi Minisztérium	http://www.honvedelem.hu/
ISM	Gyermek-, Ifjúsági és Sportminisztérium	http://www.ism.hu
BM	Belügyminisztérium	http://www.b-m.hu/
IM	Igazságügyi Minisztérium	http://www.im.hu/
IHM	Informatikai és Hírközlési Minisztérium	http://www.ihm.gov.hu/
KvVM	Környezetvédelmi és Vízügyi Minisztérium	http://www.ktm.hu/

Hungarian General Institutions		
Acronym	**Full Name**	**Home Page URL**
NKÖM	Nemzeti Kulturális Örökség Minisztériuma	http://www.nkom.hu/
OM	Oktatási Minisztérium	http://www.om.hu/
PM	Pénzügyminisztérium	http://www.p-m.hu/
FMM	Foglalkoztatáspolitikai és Munkaügyi Minisztérium	http://www.szcsm.gov.hu/
SZCSM	Egészségügyi, Szociális és Családügyi Minisztérium	http://www.eum.hu/
MeH	Miniszterelnöki Hivatal	http://www.ekormanyzat.hu/d/redirect/meh.html
MTI	Magyar Távirati Iroda Rt.	http://www.mti.hu/
TÁRKI	Társadalomkutatási Intézet és Társadalomkutatási Informatikai Egyesülés	http://www.tarki.hu/
ÁB-AEGON, ÁBAEGON	ÁB-AEGON Általános Biztosító Rt.	http://www.aegon.hu/index.html
ÁÉB	Általános Értékforgalmi Bank Rt.	http://www.gbt.hu/
ÁKK	Államadósság Kezelő Központ Rt.	http://www.allampapir.hu/
ÁSZ	Állami Számvevőszék	http://www.asz.gov.hu/ASZ/www.nsf/frame?openForm
BÉT	Budapesti Értéktőzsde	http://www.bet.hu/
CIB	CIB Bank Rt.	http://www.cib.hu/magyar/infocenter/index.jsp
EXIM	Magyar Export-Import Bank Rt.	http://www.eximbank.hu/Internet/_main.asp
FHB	Földhitel- és Jelzálogbank Rt.	http://www.fhb.hu/
Generali	Generali-Providencia Biztosító Rt.	http://www.generali.hu/
HB	Allianz Hungária Biztosító Rt.	http://www.ahbrt.hu/
HVB	HypoVereinsbank Hungária Rt.	http://www.hvb.hu/
HVG	Heti Világ Gazdaság	http://hvg.hu/
K& H, KHB	Kereskedelmi és Hitelbank Rt.	http://www.khb.hu/
KSH	Központi Statisztikai Hivatal	http://www.ksh.hu/pls/ksh/docs/index.html
MABISZ	Magyar Biztosítók Szövetsége	http://www.mabisz.hu/
MEHIB	Magyar Exporthitel Biztosító Rt.	http://www.mehib.hu/
MFB	Magyar Fejlesztési Bank Rt.	http://www.mfb.hu/
MKB	Magyar Külkereskedelmi Bank Rt.	http://www.mkb.hu
MNB	Magyar Nemzeti Bank	http://www.mnb.hu/main.asp
OTP	Országos Takarékpénztár Rt.	https://www.otpbank.hu/otpportal2000/
PSZÁF	Pénzügyi Szervezetek Állami Felügyelete	http://www.pszaf.hu/
VG	Világgazdaság	http://www.vilaggazdasag.hu/index2.php
VPOP	Vám- és Pénzügyőrség Országos Parancsnoksága	http://www.vam.hu/index2.html
FKGP	Független Kisgazdapárt	http://www.fkgp.hu/
MDF	Magyar Demokrata Fórum	http://www.mdf.hu/
MIÉP	Magyar Igazság és Élet Pártja	http://www.miep.hu/
MSZP	Magyar Szocialista Párt	http://www.mszp.hu/
NDP	Nemzeti Demokrata Párt	http://www.datanet.hu/ndp/
SZDSZ	Szabad Demokraták Szövetsége	http://www.szdsz.hu/

Hungarian General Institutions		
Acronym	**Full Name**	**Home Page URL**
SZDP	Szociáldemokrata Párt	http://www.szdp.hu/
BGF	Budapesti Gazdasági Főiskola	http://www.bgf.hu/
BKÁE, BKE	Budapesti Közgazdaságtudományi és Államigazgatási Egyetem	http://www.bke.hu/
BME, BMGE	Budapesti Műszaki és Gazdaságtudományi Egyetem	http://www.bme.hu/
BMF	Budapesti Műszaki Főiskola	http://www.bmf.hu/
DE	Debreceni Egyetem	http://www.klte.hu/0_index.html
ELTE	Eötvös Loránd Tudományegyetem	http://www.elte.hu/
KÉE	Kertészeti és Élelmiszeripari Egyetem	http://www.kee.hu/
MKE	Magyar Képzőművészeti Egyetem	http://www.mke.hu/
NyME	Nyugat-Magyarországi Egyetem	http://www.nyme.hu/hu/index.html
PTE	Pécsi Tudományegyetem	http://www.pte.hu/
SOTE	Semmelweis Egyetem	http://www.sote.hu/
SZIE	Szent István Egyetem	http://www.szie.hu/index.htm
SZE	Széchenyi István Egyetem	http://www.sze.hu/
SZTE	Szegedi Tudományegyetem	http://www.u-szeged.hu/
VE	Veszprémi Egyetem	http://www.vein.hu/
ZMNE	Zrínyi Miklós Nemzetvédelmi Egyetem	http://www.zmne.hu/

A.2 Acronyms and Home Pages of Hungarian Higher Educational Institutions

Hungarian Higher Educational Institutions		
Acronym	**Full Name**	**Home Page URL**
BKÁE TK	Budapesti Közgazdaságtudományi és Államigazgatási Egyetem Társadalomtudományi Kar	http://www.bkae.hu/subpage.php?org=7
BKÁE	Budapesti Közgazdaságtudományi és Államigazgatási Egyetem	www.bkae.hu
BKÁE GTK	Budapesti Közgazdaságtudományi és Államigazgatási Egyetem Gazdálkodástudományi Kar	http://www.bke.hu/subpage_choice_control.php?org=5&id=145&LNG=hun
BKÁE KTK	Budapesti Közgazdaságtudományi és Államigazgatási Egyetem Közgazdaságtudományi Kar	www.bkae.hu
BKE	Budapesti Közgazdaságtudományi Egyetem	www.bkae.hu
BME	Budapesti Műszaki és Gazdaságtudományi Egyetem	www.bme.hu
BME ÉSZK	Budapesti Műszaki és Gazdaságtudományi Egyetem Építészmérnöki Kar	www.epitesz.bme.hu
BME ÉÖK	Budapesti Műszaki és Gazdaságtudományi Egyetem Építőmérnöki Kar	http://www.bme.hu/hu/szervezet/karok/epito/index.html
BME GÉK	Budapesti Műszaki és Gazdaságtudományi Egyetem Gépészmérnöki Kar	http://www.gepesz.bme.hu/

Hungarian Higher Educational Institutions		
Acronym	**Full Name**	**Home Page URL**
BME GTK	Budapesti Műszaki és Gazdaságtudományi Egyetem Gazdaság és Társadalomtudományi Kar	http://www.gtk.bme.hu
BME KSK	Budapesti Műszaki és Gazdaságtudományi Egyetem Közlekedésmérnöki Kar	http://www.kozlek.bme.hu
BME TTK	Budapesti Műszaki és Gazdaságtudományi Egyetem Természettudományi Kar	www.ttk.bme.hu
BME VEK	Budapesti Műszaki és Gazdaságtudományi Egyetem Vegyészmérnöki Kar	http://www.ch.bme.hu/
BME VIK	Budapesti Műszaki és Gazdaságtudományi Egyetem Villamosmérnöki és Informatikai Kar	www.vdk.bme.hu
DE	Debreceni Egyetem	http://www.klte.hu/
DE ÁOK	Debreceni Egyetem Általános Orvostudományi Kar	http://www.dote.hu/
DE BTK	Debreceni Egyetem Bölcsésztudományi Kar	http://btk.unideb.hu/
DE JÁTI	Debreceni Egyetem Állam- és Jogtudományi Kar	http://www.law.klte.hu/jati/kezdolap/
DE K	Debreceni Egyetem Konzervatóriuma	http://delfin.klte.hu/~de-konz/de_konz_ma.html
DE KTK	Debreceni Egyetem Közgazdaságtudományi Kar	http://www.econ.klte.hu/szoveg.html
DE MTK	Debreceni Egyetem Mezőgazdaságtudományi Kar	http://www.date.hu/
DE TTK	Debreceni Egyetem Természettudományi Kar	http://hi.ttk.unideb.hu/
DRHE	Debreceni Református Hittudományi Egyetem	http://www.drhe.drk.hu/
EFE	Erdészeti és Faipari Egyetem	már nincs nyme erdőmérnöki kara
EHE	Evangélikus Hittudományi Egyetem	http://teol.lutheran.hu/
ELTE	Eötvös Loránd Tudományegyetem	http://www.elte.hu/
ELTE ÁJK	Eötvös Loránd Tudományegyetem Állam és Jogtudományi Kar	http://www.ajk.elte.hu/
ELTE BTK	Eötvös Loránd Tudományegyetem Bölcsészettudományi Kar	http://www.btk.elte.hu/
ELTE TTK	Eötvös Loránd Tudományegyetem Természettudományi Kar	http://ttk.elte.hu/
ELTE PPK	ELTE Pedagógiai és Pszichológiai Kar	http://www.ppk.elte.hu/
ELTE IK	ELTE Informatikai Kar	http://www.inf.elte.hu/
JATE	Jozsef Attila Tudomanyegyetem	http://www.jate.u-szeged.hu/indexh.html
KE ÁTK	Kaposvári Egyetem Állattudományi Kar	http://www.kaposvar.pate.hu/index.htm
KGRE	Károli Gáspár Református Egyetem	http://www.kgre.hu/
KLTE	Kossuth Lajos Tudományegyetem	már nincs ilyen DE lett
KRE	Károli Gáspár Református Egyetem	http://www.kgre.hu/
KRE ÁJK	Károli Gáspár Református Egyetem Állam és Jogtudományi Kar	http://www.kre.hu/ajk/index.php?n=1&id=1
KRE BTK	Károli Gáspár Református Egyetem Bölcsészettudományi Kar	http://www.kre.hu/btk/index.php?n=1&id=1
KRE HTK	Károli Gáspár Református Egyetem Hittudományi Kar	http://www.kre.hu/htk/index.php?n=1&id=1
LFZE	Liszt Ferenc Zeneművészeti Egyetem	http://www.musicacademy.hu/magyar/index.html
ME	Miskolci Egyetem	http://www.uni-miskolc.hu/

Hungarian Higher Educational Institutions		
Acronym	**Full Name**	**Home Page URL**
ME ÁJK	Miskolci Egyetem Állami és Jogtudományi Kar	http://www.uni-miskolc.hu/law/
ME AKK	Miskolci Egyetem Anyag és Kóhómérnöki Kar	http://www.akk.uni-miskolc.hu/
ME BTK	Miskolci Egyetem Bölcsészettudományi Kar	http://www.bolcsWeb.hu/
ME GÉK	Miskolci Egyetem Gépészmérnöki Kar	http://gepesz.uni-miskolc.hu/
ME GTK	Miskolci Egyetem Gazdaságtudományi Kar	http://www.gtk.uni-miskolc.hu/
ME MFK	Miskolci Egyetem Műszaki Földtudományi Kar	http://www.uni-miskolc.hu/~mfk/
MIE	Magyar Iparművészeti Egyetem	http://www.mie.hu/index.php
MIRT	Méréstechnika és Információs Rendszerek Tanszék BME	http://www.mit.bme.hu/
MKE	Magyar Képzőművészeti Egyetem	http://www.mke.hu/
NME	Nehézipari Műszaki Egyetem	http://www.uni-miskolc.hu/uni/univ/tortenet/mult107.htm
NYME	Nyugat Magyarországi Egyetem	http://www.nyme.hu/hu/index.php
PPKE	Pázmány Péter Katolikus Egyetem	http://www.ppke.hu/
PPKE BTK	Pázmány Péter Katolikus Egyetem Bölcsészettudományi Kar	http://www.btk.ppke.hu/
PPKE HTK	Pázmány Péter Katolikus Egyetem Hittudományi Kar	http://www.htk.ppke.hu/
PPKE ITK	Pázmány Péter Katolikus Egyetem Információs Technológiai Kar	http://www.itk.ppke.hu/
PPKE JÁK	Pázmány Péter Katolikus Egyetem Jog és Államtudományi Kar	http://www.jak.ppke.hu/
PTE	Pécsi Tudományegyetem	http://www.pte.hu/
PTE ÁJK	Pécsi Tudományegyetem Állam és Jogtudományi Kar	http://www.law.pte.hu/
PTE ÁOK	Pécsi Tudományegyetem Általános Orvostudományi Kar	http://www.pote.hu/
PTE BTK	Pécsi Tudományegyetem Bölcsészettudományi Kar	http://www.btk.pte.hu/
PTE FEEFI	Pécsi Tudományegyetem Természettudományi Kar Felnőttképzési és Emberi Erőforrás Fejlesztési Intézet	http://gaia.jpte.hu/
PTE MK	Pécsi Tudományegyetem Művészeti Kar	http://art.pte.hu/
PTE TI	Pécsi Tudományegyetem Tanárképző Intézet	www.tki.pte.hu
PTE TTK	Pécsi Tudományegyetem Természettudományi Kar	http://www.ttk.pte.hu/
SE	Semmelweis Egyetem	http://www.sote.hu/
SE ÁOK	Semmelweis Egyetem Általános Orvostudományi kar	http://www.sote.hu/oktatas/aok/
SE FOK	Semmelweis Egyetem Fogorvostudományi Kar	http://www.sote.hu/oktatas/fok/
SE GYTK	Semmelweis Egyetem Gyógyszerésztudományi Kar	http://gytk.uw.hu/
SE TSK	Semmelweis Egyetem Testnevelési és Sporttudományi Kar	http://www.hupe.hu/01.html
SOTE	Semmelveis OrvosTudományi Egyetem	már nem, SE
SZE	Széchenyi István Egyetem	http://www.sze.hu/
SZFE	Színház és Filmművészeti Egyetem	http://www.filmacademy.hu/

Hungarian Higher Educational Institutions		
Acronym	**Full Name**	**Home Page URL**
SZIE	Szent István Egyetem	http://www.szie.hu/
SZIE ÁOTK	Szent István Egyetem Állatorvostudományi Kar	http://www.univet.hu/
SZIE ÉTK	Szent István Egyetem Élelmiszertudományi Kar	http://www.physics.kee.hu/
SZIE GÉK	Szent István Egyetem Gépészmérnöki Kar	http://www.mgk.gau.hu/
SZIE GTK	Szent István Egyetem Gazdaság és Társadalomtudományi Kar	http://www.szie.hu/karok/gtk.html
SZIE KTK	Szent István Egyetem Kertészettudományi Kar	http://www.kee.hu/ktk/
SZIE KVA	Szent István Egyetem Gazdaság és Társadalomtudományi Kar Kereskedelemi és Vállalkozási Akadémia	http://www.kva.edu.hu/
SZIE MKK	Szent István Egyetem Mezőgazdaság- és Környezettudományi Kar	http://www.mkk.szie.hu/nyitolap/
SZIE TK	Szent István Egyetem Tájépítészeti -védelmi és -fejlesztési Kar	http://www.kee.hu/tajkar/
SZTE	Szegedi Tudományegyetem	http://www.u-szeged.hu/
SZTE ÁJK	Szegedi Tudományegyetem Állam és Jogtudományi Kar	http://www.juris.u-szeged.hu/index2.html
SZTE ÁOK	Szegedi Tudományegyetem Általános Orvostudományi Kar	http://www.szote.u-szeged.hu/aok/main.htm
SZTE BTK	Szegedi Tudományegyetem Bölcsészettudományi Kar	http://www.arts.u-szeged.hu/
SZTE GTK	Szegedi Tudományegyetem Gazdaságtudományi Kar	http://www.eco.u-szeged.hu/index/index.html
SZTE GYTK	Szegedi Tudományegyetem Gyógyszerésztudományi Kar	http://www.szote.u-szeged.hu/gytk/
SZTE TTK	Szegedi Tudományegyetem Természettudományi Kar	http://www.sci.u-szeged.hu/
VE	Veszprémi Egyetem	www.vein.hu
VE GMK	Veszprémi Egyetem Georgikon Mezőgazdaságtudományi Kar	http://www.georgikon.hu/
VE MK	Veszprémi Egyetem Mérnöki Kar	http://www.vein.hu/vein/karok/mk/index.php
VE TK	Veszprémi Egyetem Tanárképző Kar	http://www.vein.hu/karok/tk/index.php
VE MIK	Veszprémi Egyetem Műszaki Informatikai Kar	http://mik.vein.hu/
VE GTK	Veszprémi Egyetem Gazdaságtudományi Kar	http://www.vein.hu/www/karok/gtk/index.html
VIKKK	Vegyészmérnöki Intézet Kooperációs Kutatási Központ	http://vikkk.vein.hu/
ZMNE	Zrínyi Miklós Nemzetvédelmi Egyetem	http://www.zmka.hu/, http://www.zmne.hu/
ATF	Adventista Teológiai Főiskola	http://www.adventista.hu/index.html
ÁVF	Általános Vállalkozási Főiskola	http://www.avf.hu/
AVKF	Apor Vilmos Katolikus Főiskola	http://www.avkf.hu/
BDF	Berzsenyi Dániel Főiskola	http://www.bdtf.hu/
BDTF	Berzsenyi Dániel Tanárképző Főiskola	http://www.bdtf.hu/ már nincs ilyen
BEPF	Benedek Elek Pedagógiai Főiskola	http://www.bepf.hu/ már nincs
BGF	Budapesti Gazdasági Főiskola	http://www.bgf.hu/

Hungarian Higher Educational Institutions		
Acronym	**Full Name**	**Home Page URL**
BGF KKFK	Budapesti Gazdasági Főiskola Külkereskedelmi Főiskolai Kar	http://www.kkf.hu/
BGF KVIFK	Budapesti Gazdasági Főiskola Kereskedelmi Vendéglátóipari és Idegenforgalmi Főiskolai Kar	http://www.kvif.hu/
BGF PSZFK	Budapesti Gazdasági Főiskola Pénzügyi és Számviteli Főiskolai Kar	http://www.pszfb.hu/
BKÁE ÁFK	Budapesti Közgazdaságtudományi és Államigazgatási Egyetem Államigazgatási Főiskolai Kar	http://www.bkae.hu/subpage.php?org=4
BKF	Budapesti Kommunikációs Főiskola	http://www.bkf.hu/fooldal.php
BMF	Budapesti Műszaki Főiskola	http://www.bmf.hu/
BMF BGK	Budapesti Műszaki Főiskola Bánki Donát Gépészmérnöki Főiskolai Kar	http://www.banki.hu/
BMF KGK	Budapesti Műszaki Főiskola Keleti Károly Gazdasági Főiskolai Kar	http://www.kgk.bmf.hu/
BMF KVK	Budapesti Műszaki Főiskola Kandó Kálmán Villamosmérnöki Főiskolai Kar	http://www.kando.hu/
BMF NIK	Budapesti Műszaki Főiskola Neumann János Informatikai Főiskolai Kar	http://www.nik.hu/
BMF RKK	Budapesti Műszaki Főiskola Rejtő Sándor Könnyűipari Mérnöki Főiskolai Kar	http://www.kmf.hu/
DE EFK	Debreceni Egyetem Egészségügyi Főiskolai Kar	http://www.doteefk.hu/
DE HWPFK	Debreceni Egyetem Hajdúböszörményi Wargha István Pedagógiai Főiskolai Kar	http://www.hwpf.hu/
DE MFK	Debreceni Egyetem Műszaki Főiskolai Kar	http://www.tech.klte.hu/
DF	Dunaújvárosi Főiskola	http://www.poliod.hu/
EGHF	Egri Hittudományi Főiskola	http://www.eghf.hu/
EJF	Eötvös József Főiskola	http://www.ejf.hu/
EKF	Eszterházy Károly Főiskola	http://www.ektf.hu/index.php
ELTE GYFK	Eötvös Loránd Tudományegyetem Bárczi Gusztáv Gyógypedagógiai Főiskolai Kar	http://www.barczi.hu/
ELTE TFK	Eötvös Loránd Tudományegyetem Tanárképző Főiskolai Kar	http://iki.elte.hu/tfk/
ELTE TOFK	Eötvös Loránd Tudományegyetem Tanító és Óvóképző Főiskolai Kar	http://www.tofk.elte.hu/
ESZHF	Esztergomi Hittudományi Főiskola	http://www.ehf.hu/HONLAP/index.php
GAMF	Gépipari és Automatizálási Főiskola volt kecskeméti főisk kara	http://www.gamf.hu/
GDF	Gábor Dénes Főiskola	www.gdf.hu
GYHF	Győri Hittudományi Főiskola	http://www.gyhf.hu/
HFF	Heller Farkas Gazdasági és Turisztikai Szolgáltatások Főiskolája	http://www.hff.hu/
KE CSPFK	Kaposvári Egyetem Csokonai Vitéz Mihály Padagógiai Főiskolai Kar	http://www.csvmtkf.hu/
KF	Kecskeméti Főiskola	http://www.kefo.hu/

Hungarian Higher Educational Institutions		
Acronym	**Full Name**	**Home Page URL**
KF KFK	Kecskeméti Főiskola Kertészeti Főiskolai Kar	http://www.kfk.hu/
KF TFK	Kecskeméti Főiskola Tanítóképző Főiskolai Kar	http://www.ketif.hu/
KF MFK	Kecskeméti Főiskola Műszaki Főiskolai Kar	http://www.kefo.hu/muszaki.php
KFRTKF	Kölcsey Ferenc Református Tanítóképző Főiskola	http://www.kfrtkf.hu/
KJF	Kodolányi János Főiskola	http://www.kodolanyi.hu/
KRE TFK	Károli Gáspár Református Egyetem Tanítóképző Főiskolai Kar	http://www.reftkn.hu/
KTIF	Kölcsey Ferenc Református Tanítóképző Főiskola	http://www.kfrtkf.hu/
ME CTFK	Miskolci Egyetem Comenius Tanítóképző Főiskolai Kar	http://www.ctif.hu/
MTF	Magyar Táncművészeti Főiskola	http://www.mtf.hu/
MÜTF	Modern Üzleti Tudományok Főiskolája	http://www.mutf.hu/index.php
NYF	Nyíregyházi Főiskola	http://www.nyf.hu/
NÜF	Nemetközi Üzleti Főiskola	http://www.ibs-b.hu/index.html
NYF MMFK	Nyíregyházi Főiskola Műszaki és Mezőgazdasági Főiskolai Kar	http://mmfk.nyf.hu/
NYF TTFK	Nyíregyházi Főiskola Természettudományi Főiskolai Kar	http://zeus.nyf.hu/~nyilasi/
NYME ATFK	Nyugat Magyarországi Egyetem Apáczai Csere János Tanítóképző Főiskolai Kar	http://www.atif.hu/
NYME BPFK	Nyugat Magyarországi Egyetem Benedek Elek Pedagógiai Főiskolai Kar	http://www.bepf.hu/
NYME GEO	Nyugat Magyarországi Egyetem Geoinformatikai Főiskolai Kar	http://www.cslm.hu/
PHF	Pécsi Püspöki Hittudományi Főiskola	http://www.pphf.hu/
PMMF	Pollack Mihály Műszaki Főiskola	már nem PTE kara lett http://www.pmmf.hu/
PSZF	Pénzügyi és Számviteli Főiskola	már nem BGF kara lett http://www.pszfb.hu/
PTE EFK	Pécsi Tudományegyetem Egészségügyi Főiskolai Kar	www.efk.pte.hu/
PTE IFK	Pécsi Tudományegyetem Illyés Gyula Főiskolai Kar	http://www.igyfk.pte.hu/
PTE PMMFK	Pécsi Tudományegyetem Pollack Mihály Műszaki Főiskolai Kar	http://www.pmmf.hu/
PTF	Pünkösdi Teológiai Főiskola	http://www.ptf.hu/
RTF	Rendőrtiszti Főiskola	http://193.6.238.67/rtf/index_elemei/slide0001.htm
SE EFK	Semmelweis Egyetem Egészségügyi Főiskolai Kar	http://www.sote.hu/oktatas/efk/
SSTF	Sola Scriptura Lelkészképző és Teológiai Főiskola	http://www.sola.hu/
SSZHF	Sapientia Szerzetesi Hittudományi Főiskola	http://www.sapientia.hu/
SZAGKHF	Szent Atanáz Görög Katolikus Hittudományi Főiskola	http://www.atanaz.hu/foisk/

Hungarian Higher Educational Institutions		
Acronym	**Full Name**	**Home Page URL**
SZF	Szolnoki Főiskola	http://www.szolf.hu/
SZHF	Szegedi Hittudományi Főiskola	http://www.theol.u-szeged.hu/
SZIE GMFK	Szent István Egyetem Gazdálkodási és Mezőgazdasági Főiskolai Kar	http://www.gyfk.hu/uj/
SZIE JFK	Szent István Egyetem Jászberényi Főiskolai Kar	http://www.jtkf.hu/
SZIE YMMFK	Szent István Egyetem Ybl Miklós Műszaki Főiskolai Kar	http://www.ymmf.hu/
SZOFI	Szolnoki Főiskola	http://www.szolf.hu/
SZTE EFK	Szegedi Tudományegyetem Egészségügyi Főiskolai Kar	http://www.efk.u-szeged.hu/
SZTE JTFK	Szegedi Tudományegyetem Juhász GyulaTanárképző Főiskolai Kar	http://www.jgytf.u-szeged.hu/
SZTE MFK	Szegedi Tudományegyetem Mezőgazdasági Főiskolai Kar	http://www.mfk.u-szeged.hu/
SZTE SZÉFK	Szegedi Tudományegyetem Szegedi Élelmiszeripari Főiskolai Kar	http://www.szef.u-szeged.hu/
TF	SE Testnevelési és Sporttudományi Kar volt Testnevelési Főiskola	http://www.sote.hu/oktatas/tf/
TKBF	A Tan Kapuja Buddhista Főiskola	http://www.tkbf.hu/
TSF	Tessedik Sámuel Főiskola	http://www.kf.hu/
TSF GFK	Tessedik Sámuel Főiskola Gazdasági Főiskolai Kar	http://www.tsf.hu/new/index.php?a=1
TSF MFK	Tessedik Sámuel Főiskola Mezőgazdaasági Főiskolai Kar, Mezőtúr	http://www.mfk.hu/
TSF MVK	Tessedik Sámuel Főiskola Mezőgazdasági Víz és Környezetgazdálkodási Kar	http://www.mvk.tsf.hu/
TSF PFK	Tessedik Sámuel Főiskola Pedagógiai Főiskolai Kar	http://www.szv.kf.hu/pfk/
VHF	Veszprémi Érseki Hittudományi Főiskola	http://www.vhf.hu/vhf/text/index.html
VTIF	Vitéz János Római Katolikus Tanítóképző Főiskola	http://www.vjrktf.hu/
WJLF	Wesley János Lelkészképző Főiskola	http://www.wesley.hu/
ZSKF	Zsigmond Király Főiskola	http://www.zskf.hu/
BTA	Bptista Teológiai Akadémia	http://www.bta.hu/
MFA	Martineum Felnőttképző Akadémia	http://www.martineum.hu/
PRTA	Pápai Református Teológiai Akadémia	http://www.papacollege.hu/akademia/
SRTA	Sárospataki Református Teológiai Akadémia	http://www.srta.hu/
SZPA	Szent Pál Akadémia	www.szpa.hu
ZMKA	Zrínyi Miklós Katonai Akadémia	http://www.zmka.hu/
MTA	Magyar Tudományos Akadémia	http://mta.administrator.hu/
MPANNI	Mozgássérültek Pető András Nevelőképző és Nevelőintézete	http://www.peto.hu/
ELO	Európai Levelező Oktatás	http://www.elo.hu/

A.3 Acronyms and Home Pages of Danish Higher Educational Institutions

Danish Higher Educational Institutions		
Acronym	**Full Name**	**Home Page URL**
AAU	Aalborg Universitet	http://ekstern.aau.dk/ www.aau.dk
AAA	Arkitektskolen Aarhus / Aarhus School of Architecture	http://aarch.dk/
AH	Aalborg Handelsskole	http://www.ah.dk/
ASB	Handelshøjskolen i Århus / Aarhus School of Business	http://www.asb.dk/
AU	Aarhus Universitet	http://www.au.dk/
CBS	Copenhagen Business School	http://www.cbs.dk/
DFM	Det Fynske Musikkonservatorium	http://www.dfm.dk/
DFU	Danmarks Farmaceutiske Universitet	http://www.dfuni.dk/ www.dfh.dk
DGH	Den Grafiske Højskole	http://www.dgh.dk/
DJH	Danmarks Journalisthøjskole	http://www.djh.dk/
DKDM	Det Kongelige Danske Musikkonservatorium	http://www.dkdm.dk/
DNS	DNS International Teacher Training College	http://www.dns-tvind.dk/
DPU	Danmarks Pædagogiske Universitet	http://www.dpu.dk/ www.dnsdk.dk
DSH	Den Sociale Højskole i Aarhus	http://www.dsh-aa.dk/
DTU	Danmarks Tekniske Universitet	http://www.dtu.dk/
HIH	Handels- og IngeniørHøjskolen	http://www.hih.dk/Default.aspx
IHA	Ingeniørhøjskolen i Århus	http://www.iha.dk/Default.aspx?ID=19
IHK	Ingeniørhøjskolen i København	http://www.ihk.dk/
IHS	Idræthøjskolen i Sønderborg	http://www.ihs.dk/
IOT	Ingeniørhøjskolen Odense Teknikum	http://www.iot.dk/
ITU	IT-Universitetet i København / IT University of Copenhagen	http://www1.itu.dk/
KSS	Københavns Socialpædagogiske Seminarium	http://www.kssem.dk/
KVL	Den Kgl. Veterinær- og Landbohøjskole	http://www.kvl.dk/
NNS	Nørre Nissum Seminarium	http://Web.nns.dk/
OSS	Odense Socialpædagogiske Seminarium	http://www.oss-fyn.dk/
RMC	Rytmisk Musikkonservatorium	http://www.rmc.dk/
RUC	Roskilde Universitetscenter	http://www.ruc.dk/ruc/
SDU	Syddansk Universitet	http://www.ou.dk/ http://www.sdu.dk/
VMK	Vestjysk Musikkonservatorium	http://www.vmk.dk/1/front.asp

A.4 Acronyms and Home Pages of Hungarian and Danish parties

Hungarian parties		
Acronym	**Full Name**	**Home Page URL**
IDE	Internetes Demokrácia Pártja	http://ide-ide.hu/
MNYP	Magyar Nyugdíjasok Pártja	http://www.nyugdijasokma.hu/

Hungarian parties		
Acronym	**Full Name**	**Home Page URL**
MIÉP	Magyar Igazság és Élet Pártja	http://www.miep.hu/
MNYNP	Magyar Nemzeti Nép Párt	http://gportal.hu/portal/mnnp/
MVPP	Magyar Vidék és Polgari Párt	http://www.mvpp.hu/
NDP	Nemzeti Demokrata Párt	http://www.datanet.hu/ndp/
PPSZ	Polgárok és Polgármesterek Szövetsége az Élhető Magyarországért Párt	http://www.ppsz.hu/
SZDP	Szociáldemokrata Párt	http://www.szdp.hu/
ZD	Zöld Demokraták Szövetsége	http://www.zd.hu/
ZP	ZÖLDEK PÁRTJA	http://zoldekpartja.hu/
FIDESZ	Fidesz - Magyar Polgári Szövetség	http://www.fidesz.hu/
KDNP	KERESZTÉNYDEMOKRATA NÉPPÁRT	http://www.kdnp.hu/
MDF	Magyar Demokrata Fórum	http://www.mdf.hu/
MSZP	Magyar Szocialista Párt	http://www.mszp.hu/
SZDSZ	Szabad Demokraták Szövetsége	http://www.szdsz.hu/
Danish parties		
Acronym	**Full Name**	**Home Page URL**
CD	Centrum Demokraterne	http://www.centrumdemokraterne.dk/
DF	Dansk Folkeparti	http://www.danskfolkeparti.dk/
FrP	Fremskridtspartiet	http://www.frp.dk/
KF	DET KONSERVATIVE FOLKEPARTI	http://www.konservative.dk/
KPD	Kommunistisk Parti i Danmark	http://www.kpid.dk/Ny/index.htm
RV	Det Radikale Venstre	http://www.radikale.dk/
SD	Socialdemokraterne	http://socialdemokratiet.dk/
SF	Socialistisk Folkeparti	http://www.sf.dk/

A.5 Acronyms and Home Pages of Hungarian Government Offices

Acronym	**Full Name**	**Home Page URL**
GKM	Gazdasági és Közlekedési Minisztérium	http://www.gkm.gov.hu/
HM	Honvédelmi Minisztérium	www.honvedelem.hu
IRM	Igazságügyi és Rendészeti Minisztérium	www.im.hu
FVM	Földművelésügyi és Vidékfejlesztési Minisztérium	www.fvm.hu
OKM	Oktatási és Kulturális Minisztérium	www.om.hu
PM	Pénzügyminisztérium	www.penzugyminiszterium.hu
KvVM	Környezetvédelmi és Vízügyi Minisztérium	www.kvvm.hu
KÜM	Külügyminisztérium	www.kum.hu
MeH	Miniszterelnöki Hivatal	www.meh.hu
ÖTM	Önkormányzati és Területfejlesztési Minisztérium	www.b-m.hu
EÜM	Egészségügyi Minisztérium	www.eum.hu
SZMM	Szociális és Munkaügyi Minisztérium	www.fmm.gov.hu

APPENDIX B: EXPERIMENTAL RESULTS

In this section, the results obtained in the applications are presented in details. The appendix is organised as follows:

- Table B.1 contains the results obtained in *Section 6.4*. The table shows the identification capability of the acronyms of Hungarian general institutions in 2002.
- Table B.2 contains the results obtained in *Section 6.5*. The table shows the identification capability of the acronyms of Hungarian government offices in 2006.
- Table B.3 contains the results obtained in *Section 6.6.1*. The table shows the identification capability of the acronyms of Hungarian higher educational institutions in 2004.
- Table B.4 contains the results obtained in *Section 6.6.2*. The table shows the identification capability of the acronyms of Hungarian higher educational institutions in 2006.
- Table B.5 contains the results obtained in *Section 6.7.1*. The table shows the identification capability of the acronyms of Danish higher educational institutions in 2005.
- Table B.6 contains the results obtained in *Section 6.7.2*. The table shows the identification capability of the acronyms of Danish higher educational institutions in 2006.
- Table B.7 contains the results obtained in *Section 6.8*. The table shows the identification capability of the acronyms of Hungarian parties in 2006.
- Table B.8 contains the results obtained in *Section 6.8*. The table shows the identification capability of the acronyms of Danish parties in 2006.

Table B.1 Results of *Section 6.4*. The identification capability of the acronyms of Hungarian general institutions in 2002.

	Pseudo Precision			Pseudo Rank						Mean Pseudo Rank		
Acronym	**Hungarian**	**general**	**total**	**Heuréka**	**Alta Vizsla**	**Ariadnet**	**Google**	**Metacrawler**	**AltaVista**	**total**	**Hungarian**	**general**
APEH	0.67	1.00	0.83	0.50	0.50	1.00	1.00	1.00	1.00	0.83	0.67	1.00
ÁNTSZ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.25	0.88	1.00	0.75
ÁPV	0.33	1.00	0.67	0.13	0.06	1.00	1.00	1.00	1.00	0.70	0.40	1.00
BKIK	0.33	1.00	0.67	0.25	0.13	1.00	1.00	1.00	1.00	0.73	0.46	1.00
BKV	0.33	1.00	0.67	0.25	0.00	1.00	1.00	1.00	0.50	0.63	0.42	0.83
BVK	0.33	0.67	0.50	0.50	0.50	0.10	1.00	0.00	1.00	0.52	0.37	0.67
DOSZ	0.67	1.00	0.83	0.50	1.00	0.50	1.00	1.00	1.00	0.83	0.67	1.00
FVF	1.00	1.00	1.00	0.10	0.50	0.20	1.00	0.25	1.00	0.51	0.27	0.75
GVH	0.67	1.33	1.00	1.00	0.50	0.33	1.00	0.33	1.00	0.69	0.61	0.78
HíF	0.67	0.33	0.50	0.13	0.00	1.00	0.25	0.00	1.00	0.40	0.38	0.42
KOMA	0.67	0.67	0.67	1.00	0.00	1.00	1.00	0.00	1.00	0.67	0.67	0.67
MÁV	0.33	1.00	0.67	0.25	0.50	1.00	1.00	1.00	0.33	0.68	0.58	0.78
MBH	0.33	0.67	0.50	0.10	0.00	0.00	1.00	0.00	1.00	0.35	0.03	0.67
MEH	0.33	0.67	0.50	0.33	0.50	0.00	0.10	0.00	0.25	0.20	0.28	0.12
MÉK	0.67	0.33	0.50	0.20	0.25	0.17	0.00	0.00	0.33	0.16	0.21	0.11
MGYK	0.67	1.00	0.83	1.00	0.00	1.00	1.00	1.00	1.00	0.83	0.67	1.00
MAHART	0.00	0.00	0.00	0.00	0.25	0.50	0.50	0.17	0.50	0.32	0.25	0.39
MKIK	0.67	1.00	0.83	1.00	0.00	1.00	1.00	1.00	1.00	0.83	0.67	1.00
MKVK	0.67	1.00	0.83	1.00	0.06	1.00	1.00	1.00	1.00	0.84	0.69	1.00
AMC	0.33	0.67	0.50	0.00	0.25	0.17	1.00	0.00	1.00	0.40	0.14	0.67
MALÉV	0.33	1.00	0.67	0.25	0.25	1.00	1.00	1.00	1.00	0.75	0.50	1.00
MOL	0.67	1.00	0.83	0.11	0.17	1.00	1.00	1.00	1.00	0.71	0.43	1.00
MOKK	0.33	0.33	0.33	0.25	0.10	0.05	0.00	0.00	1.00	0.23	0.13	0.33
MOK	0.67	1.00	0.83	0.50	0.50	0.20	1.00	0.14	1.00	0.56	0.40	0.71
MÖB	0.33	1.00	0.67	0.50	0.50	0.20	1.00	1.00	1.00	0.70	0.40	1.00
MSZH	0.33	1.00	0.67	0.50	0.25	1.00	1.00	1.00	1.00	0.79	0.58	1.00
MSZT	0.67	1.00	0.83	0.11	0.25	0.50	1.00	1.00	1.00	0.64	0.29	1.00
matáv	0.67	1.00	0.83	1.00	0.13	1.00	1.00	1.00	1.00	0.85	0.71	1.00

	Pseudo Precision			Pseudo Rank						Mean Pseudo Rank		
Acronym	**Hungarian**	**general**	**total**	**Heuréka**	**Alta Vizsla**	**Ariadnet**	**Google**	**Metacrawler**	**AltaVista**	**total**	**Hungarian**	**general**
MTA	0.33	0.67	0.50	0.13	0.10	0.14	1.00	0.00	1.00	0.39	0.12	0.67
MVA	0.67	0.67	0.67	1.00	0.00	0.17	0.00	0.14	1.00	0.38	0.39	0.38
MVK	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MKGI	0.67	1.00	0.83	0.08	0.17	1.00	1.00	1.00	1.00	0.71	0.42	1.00
MGYOSZ	0.67	1.00	0.83	0.11	0.50	0.17	1.00	0.25	1.00	0.50	0.26	0.75
MBF	0.00	0.67	0.33	0.00	0.00	0.00	1.00	0.00	1.00	0.33	0.00	0.67
NBH	0.33	1.00	0.67	0.00	0.50	1.00	1.00	1.00	1.00	0.75	0.50	1.00
NIIF	0.67	1.00	0.83	0.33	0.25	1.00	1.00	1.00	1.00	0.76	0.53	1.00
NKA	0.33	1.00	0.67	0.17	0.50	0.50	1.00	1.00	1.00	0.69	0.39	1.00
NIOK	0.67	1.00	0.83	1.00	0.08	1.00	1.00	1.00	1.00	0.85	0.69	1.00
OEP	0.67	1.00	0.83	1.00	0.00	1.00	1.00	0.50	1.00	0.75	0.67	0.83
OFA	1.00	0.67	0.83	1.00	1.00	1.00	1.00	0.00	1.00	0.83	1.00	0.67
OKI	0.33	0.67	0.50	0.00	0.10	1.00	1.00	0.00	1.00	0.52	0.37	0.67
OMH	0.33	0.67	0.50	0.00	0.25	1.00	1.00	0.50	0.13	0.48	0.42	0.54
OMSZ	0.33	1.00	0.67	0.06	0.05	1.00	1.00	1.00	1.00	0.69	0.37	1.00
OMIKK	0.67	1.00	0.83	1.00	0.00	1.00	1.00	1.00	1.00	0.83	0.67	1.00
ONYF	0.67	0.67	0.67	1.00	0.50	0.50	1.00	0.17	1.00	0.69	0.67	0.72
ORTT	0.33	1.00	0.67	0.13	0.50	1.00	1.00	1.00	1.00	0.77	0.54	1.00
OTKA	0.33	1.00	0.67	0.50	0.50	1.00	1.00	1.00	1.00	0.83	0.67	1.00
SZF	0.33	1.00	0.67	0.00	0.00	1.00	1.00	1.00	1.00	0.67	0.33	1.00
VOSZ	0.33	0.67	0.50	0.00	0.00	1.00	1.00	1.00	0.00	0.50	0.33	0.67
ÁBAEGON	0.33	0.00	0.17	0.25	0.10	0.20	0.00	0.00	0.00	0.09	0.18	0.00
ÁB-AEGON	0.33	1.00	0.67	0.25	0.50	1.00	1.00	1.00	0.50	0.71	0.58	0.83
ÁÉB	0.33	0.67	0.50	0.00	0.50	1.00	1.00	1.00	0.00	0.58	0.50	0.67
ÁKK	0.33	0.33	0.33	0.00	0.00	0.20	0.00	0.10	0.00	0.05	0.07	0.03
ÁSZ	0.33	0.67	0.50	0.00	0.00	0.33	0.25	0.10	0.00	0.11	0.11	0.12
BÉT	0.33	1.00	0.67	0.25	0.00	1.00	1.00	1.00	1.00	0.71	0.42	1.00
CIB	0.33	1.00	0.67	0.00	0.13	0.25	1.00	0.14	1.00	0.42	0.13	0.71
EXIM	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	Pseudo Precision			Pseudo Rank						Mean Pseudo Rank		
Acronym	**Hungarian**	**general**	**total**	**Heuréka**	**Alta Vizsla**	**Ariadnet**	**Google**	**Metacrawler**	**AltaVista**	**total**	**Hungarian**	**general**
FHB	0.67	0.67	0.67	0.10	0.17	0.33	1.00	0.33	0.00	0.32	0.20	0.44
HB	0.00	0.33	0.17	0.00	0.00	0.00	0.10	0.00	0.20	0.05	0.00	0.10
HVB	0.67	0.67	0.67	0.20	0.50	0.33	0.10	0.00	1.00	0.36	0.34	0.37
HVG	0.67	0.67	0.67	0.33	0.05	0.33	0.25	0.17	1.00	0.36	0.24	0.47
KHB	0.33	1.00	0.67	0.00	0.25	0.20	1.00	0.33	1.00	0.46	0.15	0.78
K& H	0.67	0.67	0.67	0.00	0.50	0.20	1.00	0.50	0.00	0.37	0.23	0.50
KSH	0.67	1.00	0.83	1.00	0.00	1.00	1.00	0.50	1.00	0.75	0.67	0.83
MABISZ	1.00	1.00	1.00	0.14	1.00	1.00	1.00	1.00	1.00	0.86	0.71	1.00
MEHIB	0.67	1.00	0.83	1.00	0.25	1.00	1.00	1.00	1.00	0.88	0.75	1.00
MFB	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.13	1.00	0.85	1.00	0.71
MKB	0.67	1.00	0.83	1.00	0.50	1.00	1.00	0.33	1.00	0.81	0.83	0.78
MNB	0.33	1.00	0.67	0.00	0.06	1.00	1.00	1.00	1.00	0.68	0.35	1.00
OTP	0.33	1.00	0.67	0.50	0.08	0.50	1.00	0.11	1.00	0.53	0.36	0.70
PSZÁF	0.33	1.00	0.67	0.00	0.00	1.00	1.00	1.00	1.00	0.67	0.33	1.00
VG	0.00	0.33	0.17	0.00	0.00	0.00	0.33	0.00	0.00	0.06	0.00	0.11
VPOP	0.00	1.00	0.50	0.50	0.00	0.13	1.00	0.33	1.00	0.49	0.21	0.78
FVM	0.33	1.00	0.67	0.00	0.13	0.14	0.13	0.33	1.00	0.29	0.09	0.49
GM	0.33	0.67	0.50	0.00	0.00	0.50	1.00	0.00	1.00	0.42	0.17	0.67
HM	0.00	0.67	0.33	0.00	0.00	0.05	1.00	0.00	0.33	0.23	0.02	0.44
ISM	0.67	1.00	0.83	0.50	0.06	1.00	1.00	0.33	1.00	0.65	0.52	0.78
BM	0.00	0.33	0.17	0.00	0.00	0.13	0.00	0.00	1.00	0.19	0.04	0.33
IM	0.00	0.67	0.33	0.00	0.00	0.00	1.00	0.00	1.00	0.33	0.00	0.67
IHM	0.67	0.33	0.50	0.00	0.10	0.13	1.00	0.00	0.00	0.20	0.08	0.33
KvVM	0.00	0.67	0.33	0.00	0.06	0.00	0.25	0.13	0.00	0.07	0.02	0.13
NKÖM	0.33	1.00	0.67	0.17	0.06	1.00	1.00	1.00	1.00	0.70	0.41	1.00
OM	0.33	0.33	0.33	0.00	0.00	1.00	1.00	0.00	0.50	0.42	0.33	0.50
PM	0.00	0.67	0.33	0.00	0.00	0.00	0.50	0.00	0.33	0.14	0.00	0.28
FMM	0.00	0.33	0.17	0.08	0.00	0.00	0.17	0.00	0.00	0.04	0.03	0.06
SZCSM	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	Pseudo Precision			Pseudo Rank						Mean Pseudo Rank		
Acronym	**Hungarian**	**general**	**total**	**Heuréka**	**Alta Vizsla**	**Ariadnet**	**Google**	**Metacrawler**	**AltaVista**	**total**	**Hungarian**	**general**
MeH	0.33	1.00	0.67	1.00	0.00	0.06	1.00	1.00	1.00	0.68	0.35	1.00
MTI	0.33	1.00	0.67	0.50	0.50	1.00	1.00	0.14	1.00	0.69	0.67	0.71
TÁRKI	0.67	1.00	0.83	1.00	0.10	1.00	1.00	1.00	1.00	0.85	0.70	1.00
FKGP	0.33	1.00	0.67	0.00	0.08	1.00	1.00	1.00	1.00	0.68	0.36	1.00
MDF	0.33	1.00	0.67	0.07	0.50	1.00	1.00	1.00	1.00	0.76	0.52	1.00
MIÉP	0.33	1.00	0.67	0.00	0.00	1.00	1.00	1.00	1.00	0.67	0.33	1.00
MSZP	0.67	1.00	0.83	0.50	0.10	1.00	1.00	1.00	1.00	0.77	0.53	1.00
NDP	1.00	1.00	1.00	0.20	0.25	0.17	1.00	1.00	1.00	0.60	0.21	1.00
SZDSZ	0.33	1.00	0.67	0.10	0.50	1.00	1.00	1.00	1.00	0.77	0.53	1.00
SZDP	1.00	0.67	0.83	1.00	0.20	1.00	1.00	1.00	0.00	0.70	0.73	0.67
BGF	1.00	1.00	1.00	0.25	0.33	0.17	1.00	0.10	1.00	0.48	0.25	0.70
BKE	0.33	0.67	0.50	0.50	0.00	0.33	1.00	0.33	0.50	0.44	0.28	0.61
BKÁE	0.67	0.67	0.67	1.00	0.50	1.00	1.00	1.00	0.25	0.79	0.83	0.75
BME	0.33	0.67	0.50	0.17	0.10	1.00	1.00	0.50	0.50	0.54	0.42	0.67
BMGE	0.67	0.67	0.67	0.20	0.00	0.50	0.00	0.10	1.00	0.30	0.23	0.37
BMF	0.33	0.67	0.50	0.00	0.17	0.50	1.00	0.00	1.00	0.44	0.22	0.67
DE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ELTE	0.33	0.67	0.50	0.50	0.50	1.00	1.00	0.50	0.50	0.67	0.67	0.67
KÉE	0.67	1.00	0.83	0.25	1.00	0.10	0.50	0.50	0.50	0.48	0.45	0.50
MKE	0.33	0.33	0.33	0.00	0.00	0.25	0.10	0.00	1.00	0.23	0.08	0.37
NyME	0.33	1.00	0.67	0.25	0.17	0.17	1.00	0.20	0.50	0.38	0.19	0.57
PTE	0.33	0.67	0.50	0.25	0.13	0.14	1.00	0.00	1.00	0.42	0.17	0.67
SOTE	0.67	1.00	0.83	0.50	0.13	0.20	1.00	1.00	0.50	0.55	0.28	0.83
SZIE	0.33	1.00	0.67	0.13	0.06	1.00	1.00	1.00	1.00	0.70	0.39	1.00
SZE	0.33	0.33	0.33	0.25	0.50	0.33	0.13	0.00	1.00	0.37	0.36	0.38
SZTE	0.67	1.00	0.83	1.00	0.00	0.10	1.00	1.00	1.00	0.68	0.37	1.00
VE	0.00	0.33	0.17	0.00	0.13	0.00	0.13	0.00	0.33	0.10	0.04	0.15
ZMNE	0.67	1.00	0.83	0.50	0.50	0.25	1.00	0.50	1.00	0.63	0.42	0.83

Table B.2 Results of *Section 6.5*. The identification capability of the acronyms of Hungarian government offices in 2006.

Acronym	Pseudo Precision	% Pseudo Precision						Pseudo Rank						Mean Pseudo Rank	% Mean Pseudo Rank
		Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo	MSN	Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo	MSN		
GKM	0.67	0.670	0.000	0.060	0.025	0.004	0.000	1.00	0.08	0.13	1.00	0.25	0.00	0.41	0.72
HM	0.33	0.670	0.000	0.000	0.025	0.000	0.000	0.50	0.00	0.00	1.00	0.00	0.00	0.25	0.36
IRM	0.33	0.000	0.240	0.000	0.025	0.000	0.000	0.00	0.11	0.00	0.50	0.00	0.00	0.10	0.04
FVM	1.00	0.670	0.240	0.060	0.025	0.004	0.004	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
OKM	0.33	0.670	0.000	0.000	0.025	0.000	0.000	0.20	0.00	0.00	0.25	0.00	0.00	0.08	0.14
PM	0.33	0.000	0.000	0.060	0.025	0.000	0.000	0.00	0.00	1.00	0.50	0.00	0.00	0.25	0.07
KvVM	0.83	0.670	0.000	0.060	0.025	0.004	0.004	1.00	0.00	0.11	1.00	1.00	1.00	0.69	0.71
KÜM	0.83	0.670	0.240	0.060	0.025	0.004	0.000	1.00	0.33	0.33	0.20	0.33	0.00	0.37	0.78
MeH	1.00	0.670	0.240	0.060	0.025	0.004	0.004	1.00	1.00	0.50	0.50	0.14	0.13	0.54	0.95
ÖTM	0.33	0.670	0.240	0.000	0.000	0.000	0.000	1.00	0.14	0.00	0.00	0.00	0.00	0.19	0.70
EÜM	0.83	0.670	0.000	0.060	0.025	0.004	0.004	1.00	0.00	0.25	1.00	1.00	0.50	0.63	0.71
SZMM	0.67	0.670	0.240	0.000	0.025	0.000	0.004	0.50	0.20	0.00	0.20	0.06	0.25	0.20	0.39

Table B.3 Results of *Section 6.6.1*. The identification capability of the acronyms of Hungarian higher educational institutions in 2004.

	Search Engine Rankings						Pseudo Precision			Pseudo Rank						Mean Pseudo Rank		
Acronym	Google	Altavista	Metacrawler	Ariadnet	Altavizsla	Heuréka	general	Hungarian	total	Google	Altavista	Metacrawler	Ariadnet	Altavizsla	Heuréka	general	Hungarian	total
BKÁE TK	b*2	b*9	0	b*7	3	b*2	0.00	0.33	0.17	0.25	0.06	0.00	0.07	0.33	0.25	0.10	0.22	0.16
BKÁE	1	1	0	7	1	2	0.67	1.00	0.83	1.00	1.00	0.00	0.14	1.00	0.50	0.67	0.55	0.61
BKÁE GTK	b*2	0	0	b*4	6	b*3	0.00	0.33	0.17	0.25	0.00	0.00	0.13	0.17	0.17	0.08	0.15	0.12
BKÁE KTK	0	0	0	0	b*2	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.08	0.04
BKE	2	2	2	6	1	b*1	1.00	0.67	0.83	0.50	0.50	0.50	0.17	1.00	0.50	0.50	0.56	0.53
BME	0	1	3	b*3	1	b*2	0.67	0.33	0.50	0.00	1.00	0.33	0.17	1.00	0.25	0.44	0.47	0.46
BME ÉSZK	b*1	b*1	0	b*2	1	b*1	0.00	0.33	0.17	0.50	0.50	0.00	0.25	1.00	0.50	0.33	0.58	0.46
BME ÉÖK	b*3	b*3	0	b*2	b*1	b*1	0.00	0.00	0.00	0.17	0.17	0.00	0.25	0.50	0.50	0.11	0.42	0.26
BME GÉK	b*1	b*3	0	0	b*1	0	0.00	0.00	0.00	0.50	0.17	0.00	0.00	0.50	0.00	0.22	0.17	0.19
BME GTK	1	1	3	b*5	1	1	1.00	0.67	0.83	1.00	1.00	0.33	0.10	1.00	1.00	0.78	0.70	0.74

	Search Engine Rankings						Pseudo Precision			Pseudo Rank						Mean Pseudo Rank		
Acronym	**Google**	**Altavista**	**Metacrawler**	**Ariadnet**	**Altavizsla**	**Heuréka**	**general**	**Hungarian**	**total**	**Google**	**Altavista**	**Metacrawler**	**Ariadnet**	**Altavizsla**	**Heuréka**	**general**	**Hungarian**	**total**
BME KSK	3	1	1	b*2	b*2	b*3	1.00	0.00	0.50	0.33	1.00	1.00	0.25	0.25	0.17	0.78	0.22	0.50
BME TTK	b*1	1	b*4	b*5	9		1.00	0.33	0.67	0.50	1.00	0.13	0.10	0.11	0.13	0.54	0.11	0.33
BME VEK	b*2	b*7	b*5	0	b*3	b*8	0.00	0.00	0.00	0.25	0.07	0.10	0.00	0.17	0.06	0.14	0.08	0.11
BME VIK	2	1	2	0	b*10	0	1.00	0.00	0.50	0.50	1.00	0.50	0.00	0.05	0.00	0.67	0.02	0.34
DE	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DE ÁOK	0	0	0	0	b*2	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.08	0.04
DE BTK	0	0	0	0	b*1	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.17	0.08
DE JÁTI	b*4	b*1	0	b*1	b*1	b*1	0.00	0.00	0.00	0.13	0.50	0.00	0.50	0.50	0.50	0.21	0.50	0.35
DE K	0	0	0	0	b*4	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.04	0.02
DE KTK	0	0	0	0	b*2	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.08	0.04
DE MTK	0	0	0	0	b*2	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.08	0.04
DE TTK	0	0	0	0	b*1	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.17	0.08
DRHE	2	1	1	5	1	3	1.00	1.00	1.00	0.50	1.00	1.00	0.20	1.00	0.33	0.83	0.51	0.67
EFE	0	0	0	b*7	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.02	0.01
EHE	0	0	0	b*9	b*4	b*8	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.13	0.06	0.00	0.08	0.04
ELTE	1	1	2	b*6	1	b*2	1.00	0.33	0.67	1.00	1.00	0.50	0.08	1.00	0.25	0.83	0.44	0.64
ELTE ÁJK	b*1	b*1	0	b*2	1	9	0.00	0.67	0.33	0.50	0.50	0.00	0.25	1.00	0.11	0.33	0.45	0.39
ELTE BTK	1	1	1	6	2	b*7	1.00	0.67	0.83	1.00	1.00	1.00	0.17	0.50	0.07	1.00	0.25	0.62
ELTE TTK	1	1	8	b*4	5	b*1	1.00	0.33	0.67	1.00	1.00	0.13	0.13	0.20	0.50	0.71	0.28	0.49
ELTE PPK	1	0	4	b*6	1	0	0.67	0.33	0.50	1.00	0.00	0.25	0.08	1.00	0.00	0.42	0.36	0.39
ELTE IK	10	b*1	b*9	b*5	5	b*2	0.33	0.33	0.33	0.10	0.50	0.06	0.10	0.20	0.25	0.22	0.18	0.20
JATE	3	2	1	0	2	0	1.00	0.33	0.67	0.33	0.50	1.00	0.00	0.50	0.00	0.61	0.17	0.39
KE ÁTK	7	5	0	0	5	0	0.67	0.33	0.50	0.14	0.20	0.00	0.00	0.20	0.00	0.11	0.07	0.09
KGRE	3	1	2	2	1	0	1.00	0.67	0.83	0.33	1.00	0.50	0.50	1.00	0.00	0.61	0.50	0.56
KLTE	1	1	3	b*6	b*1	b*2	1.00	0.00	0.50	1.00	1.00	0.33	0.08	0.50	0.25	0.78	0.28	0.53
KRE	0	0	0	b*3	b*7	0	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.07	0.00	0.00	0.08	0.04
KRE ÁJK	7	1	0	b*1	9	b*2	0.67	0.33	0.50	0.14	1.00	0.00	0.50	0.11	0.25	0.38	0.29	0.33
KRE BTK	b*2	b*2	b*1	b*1	b*2	b*1	0.00	0.00	0.00	0.25	0.25	0.50	0.50	0.25	0.50	0.33	0.42	0.38
KRE HTK	b*1	b*1	b*1	b*1	b*1	b*1	0.00	0.00	0.00	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
LFZE	b*1	b*1	b*1	b*2	b*3	b*1	0.00	0.00	0.00	0.50	0.50	0.50	0.25	0.17	0.50	0.50	0.31	0.40
ME	0	0	0	0	1	0	0.00	0.33	0.17	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.33	0.17
ME ÁJK	2	1	0	b*1	2	0	0.67	0.33	0.50	0.50	1.00	0.00	0.50	0.50	0.00	0.50	0.33	0.42

	Search Engine Rankings						Pseudo Precision			Pseudo Rank						Mean Pseudo Rank		
Acronym	**Google**	**Altavista**	**Metacrawler**	**Ariadnet**	**Altavizsla**	**Heuréka**	**general**	**Hungarian**	**total**	**Google**	**Altavista**	**Metacrawler**	**Ariadnet**	**Altavizsla**	**Heuréka**	**general**	**Hungarian**	**total**
ME AKK	b*3	0	b*4	3	b*1	0	0.00	0.33	0.17	0.17	0.00	0.13	0.33	0.50	0.00	0.10	0.28	0.19
ME BTK	7	0	0	0	b*3	0	0.33	0.00	0.17	0.14	0.00	0.00	0.00	0.17	0.00	0.05	0.06	0.05
ME GÉK	5	0	0	0	0	0	0.33	0.00	0.17	0.20	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.03
ME GTK	0	b*1	0	7	b*1	0	0.00	0.33	0.17	0.00	0.50	0.00	0.14	0.50	0.00	0.17	0.21	0.19
ME MFK	b*10	0	0	0	b*1	0	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.50	0.00	0.02	0.17	0.09
MIE	0	8	0	2	1	1	0.33	1.00	0.67	0.00	0.13	0.00	0.50	1.00	1.00	0.04	0.83	0.44
MIRT	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MKE	7	3	0	6	1	0	0.67	0.67	0.67	0.14	0.33	0.00	0.17	1.00	0.00	0.16	0.39	0.27
NME	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NYME	4	1	1	b*1	2	b*1	1.00	0.33	0.67	0.25	1.00	1.00	0.50	0.50	0.50	0.75	0.50	0.63
PPKE	1	1	1	b*1	1	1	1.00	0.67	0.83	1.00	1.00	1.00	0.50	1.00	1.00	1.00	0.83	0.92
PPKE BTK	1	1	1	3	1	2	1.00	1.00	1.00	1.00	1.00	1.00	0.33	1.00	0.50	1.00	0.61	0.81
PPKE HTK	1	1	b*4	8	b*4	2	0.67	0.67	0.67	1.00	1.00	0.13	0.13	0.13	0.50	0.71	0.25	0.48
PPKE ITK	b*7	1	b*2	b*2	b*1	b*1	0.33	0.00	0.17	0.07	1.00	0.25	0.25	0.50	0.50	0.44	0.42	0.43
PPKE JÁK	b*2	b*6	b*1	1	b*1	3	0.00	0.67	0.33	0.25	0.08	0.50	1.00	0.50	0.33	0.28	0.61	0.44
PTE	0	b*5	1	5	1	b*2	0.33	0.67	0.50	0.00	0.10	1.00	0.20	1.00	0.25	0.37	0.48	0.43
PTE ÁJK	1	0	0	b*1	b*2	b*1	0.33	0.00	0.17	1.00	0.00	0.00	0.50	0.25	0.50	0.33	0.42	0.38
PTE ÁOK	1	1	0	6	1	1	0.67	1.00	0.83	1.00	1.00	0.00	0.17	1.00	1.00	0.67	0.72	0.69
PTE BTK	b*7	1	1	4	1	0	0.67	0.67	0.67	0.07	1.00	1.00	0.25	1.00	0.00	0.69	0.42	0.55
PTE FEEFI	b*7	3	0	b*1	4	7	0.33	0.67	0.50	0.07	0.33	0.00	0.50	0.25	0.14	0.13	0.30	0.22
PTE MK	1	2	1	b*9	b*1	b*9	1.00	0.00	0.50	1.00	0.50	1.00	0.06	0.50	0.06	0.83	0.20	0.52
PTE TI	b*3	0	b*4	n*4	1	b*3	0.00	0.33	0.17	0.17	0.00	0.13	0.13	1.00	0.17	0.10	0.43	0.26
PTE TTK	2	1	2	b*1	b*10	b*1	1.00	0.00	0.50	0.50	1.00	0.50	0.50	0.05	0.50	0.67	0.35	0.51
SE	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SE ÁOK	b*2	0	0	0	b*2	0	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.25	0.00	0.08	0.08	0.08
SE FOK	b*3	0	b*4	0	b*1	0	0.00	0.00	0.00	0.17	0.00	0.13	0.00	0.50	0.00	0.10	0.17	0.13
SE GYTK	b*3	0	b*2	0	b*1	0	0.00	0.00	0.00	0.17	0.00	0.25	0.00	0.50	0.00	0.14	0.17	0.15
SE TSK	b*2	0	b*2	0	b*1	0	0.00	0.00	0.00	0.25	0.00	0.25	0.00	0.50	0.00	0.17	0.17	0.17
SOTE	1	1	1	0	1	0	1.00	0.33	0.67	1.00	1.00	1.00	0.00	1.00	0.00	1.00	0.33	0.67
SZE	0	2	0	0	1	0	0.33	0.33	0.33	0.00	0.50	0.00	0.00	1.00	0.00	0.17	0.33	0.25
SZFE	b*1	b*5	b*1	b*2	b*2	b*1	0.00	0.00	0.00	0.50	0.10	0.50	0.25	0.25	0.50	0.37	0.33	0.35
SZIE	1	1	1	b*2	1	b*1	1.00	0.33	0.67	1.00	1.00	1.00	0.25	1.00	0.50	1.00	0.58	0.79

	Search Engine Rankings						Pseudo Precision			Pseudo Rank						Mean Pseudo Rank		
Acronym	**Google**	**Altavista**	**Metacrawler**	**Ariadnet**	**Altavizsla**	**Heuréka**	**general**	**Hungarian**	**total**	**Google**	**Altavista**	**Metacrawler**	**Ariadnet**	**Altavizsla**	**Heuréka**	**general**	**Hungarian**	**total**
SZIE ÁOTK	b*5	b*1	b*9	0	1	0	0.00	0.33	0.17	0.10	0.50	0.06	0.00	1.00	0.00	0.22	0.33	0.28
SZIE ÉTK	b*3	b*1	0	b*1	b*1	b*1	0.00	0.00	0.00	0.17	0.50	0.00	0.50	0.50	0.50	0.22	0.50	0.36
SZIE GÉK	b*3	0	0	0	b*1	b*2	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.50	0.25	0.06	0.25	0.15
SZIE GTK	b*1	b*6	b*1	b*3	b*5	b*2	0.00	0.00	0.00	0.50	0.08	0.50	0.17	0.10	0.25	0.36	0.17	0.27
SZIE KTK	0	0	b*10	0	b*1	0	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.50	0.00	0.02	0.17	0.09
SZIE KVA	b*3	0	b*4	0	b*1	b*5	0.00	0.00	0.00	0.17	0.00	0.13	0.00	0.50	0.10	0.10	0.20	0.15
SZIE MKK	b*1	b*1	b*1	3	b*1	1	0.00	0.33	0.17	0.50	0.50	0.50	0.33	0.50	1.00	0.50	0.61	0.56
SZIE TK	b*6	b*6	b*2	b*9	0	0	0.00	0.00	0.00	0.08	0.08	0.25	0.06	0.00	0.00	0.14	0.02	0.08
SZTE	1	1	1	10	1	1	1.00	1.00	1.00	1.00	1.00	1.00	0.10	1.00	1.00	1.00	0.70	0.85
SZTE ÁJK	4	b*1	0	9	1	b*1	1.00	0.67	0.83	0.25	0.50	0.00	0.11	1.00	0.50	0.25	0.54	0.39
SZTE ÁOK	b*1	b*1	0	b*8	b*3	6	0.00	0.33	0.17	0.50	0.50	0.00	0.06	0.17	0.17	0.33	0.13	0.23
SZTE BTK	1	b*1	1	0	b*1	0	0.67	0.00	0.33	1.00	0.50	1.00	0.00	0.50	0.00	0.83	0.17	0.50
SZTE GTK	1	2	1	0	b*2	b*10	1.00	0.00	0.50	1.00	0.50	1.00	0.00	0.25	0.05	0.83	0.10	0.47
SZTE GYTK	b*3	b*1	b*5	0	b*2	b*6	0.00	0.00	0.00	0.17	0.50	0.10	0.00	0.25	0.08	0.26	0.11	0.18
SZTE TTK	1	1	1	0	b*1	b*2	1.00	0.00	0.50	1.00	1.00	1.00	0.00	0.50	0.25	1.00	0.25	0.63
VE	0	0	0	b*5	1	0	0.00	0.33	0.17	0.00	0.00	0.00	0.10	1.00	0.00	0.00	0.37	0.18
VE GMK	b*1	0	b*2	0	b*2	0	0.00	0.00	0.00	0.50	0.00	0.25	0.00	0.25	0.00	0.25	0.08	0.17
VE MK	b*1	0	b*4	0	b*1	0	0.00	0.00	0.00	0.50	0.00	0.13	0.00	0.50	0.00	0.21	0.17	0.19
VE TK	b*1	0	0	0	b*1	0	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.50	0.00	0.17	0.17	0.17
VE MIK	b*1	0	b*2	0	0	0	0.00	0.00	0.00	0.50	0.00	0.25	0.00	0.00	0.00	0.25	0.00	0.13
VE GTK	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
VIKKK	1	1	3	2	1	0	1.00	0.67	0.83	1.00	1.00	0.33	0.50	1.00	0.00	0.78	0.50	0.64
ZMNE	1	1	1	5	1	9	1.00	1.00	1.00	1.00	1.00	1.00	0.20	1.00	0.11	1.00	0.44	0.72
ATF	0	0	0	0	b*3	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.06	0.03
ÁVF	1	2	0	2	1	1	0.67	1.00	0.83	1.00	0.50	0.00	0.50	1.00	1.00	0.50	0.83	0.67
AVKF	1	1	1	2	1	1	1.00	1.00	1.00	1.00	1.00	1.00	0.50	1.00	1.00	1.00	0.83	0.92
BDF	0	0	0	8	2	b*3	0.00	0.67	0.33	0.00	0.00	0.00	0.13	0.50	0.17	0.00	0.26	0.13
BDTF	1	1	1	1	1	0	1.00	0.67	0.83	1.00	1.00	1.00	1.00	1.00	0.00	1.00	0.67	0.83
BEPF	2	1	7	b*2	1	0	1.00	0.33	0.67	0.50	1.00	0.14	0.25	1.00	0.00	0.55	0.42	0.48
BGF	3	5	10	b*2	1	b*1	1.00	0.33	0.67	0.33	0.20	0.10	0.25	1.00	0.50	0.21	0.58	0.40
BGF KKFK	b*1	b*4	b*1	6	b*1	9	0.00	0.67	0.33	0.50	0.13	0.50	0.17	0.50	0.11	0.38	0.26	0.32
BGF KVIFK	1	1	1	2	1	0	1.00	0.67	0.83	1.00	1.00	1.00	0.50	1.00	0.00	1.00	0.50	0.75

	Search Engine Rankings						Pseudo Precision			Pseudo Rank						Mean Pseudo Rank		
Acronym	**Google**	**Altavista**	**Metacrawler**	**Ariadnet**	**Altavizsla**	**Heuréka**	**general**	**Hungarian**	**total**	**Google**	**Altavista**	**Metacrawler**	**Ariadnet**	**Altavizsla**	**Heuréka**	**general**	**Hungarian**	**total**
BGF PSZFK	b*9	1	b*3	b*10	b*3	b*5	0.33	0.00	0.17	0.06	1.00	0.17	0.05	0.17	0.10	0.41	0.11	0.26
BKÁE ÁFK	b*5	0	0	0	b*4	0	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.13	0.00	0.03	0.04	0.04
BKF	0	5	0	2	1	0	0.33	0.67	0.50	0.00	0.20	0.00	0.50	1.00	0.00	0.07	0.50	0.28
BMF	b*1	5	0	b*4	1	0	0.33	0.33	0.33	0.50	0.20	0.00	0.13	1.00	0.00	0.23	0.38	0.30
BMF BGK	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
BMF KGK	1	1	1	1	2	0	1.00	0.67	0.83	1.00	1.00	1.00	1.00	0.50	0.00	1.00	0.50	0.75
BMF KVK	4	1	1	0	4	0	1.00	0.33	0.67	0.25	1.00	1.00	0.00	0.25	0.00	0.75	0.08	0.42
BMF NIK	0	1	b*1	0	9	4	0.33	0.67	0.50	0.00	1.00	0.50	0.00	0.11	0.25	0.50	0.12	0.31
BMF RKK	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DE EFK	0	4	0	0	8	0	0.33	0.33	0.33	0.00	0.25	0.00	0.00	0.13	0.00	0.08	0.04	0.06
DE HWPFK	0	0	0	b*9	0	b*5	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.10	0.00	0.05	0.03
DE MFK	0	0	0	0	b*3	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.06	0.03
DF	0	0	0	0	10	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.03	0.02
EGHF	b*4	3	b*7	0	2	b*6	0.33	0.33	0.33	0.13	0.33	0.07	0.00	0.50	0.08	0.18	0.19	0.19
EJF	0	3	0	b*5	1	b*2	0.33	0.33	0.33	0.00	0.33	0.00	0.10	1.00	0.25	0.11	0.45	0.28
EKF	0	0	0	3	1	0	0.00	0.67	0.33	0.00	0.00	0.00	0.33	1.00	0.00	0.00	0.44	0.22
ELTE GYFK	b*5	0	b*7	0	b*2	0	0.00	0.00	0.00	0.10	0.00	0.07	0.00	0.25	0.00	0.06	0.08	0.07
ELTE TFK	2	0	8	0	0	0	0.67	0.00	0.33	0.50	0.00	0.13	0.00	0.00	0.00	0.21	0.00	0.10
ELTE TOFK	2	1	1	0	1	0	1.00	0.33	0.67	0.50	1.00	1.00	0.00	1.00	0.00	0.83	0.33	0.58
ESZHF	b*1	0	b*1	b*4	b*1	b*8	0.00	0.00	0.00	0.50	0.00	0.50	0.13	0.50	0.06	0.33	0.23	0.28
GAMF	1	1	1	1	1	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.33	1.00	0.78	0.89
GDF	0	2	0	0	1	0	0.33	0.33	0.33	0.00	0.50	0.00	0.00	1.00	0.00	0.17	0.33	0.25
GYHF	3	1	3	3	1	2	1.00	1.00	1.00	0.33	1.00	0.33	0.33	1.00	0.50	0.56	0.61	0.58
HFF	9	4	0	7	1	3	0.67	1.00	0.83	0.11	0.25	0.00	0.14	1.00	0.33	0.12	0.49	0.31
KE CSPFK	b*2	b*1	b*2	b*1	b*1	b*1	0.00	0.00	0.00	0.25	0.50	0.25	0.50	0.50	0.50	0.33	0.50	0.42
KF	0	0	0	0	b*1	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.17	0.08
KF KFK	b*1	0	b*2	0	b*1	0	0.00	0.00	0.00	0.50	0.00	0.25	0.00	0.50	0.00	0.25	0.17	0.21
KF TFK	b*1	b*1	0	0	b*1	0	0.00	0.00	0.00	0.50	0.50	0.00	0.00	0.50	0.00	0.33	0.17	0.25
KF MFK	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
KFRTKF	1	1	1	b*5	b*2	7	1.00	0.33	0.67	1.00	1.00	1.00	0.10	0.25	0.14	1.00	0.16	0.58
KJF	0	5	0	2	1	b*2	0.33	0.67	0.50	0.00	0.20	0.00	0.50	1.00	0.25	0.07	0.58	0.33
KRE TFK	b*1	b*7	b*1	3	b*1	1	0.00	0.67	0.33	0.50	0.07	0.50	0.33	0.50	1.00	0.36	0.61	0.48

	Search Engine Rankings						Pseudo Precision			Pseudo Rank						Mean Pseudo Rank		
Acronym	**Google**	**Altavista**	**Metacrawler**	**Ariadnet**	**Altavizsla**	**Heuréka**	**general**	**Hungarian**	**total**	**Google**	**Altavista**	**Metacrawler**	**Ariadnet**	**Altavizsla**	**Heuréka**	**general**	**Hungarian**	**total**
KTIF	0	0	0	0	b*3	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.06	0.03
ME CTFK	b*3	b*1	b*7	b*3	b*1	0	0.00	0.00	0.00	0.17	0.50	0.07	0.17	0.50	0.00	0.25	0.22	0.23
MTF	0	0	0	2	b*3	2	0.00	0.67	0.33	0.00	0.00	0.00	0.50	0.17	0.50	0.00	0.39	0.19
MÜTF	b*1	2	0	1	b*1	1	0.33	0.67	0.50	0.50	0.50	0.00	1.00	0.50	1.00	0.33	0.83	0.58
NYF	4	2	8	0	1	0	1.00	0.33	0.67	0.25	0.50	0.13	0.00	1.00	0.00	0.29	0.33	0.31
NÜF	b*5	0	0	b*2	b*1	0	0.00	0.00	0.00	0.10	0.00	0.00	0.25	0.50	0.00	0.03	0.25	0.14
NYF MMFK	b*1	0	b*3	0	0	0	0.00	0.00	0.00	0.50	0.00	0.17	0.00	0.00	0.00	0.22	0.00	0.11
NYF TTFK	0	0	0	0	2	3	0.00	0.67	0.33	0.00	0.00	0.00	0.00	0.50	0.33	0.00	0.28	0.14
NYME ATFK	1	0	5	6	1	0	0.67	0.67	0.67	1.00	0.00	0.20	0.17	1.00	0.00	0.40	0.39	0.39
NYME BPFK	1	1	b*9	0	1	b*6	0.67	0.33	0.50	1.00	1.00	0.06	0.00	1.00	0.08	0.69	0.36	0.52
NYME GEO	2	1	2	0	1	0	1.00	0.33	0.67	0.50	1.00	0.50	0.00	1.00	0.00	0.67	0.33	0.50
PHF	0	0	0	0	b*6	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.03	0.01
PMMF	3	4	3	3	4	0	1.00	0.67	0.83	0.33	0.25	0.33	0.33	0.25	0.00	0.31	0.19	0.25
PSZF	1	3	1	1	1	2	1.00	1.00	1.00	1.00	0.33	1.00	1.00	1.00	0.50	0.78	0.83	0.81
PTE EFK	2	1	5	b*7	b*4	0	1.00	0.00	0.50	0.50	1.00	0.20	0.07	0.13	0.00	0.57	0.07	0.32
PTE IFK	b*1	0	b*1	7	b*1	b*9	0.00	0.33	0.17	0.50	0.00	0.50	0.14	0.50	0.06	0.33	0.23	0.28
PTE PMMFK	1	b*1	6	b*4	b*2	b*5	0.67	0.00	0.33	1.00	0.50	0.17	0.13	0.25	0.10	0.56	0.16	0.36
PTF	0	0	0	1	1	5	0.00	1.00	0.50	0.00	0.00	0.00	1.00	1.00	0.20	0.00	0.73	0.37
RTF	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SE EFK	0	0	b*10	0	b*8	0	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.06	0.00	0.02	0.02	0.02
SSTF	0	0	0	0	b*1	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.17	0.08
SSZHF	b*1	0	b*1	b*6	b*1	b*5	0.00	0.00	0.00	0.50	0.00	0.50	0.08	0.50	0.10	0.33	0.23	0.28
SZAGKHF	0	0	b*2	b*4	b*1	b*4	0.00	0.00	0.00	0.00	0.00	0.25	0.13	0.50	0.13	0.08	0.25	0.17
SZF	0	0	b*9	0	1	0	0.00	0.33	0.17	0.00	0.00	0.06	0.00	1.00	0.00	0.02	0.33	0.18
SZHF	1	10	4	0	3	0	1.00	0.33	0.67	1.00	0.10	0.25	0.00	0.33	0.00	0.45	0.11	0.28
SZIE GMFK	b*2	b*8	b*3	0	b*4	0	0.00	0.00	0.00	0.25	0.06	0.17	0.00	0.13	0.00	0.16	0.04	0.10
SZIE JFK	b*1	b*1	b*1	0	1	b*8	0.00	0.33	0.17	0.50	0.50	0.50	0.00	1.00	0.06	0.50	0.35	0.43
SZIE YMMFK	b*1	b*5	b*4	6	4	3	0.00	1.00	0.50	0.50	0.10	0.13	0.17	0.25	0.33	0.24	0.25	0.25
SZOFI	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SZTE EFK	1	1	1	0	b*3	0	1.00	0.00	0.50	1.00	1.00	1.00	0.00	0.17	0.00	1.00	0.06	0.53
SZTE JTFK	b*3	0	b*2	0	b*1	b*9	0.00	0.00	0.00	0.17	0.00	0.25	0.00	0.50	0.06	0.14	0.19	0.16
SZTE MFK	1	1	1	2	1	1	1.00	1.00	1.00	1.00	1.00	1.00	0.50	1.00	1.00	1.00	0.83	0.92

	Search Engine Rankings						Pseudo Precision			Pseudo Rank						Mean Pseudo Rank		
Acronym	**Google**	**Altavista**	**Metacrawler**	**Ariadnet**	**Altavizsla**	**Heuréka**	**general**	**Hungarian**	**total**	**Google**	**Altavista**	**Metacrawler**	**Ariadnet**	**Altavizsla**	**Heuréka**	**general**	**Hungarian**	**total**
SZTE SZÉFK	3	b*5	0	5	1	2	0.33	1.00	0.67	0.33	0.10	0.00	0.20	1.00	0.50	0.14	0.57	0.36
TF	0	0	0	0	4	0	0.00	0.33	0.17	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.08	0.04
TKBF	1	1	1	6	1	6	1.00	1.00	1.00	1.00	1.00	1.00	0.17	1.00	0.17	1.00	0.44	0.72
TSF	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TSF GFK	1	0	4	0	2	0	0.67	0.33	0.50	1.00	0.00	0.25	0.00	0.50	0.00	0.42	0.17	0.29
TSF MFK	1	1	1	2	b*2	1	1.00	0.67	0.83	1.00	1.00	1.00	0.50	0.25	1.00	1.00	0.58	0.79
TSF MVK	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TSF PFK	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
VHF	0	0	0	0	1	0	0.00	0.33	0.17	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.33	0.17
VTIF	b*3	b*3	0	0	b*1	0	0.00	0.00	0.00	0.17	0.17	0.00	0.00	0.50	0.00	0.11	0.17	0.14
WJLF	b*8	0	b*5	0	b*1	0	0.00	0.00	0.00	0.06	0.00	0.10	0.00	0.50	0.00	0.05	0.17	0.11
ZSKF	1	1	1	1	1	1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
BTA	0	0	0	0	1	9	0.00	0.67	0.33	0.00	0.00	0.00	0.00	1.00	0.11	0.00	0.37	0.19
MFA	0	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PRTA	0	10	0	b*9	b*2	b*8	0.33	0.00	0.17	0.00	0.10	0.00	0.06	0.25	0.06	0.03	0.12	0.08
SRTA	0	0	0	b*7	3	b*7	0.00	0.33	0.17	0.00	0.00	0.00	0.07	0.33	0.07	0.00	0.16	0.08
SZPA	1	1	1	10	1	5	1.00	1.00	1.00	1.00	1.00	1.00	0.10	1.00	0.20	1.00	0.43	0.72
ZMKA	1	0	1	1	0	9	0.67	0.67	0.67	1.00	0.00	1.00	1.00	0.00	0.11	0.67	0.37	0.52
MTA	0	4	0	0	1	0	0.33	0.33	0.33	0.00	0.25	0.00	0.00	1.00	0.00	0.08	0.33	0.21
MPANNI	b*2	0	b*6	0	b*2	0	0.00	0.00	0.00	0.25	0.00	0.08	0.00	0.25	0.00	0.11	0.08	0.10
ELO	0	0	0	0	2	0	0.00	0.33	0.17	0.00	0.00	0.00	0.00	0.50	0.00	0.00	0.17	0.08

Table B.4 Results of *Section 6.6.2*. The identification capability of the acronyms of Hungarian higher educational institutions in 2006.

Acronym	**Pseudo Precision**	**%Pseudo Precision**					**Pseudo Rank**					**Mean Pseudo Rank**	**% Mean Pseudo Rank**
		Google.co.hu	**http://lap.hu/**	**Kurzor**	**vizsla24**	**Yahoo**	**Google.co.hu**	**http://lap.hu/**	**Kurzor**	**vizsla24**	**Yahoo**		
BKÁE TK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.36
BKÁE	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.72
BKÁE GTK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.17
BKÁE KTK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.34
BKE	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.70

Acronym	Pseudo Precision	%Pseudo Precision					Pseudo Rank					Mean Pseudo Rank	% Mean Pseudo Rank
		Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo	Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo		
BME	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.83
BME ÉSZK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.07
BME ÉÖK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.08
BME GÉK	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.27
BME GTK	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	1.00
BME KSK	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.72
BME TTK	0.80	0.67	0	0.06	0.025	0.004	1.25	1.49	0.00	16.67	40.00	11.88	0.76
BME VEK	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.21
BME VIK	0.80	0.67	0	0.06	0.025	0.004	1.25	1.49	0.00	16.67	40.00	11.88	0.72
DE	0.40	0.67	0	0	0.025	0	2.50	1.49	0.00	0.00	40.00	8.80	0.68
DE ÁOK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DE BTK	0.40	0.67	0.24	0.06	0	0	2.50	1.49	4.17	16.67	0.00	4.97	0.72
DE JÁTI	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.71
DE K	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DE KTK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.29
DE MTK	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.08
DE TTK	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.71
DRHE	0.80	0.67	0	0.06	0.025	0.004	1.25	1.49	0.00	16.67	40.00	11.88	0.75
EFE	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
EHE	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.69
ELTE	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.85
ELTE ÁJK	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.70
ELTE BTK	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.74
ELTE TTK	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	1.00
ELTE PPK	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	0.77
ELTE IK	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.71
JATE	0.80	0.67	0.24	0	0.025	0.004	1.25	1.49	4.17	0.00	40.00	9.38	0.60
KE ÁTK	0.60	0.67	0	0.06	0.025	0	1.67	1.49	0.00	16.67	40.00	11.97	0.21
KGRE	0.40	0	0.24	0	0.025	0	2.50	0.00	4.17	0.00	40.00	9.33	0.61
KLTE	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.70
KRE	0.60	0.67	0	0.06	0.025	0	1.67	1.49	0.00	16.67	40.00	11.97	0.71
KRE ÁJK	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	0.28

Acronym	Pseudo Precision	%Pseudo Precision					Pseudo Rank					Mean Pseudo Rank	% Mean Pseudo Rank
		Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo	Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo		
KRE BTK	0.60	0.67	0.24	0	0.025	0	1.67	1.49	4.17	0.00	40.00	9.47	0.73
KRE HTK	0.60	0.67	0.24	0	0.025	0	1.67	1.49	4.17	0.00	40.00	9.47	0.22
LFZE	0.40	0.67	0	0	0	0.004	2.50	1.49	0.00	0.00	0.00	0.80	0.74
ME	0.40	0.67	0	0	0.025	0	2.50	1.49	0.00	0.00	40.00	8.80	0.70
ME ÁJK	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.23
ME AKK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.34
ME BTK	0.40	0.67	0	0	0.025	0	2.50	1.49	0.00	0.00	40.00	8.80	0.70
ME GÉK	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.71
ME GTK	0.40	0.67	0.24	0	0	0	2.50	1.49	4.17	0.00	0.00	1.63	0.91
ME MFK	0.40	0.67	0	0.06	0	0	2.50	1.49	0.00	16.67	0.00	4.13	0.17
MIE	0.40	0.67	0.24	0	0	0	2.50	1.49	4.17	0.00	0.00	1.63	0.94
MIRT	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.07
MKE	0.60	0.67	0.24	0	0.025	0.004	1.67	1.49	4.17	0.00	40.00	9.47	0.92
NME	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NYME	0.80	0.67	0.24	0	0.025	0.004	1.25	1.49	4.17	0.00	40.00	9.38	0.94
PPKE	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.72
PPKE BTK	0.80	0.67	0	0.06	0.025	0.004	1.25	1.49	0.00	16.67	40.00	11.88	0.79
PPKE HTK	0.60	0.67	0.24	0	0.025	0	1.67	1.49	4.17	0.00	40.00	9.47	0.94
PPKE ITK	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.70
PPKE JÁK	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	0.88
PTE	0.80	0.67	0	0.06	0.025	0.004	1.25	1.49	0.00	16.67	40.00	11.88	0.88
PTE ÁJK	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.70
PTE ÁOK	0.60	0.67	0.24	0.06	0	0	1.67	1.49	4.17	16.67	0.00	4.80	0.93
PTE BTK	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.76
PTE FEEFI	0.40	0.67	0	0.06	0	0	2.50	1.49	0.00	16.67	0.00	4.13	0.73
PTE MK	0.40	0.67	0	0.06	0	0.004	2.50	1.49	0.00	16.67	0.00	4.13	0.68
PTE TI	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PTE TTK	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.68
SE	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SE ÁOK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.34
SE FOK	0.20	0	0	0.06	0	0	5.00	0.00	0.00	16.67	0.00	4.33	0.35
SE GYTK	0.80	0.67	0	0.06	0.025	0.004	1.25	1.49	0.00	16.67	40.00	11.88	0.22

Acronym	Pseudo Precision	%Pseudo Precision					Pseudo Rank					Mean Pseudo Rank	% Mean Pseudo Rank
		Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo	Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo		
SE TSK	0.20	0	0	0	0.025	0	5.00	0.00	0.00	0.00	40.00	9.00	0.34
SOTE	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.71
SZE	0.60	0.67	0.24	0	0.025	0	1.67	1.49	4.17	0.00	40.00	9.47	0.78
SZFE	0.60	0.67	0.24	0.06	0	0	1.67	1.49	4.17	16.67	0.00	4.80	0.85
SZIE	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	0.97
SZIE ÁOTK	0.40	0	0	0	0.025	0.004	2.50	0.00	0.00	0.00	40.00	8.50	0.36
SZIE ÉTK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.06
SZIE GÉK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.39
SZIE GTK	0.80	0.67	0.24	0.06	0.025	0	1.25	1.49	4.17	16.67	40.00	12.72	0.74
SZIE KTK	0.20	0	0	0.06	0	0	5.00	0.00	0.00	16.67	0.00	4.33	0.01
SZIE KVA	0.60	0.67	0	0	0	0	1.67	1.49	0.00	0.00	0.00	0.63	0.27
SZIE MKK	0.80	0.67	0	0.06	0.025	0.004	1.25	1.49	0.00	16.67	40.00	11.88	0.76
SZIE TK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.05
SZTE	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.77
SZTE ÁJK	0.60	0.67	0	0	0.025	0	1.67	1.49	0.00	0.00	40.00	8.63	0.70
SZTE ÁOK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.09
SZTE BTK	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.75
SZTE GTK	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.73
SZTE GYTK	0.40	0.67	0	0	0.025	0	2.50	1.49	0.00	0.00	40.00	8.80	0.73
SZTE TTK	0.80	0.67	0	0.06	0.025	0.004	1.25	1.49	0.00	16.67	40.00	11.88	0.77
VE	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
VE GMK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.34
VE MK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
VE TK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
VE MIK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.34
VE GTK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
VIKKK	0.80	0.67	0.24	0	0.025	0.004	1.25	1.49	4.17	0.00	40.00	9.38	0.94
ZMNE	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	1.00
ATF	0.20	0	0	0.06	0	0	5.00	0.00	0.00	16.67	0.00	4.33	0.02
ÁVF	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.72
AVKF	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	0.95
BDF	0.40	0.67	0	0.06	0	0	2.50	1.49	0.00	16.67	0.00	4.13	0.82

Acronym	Pseudo Precision	%Pseudo Precision					Pseudo Rank					Mean Pseudo Rank	% Mean Pseudo Rank
		Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo	Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo		
BDTF	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.67
BEPF	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.70
BGF	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.85
BGF KKFK	0.80	0.67	0	0.06	0.025	0.004	1.25	1.49	0.00	16.67	40.00	11.88	0.81
BGF KVIFK	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.76
BGF PSZFK	0.80	0.67	0	0.06	0.025	0.004	1.25	1.49	0.00	16.67	40.00	11.88	0.76
BKÁE ÁFK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.34
BKF	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	1.00
BMF	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.82
BMF BGK	0.20	0	0	0	0.025	0	5.00	0.00	0.00	0.00	40.00	9.00	0.36
BMF KGK	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	0.96
BMF KVK	0.20	0	0	0	0.025	0	5.00	0.00	0.00	0.00	40.00	9.00	0.16
BMF NIK	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.26
BMF RKK	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.72
DE EFK	0.80	0.67	0.24	0.06	0.025	0	1.25	1.49	4.17	16.67	40.00	12.72	0.85
DE HWPFK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.14
DE MFK	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.67
DF	0.40	0.67	0	0	0.025	0	2.50	1.49	0.00	0.00	40.00	8.80	0.67
EGHF	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	1.00
EJF	0.60	0.67	0.24	0	0.025	0	1.67	1.49	4.17	0.00	40.00	9.47	0.94
EKF	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.69
ELTE GYFK	0.40	0.67	0	0	0.025	0	2.50	1.49	0.00	0.00	40.00	8.80	0.23
ELTE TFK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ELTE TOFK	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.70
ESZHF	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.34
GAMF	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.71
GDF	0.40	0.67	0	0	0.025	0	2.50	1.49	0.00	0.00	40.00	8.80	0.70
GYHF	0.60	0.67	0.24	0.06	0	0	1.67	1.49	4.17	16.67	0.00	4.80	0.85
HFF	0.80	0.67	0.24	0.06	0.025	0	1.25	1.49	4.17	16.67	40.00	12.72	0.97
KE CSPFK	0.40	0	0	0	0.025	0.004	2.50	0.00	0.00	0.00	40.00	8.50	0.39
KF	0.20	0	0	0	0.025	0	5.00	0.00	0.00	0.00	40.00	9.00	0.11
KF KFK	0.40	0	0	0.06	0.025	0	2.50	0.00	0.00	16.67	40.00	11.83	0.54

Acronym	Pseudo Precision	%Pseudo Precision					Pseudo Rank					Mean Pseudo Rank	% Mean Pseudo Rank
		Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo	Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo		
KF TFK	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.18
KF MFK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
KFRTKF	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.82
KJF	0.60	0	0	0.06	0.025	0.004	1.67	0.00	0.00	16.67	40.00	11.67	0.12
KRE TFK	0.40	0	0	0.06	0	0.004	2.50	0.00	0.00	16.67	0.00	3.83	0.47
KTIF	0.40	0	0	0	0.025	0.004	2.50	0.00	0.00	0.00	40.00	8.50	0.49
ME CTFK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MTF	0.60	0.67	0.24	0	0.025	0	1.67	1.49	4.17	0.00	40.00	9.47	0.94
MÜTF	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.77
NYF	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.77
NÜF	0.20	0	0	0	0.025	0	5.00	0.00	0.00	0.00	40.00	9.00	0.43
NYF MMFK	0.20	0	0	0.06	0	0	5.00	0.00	0.00	16.67	0.00	4.33	0.41
NYF TTFK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.35
NYME ATFK	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.69
NYME BPFK	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.26
NYME GEO	0.20	0.67	0	0	0.025	0.004	5.00	1.49	0.00	0.00	40.00	9.30	0.70
PHF	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PMMF	0.80	0.67	0.24	0	0.025	0.004	1.25	1.49	4.17	0.00	40.00	9.38	0.78
PSZF	0.80	0.67	0	0.06	0.025	0.004	1.25	1.49	0.00	16.67	40.00	11.88	0.71
PTE EFK	0.40	0.67	0	0	0.025	0	2.50	1.49	0.00	0.00	40.00	8.80	0.70
PTE IFK	0.20	0	0	0	0.025	0	5.00	0.00	0.00	0.00	40.00	9.00	0.09
PTE PMMFK	0.80	0.67	0.24	0	0.025	0.004	1.25	1.49	4.17	0.00	40.00	9.38	0.40
PTF	0.60	0.67	0.24	0.06	0	0	1.67	1.49	4.17	16.67	0.00	4.80	0.78
RTF	0.40	0.67	0	0.06	0	0	2.50	1.49	0.00	16.67	0.00	4.13	0.73
SE EFK	0.60	0	0.24	0	0.025	0.004	1.67	0.00	4.17	0.00	40.00	9.17	0.48
SSTF	0.20	0	0	0	0.025	0	5.00	0.00	0.00	0.00	40.00	9.00	0.43
SSZHF	0.40	0.67	0	0	0.025	0	2.50	1.49	0.00	0.00	40.00	8.80	0.49
SZAGKHF	0.20	0	0	0.06	0	0	5.00	0.00	0.00	16.67	0.00	4.33	0.53
SZF	0.60	0.67	0	0.06	0.025	0	1.67	1.49	0.00	16.67	40.00	11.97	0.15
SZHF	0.60	0.67	0.24	0.06	0	0	1.67	1.49	4.17	16.67	0.00	4.80	0.59

Acronym	Pseudo Precision	%Pseudo Precision					Pseudo Rank					Mean Pseudo Rank	% Mean Pseudo Rank
		Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo	Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo		
SZIE GMFK	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.67
SZIE JFK	0.80	0.67	0	0.06	0.025	0.004	1.25	1.49	0.00	16.67	40.00	11.88	0.88
SZIE YMMFK	0.60	0	0	0.06	0.025	0.004	1.67	0.00	0.00	16.67	40.00	11.67	0.36
SZOFI	0.20	0	0	0	0	0.004	5.00	0.00	0.00	0.00	0.00	1.00	0.00
SZTE EFK	0.60	0.67	0	0.06	0.025	0.004	1.67	1.49	0.00	16.67	40.00	11.97	0.71
SZTE JTFK	0.60	0	0.24	0	0.025	0.004	1.67	0.00	4.17	0.00	40.00	9.17	0.47
SZTE MFK	0.80	0.67	0.24	0	0.025	0.004	1.25	1.49	4.17	0.00	40.00	9.38	0.95
SZTE SZÉFK	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.01
TF	0.40	0.67	0	0	0.025	0	2.50	1.49	0.00	0.00	40.00	8.80	0.70
TKBF	0.80	0.67	0.24	0	0.025	0.004	1.25	1.49	4.17	0.00	40.00	9.38	0.78
TSF	0.60	0.67	0	0.06	0.025	0	1.67	1.49	0.00	16.67	40.00	11.97	0.76
TSF GFK	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	0.82
TSF MFK	0.20	0.67	0	0	0	0	5.00	1.49	0.00	0.00	0.00	1.30	0.67
TSF MVK	0.80	0.67	0.24	0	0.025	0.004	1.25	1.49	4.17	0.00	40.00	9.38	0.94
TSF PFK	0.40	0.67	0	0	0	0.004	2.50	1.49	0.00	0.00	0.00	0.80	0.67
VHF	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
VTIF	0.20	0	0	0	0.025	0	5.00	0.00	0.00	0.00	40.00	9.00	0.50
WJLF	0.40	0.67	0	0	0.025	0	2.50	1.49	0.00	0.00	40.00	8.80	0.82
ZSKF	0.80	0.67	0	0.06	0.025	0.004	1.25	1.49	0.00	16.67	40.00	11.88	0.82
BTA	0.60	0.67	0	0.06	0.025	0	1.67	1.49	0.00	16.67	40.00	11.97	0.71
MFA	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PRTA	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	1.00
SRTA	0.60	0.67	0	0.06	0.025	0	1.67	1.49	0.00	16.67	40.00	11.97	0.75
SZPA	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	0.88
ZMKA	0.00	0	0	0	0	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MTA	1.00	0.67	0.24	0.06	0.025	0.004	1.00	1.49	4.17	16.67	40.00	12.67	0.94
MPANNI	0.60	0.67	0	0	0.025	0.004	1.67	1.49	0.00	0.00	40.00	8.63	0.42
ELO	0.80	0.67	0.24	0.06	0.025	0	1.25	1.49	4.17	16.67	40.00	12.72	0.97

Table B.5 Results of *Section 6.7.1*. The identification capability of the acronyms of Danish higher educational institutions in 2005.

	Search Engine Rankings							Pseudo Precision			Pseudo Rank							Mean Pseudo Rank		
Acronym	**Google**	**Alta-vista**	**Meta-crawler**	**I2R**	**Ofir**	**Soegning**	**Jubii**	**General**	**Danish**	**Total**	**Google**	**Alta-vista**	**Meta-crawler**	**I2R**	**Ofir**	**Soegning**	**Jubii**	**General**	**Danish**	**Total**
AAU	5	6	b*0	1	1	b*1	2	0.75	0.67	0.71	0.20	0.17	0.00	1.00	1.00	0.50	0.50	0.34	0.67	0.48
AAA	b*0	b*0	b*0	b*0	b*2	b*0	b*2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.25	0.00	0.17	0.07
AH	b*0	b*0	b*0	b*0	2	1	b*0	0.00	0.67	0.29	0.00	0.00	0.00	0.00	0.50	1.00	0.00	0.00	0.50	0.21
ASB	6	b*0	b*0	b*0	1	b*1	1	0.25	0.67	0.43	0.17	0.00	0.00	0.00	1.00	0.50	1.00	0.04	0.83	0.38
AU	b*0	b*0	b*0	b*0	1	9	b*0	0.00	0.67	0.29	0.00	0.00	0.00	0.00	1.00	0.11	0.00	0.00	0.37	0.16
CBS	7	b*0	b*0	b*0	1	b*2	2	0.25	0.67	0.43	0.14	0.00	0.00	0.00	1.00	0.25	0.50	0.04	0.58	0.27
DFM	b*0	b*0	b*0	b*0	3	b*0	1	0.00	0.67	0.29	0.00	0.00	0.00	0.00	0.33	0.00	1.00	0.00	0.44	0.19
DFU	2	b*0	4	2	3	8	7	0.75	1.00	0.86	0.50	0.00	0.25	0.50	0.33	0.13	0.14	0.31	0.20	0.26
DGH	b*0	2	b*0	1	1	1	2	0.50	1.00	0.71	0.00	0.50	0.00	1.00	1.00	1.00	0.50	0.38	0.83	0.57
DJH	5	3	5	8	1	b*1	5	1.00	0.67	0.86	0.20	0.33	0.20	0.13	1.00	0.50	0.20	0.21	0.57	0.37
DKDM	1	1	1	5	1	b*1	1	1.00	0.67	0.86	1.00	1.00	1.00	0.20	1.00	0.50	1.00	0.80	0.83	0.81
DNS	b*0	b*0	b*0	b*0	1	b*0	b*0	0.00	0.33	0.14	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.33	0.14
DPU	3	1	1	3	1	3	1	1.00	1.00	1.00	0.33	1.00	1.00	0.33	1.00	0.33	1.00	0.67	0.78	0.71
DSH	7	b*0	b*0	b*0	b*0	b*0	2	0.25	0.33	0.29	0.14	0.00	0.00	0.00	0.00	0.00	0.50	0.04	0.17	0.09
DTU	1	1	2	1	1	b*5	1	1.00	0.67	0.86	1.00	1.00	0.50	1.00	1.00	0.10	1.00	0.88	0.70	0.80
HIH	10	1	9	5	1	3	2	1.00	1.00	1.00	0.10	1.00	0.11	0.20	1.00	0.33	0.50	0.35	0.61	0.46
IHA	b*0	b*0	b*0	b*0	1	b*0	1	0.00	0.67	0.29	0.00	0.00	0.00	0.00	1.00	0.00	1.00	0.00	0.67	0.29
IHK	4	2	2	4	1	1	4	1.00	1.00	1.00	0.25	0.50	0.50	0.25	1.00	1.00	0.25	0.38	0.75	0.54
IHS	b*0	b*0	b*0	b*0	1	1	2	0.00	1.00	0.43	0.00	0.00	0.00	0.00	1.00	1.00	0.50	0.00	0.83	0.36
IOT	4	3	b*0	b*0	1	4	2	0.50	1.00	0.71	0.25	0.33	0.00	0.00	1.00	0.25	0.50	0.15	0.58	0.33
ITU	8	9	b*0	b*0	1	b*0	1	0.50	0.67	0.57	0.13	0.11	0.00	0.00	1.00	0.00	1.00	0.06	0.67	0.32
KSS	b*0	b*0	b*0	b*0	b*0	b*0	3	0.00	0.33	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.00	0.11	0.05
KVL	1	1	1	3	1	6	b*2	1.00	0.67	0.86	1.00	1.00	1.00	0.33	1.00	0.17	0.25	0.83	0.47	0.68
NNS	9	b*0	b*0	b*0	1	4	1	0.25	1.00	0.57	0.11	0.00	0.00	0.00	1.00	0.25	1.00	0.03	0.75	0.34
OSS	b*0	b*0	b*0	b*0	b*0	b*0	3	0.00	0.33	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.00	0.11	0.05
RMC	b*0	b*0	b*0	b*0	5	6	3	0.00	1.00	0.43	0.00	0.00	0.00	0.00	0.20	0.17	0.33	0.00	0.23	0.10
RUC	1	1	b*0	1	1	9	8	0.75	1.00	0.86	1.00	1.00	0.00	1.00	1.00	0.11	0.13	0.75	0.41	0.61

	Search Engine Rankings							Pseudo Precision			Pseudo Rank							Mean Pseudo Rank		
Acronym	**Google**	**Alta-vista**	**Meta-crawler**	**I2R**	**Ofir**	**Soegning**	**Jubii**	**General**	**Danish**	**Total**	**Google**	**Alta-vista**	**Meta-crawler**	**I2R**	**Ofir**	**Soegning**	**Jubii**	**General**	**Danish**	**Total**
SDU	3	2	2	b*0	1	10	6	0.75	1.00	0.86	0.33	0.50	0.50	0.00	1.00	0.10	0.17	0.33	0.42	0.37
VMK	b*0	5	b*0	b*0	1	1	4	0.25	1.00	0.71	0.00	0.20	0.00	0.00	1.00	1.00	0.25	0.05	0.75	0.35

Table B.6 Results of *Section 6.7.2*. The identification capability of the acronyms of Danish higher educational institutions in 2006.

Acronym	**Pseudo Precision**	**% Pseudo Precision**						**Pseudo Rank**						**Mean Pseudo Rank**	**% Mean Pseudo Rank**
		Google.dk	**Google.com**	**Eniro**	**MSN.dk**	**Jubii**	**Yahoo.dk**	**Google.dk**	**Google.com**	**Eniro**	**MSN.dk**	**Jubii**	**Yahoo.dk**		
AAU	1.00	0.67	0.11	0.09	0.07	0.04	0.02	1.00	0.20	1.00	1.00	0.50	1.00	0.78	0.89
AAA	1.50	0.00	0.00	0.09	0.07	0.00	0.02	0.00	0.00	0.33	0.17	0.25	0.10	0.14	0.05
AH	0.67	0.67	0.00	0.09	0.07	0.00	0.02	1.00	0.00	1.00	0.50	0.00	0.17	0.44	0.80
ASB	0.83	0.67	0.11	0.09	0.07	0.04	0.02	1.00	0.17	1.00	0.50	1.00	1.00	0.78	0.87
AU	0.67	0.67	0.00	0.09	0.00	0.00	0.02	1.00	0.00	1.00	1.00	0.00	0.50	0.58	0.84
CBS	1.00	0.67	0.11	0.09	0.07	0.04	0.02	1.00	0.14	1.00	0.50	0.50	0.33	0.58	0.84
DFM	0.83	0.67	0.00	0.09	0.07	0.04	0.02	0.33	0.00	0.25	1.00	1.00	0.25	0.47	0.36
DFU	1.00	0.67	0.11	0.09	0.07	0.04	0.02	0.20	0.50	0.20	1.00	0.14	0.50	0.42	0.29
DGH	0.83	0.67	0.00	0.09	0.07	0.04	0.02	1.00	0.00	1.00	1.00	0.50	1.00	0.75	0.87
DJH	1.00	0.67	0.11	0.09	0.07	0.04	0.02	1.00	0.20	1.00	1.00	0.20	1.00	0.73	0.88
DKDM	1.00	0.67	0.11	0.09	0.07	0.04	0.02	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DNS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DPU	1.00	0.67	0.11	0.09	0.07	0.04	0.02	1.00	0.33	1.00	1.00	1.00	1.00	0.89	0.93
DSH	0.83	0.67	0.11	0.09	0.07	0.04	0.00	0.50	0.14	0.33	0.20	0.50	0.00	0.28	0.41
DTU	1.00	0.67	0.11	0.09	0.07	0.04	0.02	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
HIH	1.00	0.67	0.11	0.09	0.07	0.04	0.02	1.00	0.10	1.00	1.00	0.50	1.00	0.77	0.88
IHA	0.83	0.67	0.00	0.09	0.07	0.04	0.02	1.00	0.00	1.00	1.00	1.00	1.00	0.83	0.89
IHK	1.00	0.67	0.11	0.09	0.07	0.04	0.02	1.00	0.25	1.00	1.00	0.25	0.50	0.67	0.88
IHS	0.67	0.00	0.00	0.09	0.07	0.04	0.02	0.00	0.00	1.00	1.00	0.50	0.25	0.46	0.19
IOT	1.00	0.67	0.11	0.09	0.07	0.04	0.02	1.00	0.25	1.00	0.50	0.50	1.00	0.71	0.86
ITU	1.00	0.67	0.11	0.09	0.07	0.04	0.02	1.00	0.13	1.00	1.00	1.00	0.50	0.77	0.89
KSS	0.50	0.67	0.00	0.09	0.00	0.04	0.00	1.00	0.00	1.00	0.00	0.33	0.00	0.39	0.77

Acronym	Pseudo Precision	% Pseudo Precision						Pseudo Rank						Mean Pseudo Rank	% Mean Pseudo Rank
		Google.dk	Google.com	Eniro	MSN.dk	Jubii	Yahoo.dk	Google.dk	Google.com	Eniro	MSN.dk	Jubii	Yahoo.dk		
KVL	0.83	0.67	0.11	0.09	0.07	0.00	0.02	1.00	1.00	1.00	1.00	0.25	1.00	0.88	0.97
NNS	0.83	0.67	0.11	0.09	0.07	0.04	0.00	1.00	0.11	1.00	1.00	1.00	0.06	0.69	0.88
OSS	0.50	0.67	0.00	0.09	0.07	0.04	0.00	1.00	0.00	1.00	0.05	0.33	0.00	0.40	0.78
RMC	0.83	0.67	0.00	0.09	0.00	0.04	0.02	1.00	0.00	1.00	1.00	0.33	0.33	0.61	0.85
RUC	0.83	0.67	0.11	0.09	0.07	0.04	0.00	1.00	1.00	1.00	1.00	0.13	0.00	0.69	0.95
SDU	1.00	0.67	0.11	0.09	0.07	0.04	0.02	1.00	0.33	1.00	1.00	0.17	1.00	0.75	0.89
VMK	0.83	0.67	0.00	0.09	0.07	0.04	0.02	1.00	0.00	1.00	1.00	0.25	1.00	0.71	0.86

Table B.7 Results of *Section 6.8*. The identification capability of the acronyms of Hungarian parties in 2006.

Acronym	Pseudo Precision	% Pseudo Precision						Pseudo Rank						Mean Pseudo Rank	% Mean Pseudo Rank
		Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo	MSN	Google.co.hu	http://lap.hu/	Kurzor	vizsla24	Yahoo	MSN		
IDE	0.50	0.000	0.240	0.060	0.025	0.000	0.000	0.00	0.33	1.00	1.00	0.00	0.00	0.39	0.17
MNYP	0.00	0.000	0.000	0.000	0.000	0.000	0.000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MIÉP	0.67	0.670	0.240	0.000	0.000	0.004	0.000	1.00	1.00	0.17	0.17	1.00	1.00	0.72	0.93
MNYNP	0.67	0.670	0.000	0.060	0.025	0.004	0.000	1.00	0.00	1.00	1.00	1.00	0.00	0.67	0.76
MVPP	0.67	0.670	0.240	0.000	0.025	0.004	0.000	1.00	1.00	0.05	1.00	0.33	0.00	0.56	0.94
NDP	0.67	0.670	0.240	0.060	0.025	0.000	0.000	0.11	1.00	1.00	1.00	0.00	0.00	0.52	0.40
PPSZ	0.67	0.670	0.000	0.060	0.025	0.000	0.000	1.00	0.00	1.00	0.10	0.00	0.33	0.41	0.73
SZDP	0.83	0.670	0.240	0.000	0.025	0.004	0.000	1.00	1.00	0.13	1.00	1.00	1.00	0.85	0.95
ZD	0.33	0.670	0.000	0.060	0.000	0.000	0.000	1.00	0.00	0.50	0.00	0.00	0.00	0.25	0.70
ZP	0.00	0.000	0.000	0.000	0.000	0.000	0.000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FIDESZ	1.00	0.670	0.240	0.060	0.025	0.004	0.000	1.00	1.00	1.00	1.00	1.00	0.33	0.89	1.00
KDNP	0.67	0.670	0.000	0.060	0.025	0.004	0.000	1.00	0.00	1.00	1.00	1.00	0.00	0.67	0.76
MDF	0.67	0.670	0.000	0.060	0.025	0.004	0.000	1.00	0.50	0.50	1.00	1.00	0.00	0.67	0.85
MSZP	1.00	0.670	0.240	0.060	0.025	0.004	0.000	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
SZDSZ	0.83	0.670	0.240	0.000	0.025	0.004	0.000	1.00	1.00	0.50	1.00	1.00	1.00	0.92	0.97

Table B.8 Results of *Section 6.8*. The identification capability of the acronyms of Danish parties in 2006.

Acronym	Pseudo Precision	% Pseudo Precision						Pseudo Rank						Mean Pseudo Rank	% Mean Pseudo Rank
		Google.dk	Google.com	Eniro	MSN.dk	Jubii	Yahoo.dk	Google.dk	Google.com	Eniro	MSN.dk	Jubii	Yahoo.dk		
CD	0.17	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.04	0.02
DF	0.83	0.67	0.00	0.09	0.07	0.04	0.02	1.00	0.00	1.00	0.50	0.33	1.00	0.64	0.83
FrP	0.67	0.67	0.00	0.00	0.07	0.04	0.02	1.00	0.00	0.00	0.11	0.33	0.33	0.30	0.70
KF	0.83	0.67	0.00	0.09	0.07	0.04	0.02	0.33	0.00	0.33	0.17	0.50	0.50	0.31	0.30
KPD	0.33	0.00	0.00	0.00	0.00	0.04	0.02	0.00	0.00	0.00	0.00	1.00	1.00	0.33	0.06
RV	0.50	0.67	0.00	0.09	0.00	0.00	0.02	0.33	0.00	1.00	0.00	0.00	0.13	0.24	0.32
SD	0.50	0.00	0.00	0.00	0.07	0.04	0.02	0.00	0.00	0.00	0.17	0.50	0.50	0.19	0.04
SF	0.83	0.67	0.00	0.09	0.07	0.04	0.02	1.00	0.00	1.00	1.00	1.00	1.00	0.83	0.89