Ein „Fast-Forward“-Ansatz zur Erstellung eines Systems zum cross-lingualen Question Answering

Thomas *Mandl*, René *Schneider* und Robert *Strötgen*, Hildesheim

1 Einleitung

Die Lehre im Studiengang Internationales Informationsmanagement an der Universität Hildesheim[1] betont die konstruktiven Fähigkeiten bei der Erstellung von Informationssystemen. Ein wichtiges didaktisches Konzept hierzu bilden Projektkurse im Hauptstudium, in denen größere Aufgaben als Projekt bearbeitet werden. Dabei stehen die Lehrenden vor der Herausforderung, zum einen den Studierenden eine ganzheitliche und sinnvolle Aufgabe zu stellen und zum anderen zu garantieren, dass diese im Rahmen eines Semesters bearbeitet werden kann. Die Ganzheitlichkeit fördert die Motivation der Studierenden und erhöht den Realitätsgrad der Projekte, erschwert aber gleichzeitig die Realisierbarkeit. Das hier vorgestellte Question Answering System eignet sich aufgrund seiner modularen Architektur sehr gut für einen Projektkurs. Wie in einem realen Projekt, erforderte es die Einarbeitung in mehrere bereits vorliegende Komponenten, deren Modifikation für den vorliegenden Anwendungsfall und die Integration in eine Gesamtarchitektur. Die Komponenten bildeten für den Projektverlauf die Arbeitspakete und Meilensteine. Sie werden in den folgenden Abschnitten beschrieben.

Die theoretische Grundlagen zum Question Answering wurden im Semester vor dem Projekt in einem Hauptseminar gelegt. Dabei zeigte sich, dass die Wurzeln der Frage-Antwort-Systeme zurückreichen (Lehnert 1978, Nebel & Marburger 1982). In den vergangenen Jahren hat die Thematik stark an Dynamik gewonnen und besonders das sog. Open-Domain Question Answering hat die Entwicklungen im Bereich der maschinellen Sprachverarbeitung nachhaltig beeinflusst.

Ein Open-Domain Question Answering System akzeptiert eine natürlichsprachliche Anfrage jedweder Art und liefert dazu eine passende Antwort. Diese Definition zeigt gleichzeitig die Nähe bzw. Nachfolgeschaft des Question Answering zu den bereits gut erforschten Bereichen des Information Retrieval und der Informationsextraktion. Der grundsätzliche Unterschied, dass der Ausgangspunkt des Question Answering im Gegensatz zur schlüsselwortorientierten Verarbeitung von Retrievalsystemen bzw. des schablonenartigen Mapping der Informationsextraktion ein vollständiger Satz oder besser eine ausformulierte Frage ist, zeigt, dass aufgrund des höheren Anspruchs an die Komponenten der maschinellen Sprachverarbeitung die Systemkomplexität zunimmt, insbesondere dann, wenn Frage und Antwort in unterschiedlichen Sprachen formuliert sind, wie dies im multilingualen Question Answering der Fall ist.

Die andauernde Aktualität dieser Thematik hat dazu geführt, dass Question Answering mittlerweile ein fester Bestandteil der Evaluierungsinitiativen von TREC, CLEF und NTCIIR[2] ist. Die dort verwendeten Methoden sowie die erreichten Ergebnisse (mit einer durchschnitt-

1. http://www.uni-hildesheim.de/de/5806.htm
2. http://trec.nist.gov; http://www.clef-campaign.org, http://research.nii.ac.jp/ntcir

lichen Precision von 20 Prozent) können als repräsentativ angesehen werden und zeigen gleichzeitig den noch ausstehenden Reifegrad der Technologie.

Der Beitrag, den das derzeit von der Informationswissenschaft der Universität Hildesheim entwickelte System zum Question Answering darstellt, zeichnet sich dadurch aus, dass er bereits ausgereifte Techniken und Systeme der open-source community verwendet, sowie auf an der Universität Hildesheim im Rahmen der Forschung zum Information Retrieval entwikkelte Komponenten zurückgreift (cf. Hackl et al. 2004) und eigene Anstrengungen im Bereich der Frage-Taxonomien und der Antwortextraktion unternimmt. Da während der langjährigen Arbeiten im Bereich des Information Retrieval ein besonderer Fokus auf die Erkennung von Eigennamen gerichtet wurde, die sog. Named-Entity Recognition, und diese gezeigt haben, wie stark der Einfluss eines effizienten Retrievals auf die Retrievalergebnisse ist (cf. Mandl & Womser-Hacker 2005), wurde ein Schwerpunkt in der ersten Phase der Systemerstellung auf diesen Bereich gelegt. Das Paper beschreibt im Folgenden die Systemarchitektur, einzelne Systemmodule und berichtet über erste Ergebnisse.

2 Systemüberblick

Das bislang erstellte System verarbeitet in einem einzigen Durchlauf ohne rekursive Optimierung einzelne Fragen und liefert die vom System gefundenen Antworten in den Textkorpora der Sprachen, die zu Beginn durch den Benutzer spezifiziert wurden. Zum gegenwärtigen Stand erlaubt das System Question Answering für die Sprachpaare deutsch-deutsch, deutsch-englisch, englisch-deutsch und englisch-englisch.

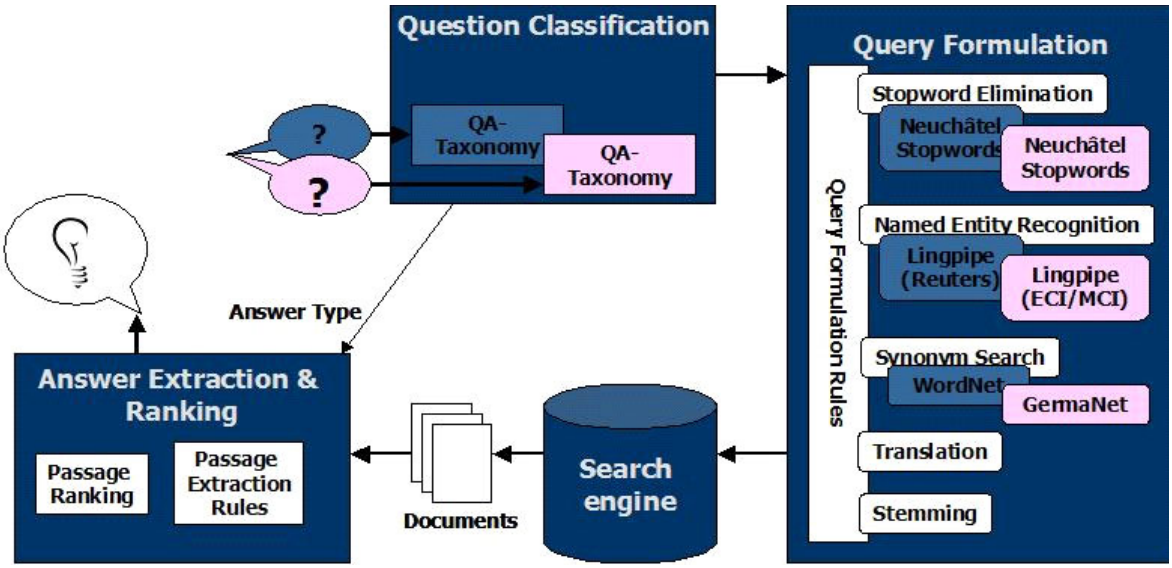

Abbildung 1: Systemüberblick

2.1 Systemarchitektur

Die Systemarchitektur folgt im Großen und Ganzen der Architektur, die sich in den vergangenen Jahren innerhalb der Forschungsgemeinschaft durchgesetzt hat und besteht aus den nachstehenden Hauptkomponenten:

- einer Frageanalyse entsprechend einer eigens entwickelten Fragen-Taxonomie,

- einer Frageverarbeitung bestehend aus
  Stoppworteliminierung,
  Eigennamenerkennung,
  Schlüsselwortextraktion sowie Schlüsselwortexpansion
- einer Übersetzung der Frage, je nach Maßgabe durch den Benutzer
- einer Retrieval- und Extraktionskomponente zum
  Dokumentretrieval relevanter Dokumente entsprechend der generierten Schlüsselwörter
  Passagenretrieval entsprechend einer Formel zur Berechnung von Schlüsselwortauftreten und –dichte.
  Ranking der extrahierten Passagen entsprechend a priori erstellter Regeln, die mit dem jeweiligen Frage- und daraus folgenden Antworttyp korrespondieren.

2.2 Frage- und Antwort-Taxonomie

Aufgrund der Einbindung der Forschungsaktivitäten am Arbeitsbereich Informationswissenschaft der Universität Hildesheim in die CLEF-Evaluierungsinitiative wurde die Erstellung der Frage- und Antwort-Taxonomie an den Fragen der CLEF-Question-Answering Tracks der Jahre 2003 und 2004 (Magnini et al. 2004, Magnini et al. 2005) und einem Erfahrungsbericht der Gruppe um Moldovan und Harabagiu (1999) ausgerichtet.

Fragewort	Unterklassen	Antwortkategorien
WER?	Wer ist [Eigenname]?	Definition
	Wer ist [Definition]?	Person/Organisation
	Wem ?	Person/Organisation
	[Präp.] wem?	
	Wen?	Person/Organisation
	[Präp] wen?	
WAS?	Was ist [Sache/Konzept]?	Eigenname/Organisation/Erklärungen/Definition
	Was bedeutet/symbolisiert?	Erklärung/Definition
	Was [Handlungsverben]?	Artefakte/Handlungen
	Welche [Lebenwesen/Konzept]?	
	Welcher [Lebewesen/Konzept]?	
	Welches [Lebewesen/Sache/Konzept]?	Eigennamen/Organisation/Bezeichnung
	Welchen [Lebewesen/Sache/Konzept]?	
	Welchem?	
	[Präp.] welche/r/s/n/m?	
WANN?		Datum/Zeitangabe
WO?	Wo ist/befindet sich/liegt?	
	Wo arbeitet?	Ort
	Woher?	
WIE?		Art und Weise
	Wie viele?	Zahl
	Wie viel?	Zahl [&Maßeinheit]/Geld/Preis
	Wie lang?	Zeit/Entfernung
	Wie weit?	Entfernung
	Wie groß?	Zahl
	Wie lautet/heißt/ist der Name?	Name/Organisation/Abkürzung
	Wie alt?	Zahl
NENNE		Person/Organisation/Ort/Titel
WOMIT?		Werkzeug/Material
WOZU?		Grund
WODURCH?		Grund/Anlass
WORAUS?		Material/Teile eines Ganzen
WARUM?		Grund

Abbildung 2: Frage-Taxonomie

Entsprechend den Vorgaben für die Erstellung des Systems besteht die Hauptaufgabe dieses Arbeitspakets in der Erstellung einer Frage-Taxonomie für faktische Fragen mit einfacher Semantik, vor allen Dingen um die charakteristischen W-Fragen wie „Wer, Was, Wann, Wo, Wie?, usw.“, sowie der Zuweisung eines oder mehrerer korrespondierender Antworttypen, die die Minimierung bzw. das Ranking der Antworten aus dem Dokumentkorpus minimiert.

Antwortkategorie		**Fragewort (+ Erweiterung)**
Alter	Quantitativ	Wie alt?
Anzahl	Quantitativ	Wie viele?
Art und Weise	Prozedural	Wie?
Datum	Temporal	Wann?
Eigenname Person/Firma		Wer ist + der/die/das Wem? Wen? Welcher? Welche? Welches? Welchen? Welchem? Wie + lautet? Wie + heißt? Wie ist (der Name)?
Entfernung	Quantitativ	Wie lang? Wie weit?
Grund	Kausal	Warum? Wodurch? Wozu?
Liste		Nenne Nennen Welche + [..]?
Maßeinheit/Währung	Quantitativ	Wie viel?
Material, Teile eines Ganzen		Woraus?
Mittel	Instrumental	Womit?
Ort	Lokativ	Wo? Woher? Wohin?
Zeitangabe mit Einheit	Temporal	Wie + lange?
Zweck	Final	Wozu?

Abbildung 3: Antwort-Taxonomie

Die ersten Arbeitsschritte ergaben nachstehende Frage-Taxonomie für das DeutscheDas System überprüft die ihm vorgelegten Antworten in zwei Schritten. Zunächst wird nach dem Vorkommen eines Frageworts gesucht, in einem zweiten Schritt nach Signalwörtern, die auf einen bzw. den speziellen Fragetyp schließen lassen, sofern eine Subspezifikation nötig ist. Wird weder Frage- noch Signalwort gefunden, wird kein Antworttyp generiert.

Zur Illustration dieser Vorgehensweise sei als Beispiel die Frage: „In **welcher** Stadt wurde Emil Zapotek geboren?“ gegeben. Der Fragetyp, der hierbei ermittelt wird, ist “welcher”, der Antworttyp ist jedoch nur anhand des nachstehenden Worts “Stadt“ zu ermitteln und ergibt Location, also einen bestimmten Ort, nach dem gesucht wird.

Im Gegensatz zur Frage-Taxonomie, die möglichst fein gehalten wurde, ging das Bemühen bei Erstellung der Antwort-Taxonomie dahin, diese eher allgemein zu halten, um möglichst viele Fragen abzudecken. Nach Erstellung der beiden Taxonomien wurde diese anhand der Fragen aus den Question Answering Tracks in CLEF aus den Jahren 2003 und 2004 über-

prüft. Das Gesamtergebnis sieht so aus, dass 73 Prozent der CLEF-Fragen von 2004 einer korrekten Antwortkategorie zugewiesen wurden, bei weiteren 14 Prozent der Fragen wurde die Richtung der Antwortkategorie richtig identifiziert und 14 Prozent der Fragen wurde falsch bestimmt. Dies entspricht im Fall des 200 Fragen umfassenden Fragenkorpus aus CLEF 2004 einer Quote von 86 Prozent korrekt identifizierter Fragen. Am zuverlässigsten bestimmt das System Fragen, die *Wann*, *Womit* und *Wozu* enthalten. Dabei gab es niemals eine falsche Bestimmung des Antworttyps. Weiterhin sehr zuverlässig ist die Bestimmung von *Wo*-Fragen und die Aufforderung irgendetwas zu *Nennen*. Fragen, die mit *Wer* oder *Wie* beginnen werden gut, *Was*-Fragen eher ungenügend gut bestimmt.

Daraus wird ersichtlich, dass Fragen, die Unterklassen zu ihrer Bestimmung benötigen, weniger präzise erkannt werden, als Fragen, die nach der ersten Stufe sofort eingeordnet werden können

2.3 Stoppworteliminierung

Für die Stoppworteliminierung wurden die jeweiligen Nêuchatel-Stoppwortlisten[3] für Englisch und Deutsch verwendet und ohne weitere Modifikation integriert. Anhand dieser Positiv-Liste wurden alle Vorkommen dieser Wörter aus den Fragen entfernt, um diese daraufhin zu überprüfen, ob es sich bei einem oder mehreren um Eigennamen handelt und ob sich für die verbleibenden Wortformen Synonyme zur Expansion der Schlüsselwörter finden lassen.

Beide Arbeitsschritte machten einen zentralen Teil der bisher durchgeführten Arbeiten aus und sollen im Folgenden genauer erläutert werden.

2.4 Eigennamenerkennung

Zur Eigennamenerkennung wurde auf das für Forschungszwecke frei verfügbare System LingPipe[4] zurück gegriffen. Lingpipe verwendet einen statistischen Ansatz zum maschinellen Lernen von Eigennamen und Kategorien, die anhand von Korpora trainiert werden, welche vorab mit Labels versehen wurden. In einer früheren Arbeit wurde LingPipe als bestes Werkzeug zur Eigennamenerkennung evaluiert (Mandl et al. 2005). Es entstanden somit zwei Arbeitspakete: a) für das Training anhand geeigneter Korpora, b) für die Anwendung der Trainingsdaten bzw. der Anpassung des Systems an das Gesamtsystem zum Question Answering. Dabei sollten sowohl in den Fragen als auch in den Textkorpora, in denen die Antworten zu finden waren, die Eigennamen erkannt und markiert werden.

Für Deutsch wurde LingPipe anhand des Linguistic Data Consortium (LDC[5]) Korpus der Frankfurter Rundschau trainiert, dessen Umfang ca. 36 Millionen fortlaufende Wortformen entspricht, für das Englische wurde das English Language News von Reuters verwendet, das 810.000 Nachrichtentexte beinhaltet. Eine Evaluation der Erkennungsrate ergab einen Wert von 60 Prozent für die korrekt erkannten Eigennamen sowie einen Wert von 42 Prozent für die korrekte Kategorisierung. LingPipe kategorisiert die erkannten Eigennamen in vier verschiedene Gruppen: Person (PER), Organisation (ORG), Ort (LOC) und Sonstiges (MISC).

3. http://www.unine.ch/info/clef/
4. http://www.alias-i.com/lingpipe/
5. http://www.ldc.upenn.edu/

Anschließend wurde das trainierte System an das Question Answering-System so eingebunden, dass aus den Fragen die jeweiligen Eigennamen extrahiert und aus den Korpora, auf denen die Fragen gesucht wurden, eine Eigennamenindexierung erstellt wurde.

2.5 Synonymsuche

Zur Verbesserung des Retrievalvorgangs wurde wie in vielen anderen Systemen zur Expansion der Schlüsselwörter die Synonymiefunktion von WordNet[6] (Fellbaum 1996) verwendet. Da es sich um ein cross-linguales System handelt wurden für deutsche Anfragen auf GermaNet[7] (Hamp & Feldwig 1997) zurückgegriffen. Bei WordNet handelt es sich um ein monolinguales Lexikon für das Englische, das sich trotz einiger Unzulänglichkeiten zu einem Quasi-Standard für eine Vielzahl von Anwendungen der maschinellen Sprachverarbeitung entwickelt hat. In beiden Fällen wurden die jeweils aktuellen Versionen lokal installiert und in das System eingebunden.

2.6 Übersetzung

Die Übersetzungsanteile wurden über Anfragen an Babylon, einer im Rahmen eines früheren Projektseminars entwickelten Java-Anwendung zur Meta-Online-Überseztung bewerkstelligt (Krauss & Petzold 2004). Dabei werden die Anfragen von Babylon an vier freie Übersetzungsanbieter im Internet gesendet und an das Question Answering-System zurück geliefert. Anschließend werden von Systemseite die relevanten Ergebnisse der Übersetzung als Vereinigungsmenge dargestellt.

2.7 Passagenretrieval

Die überwiegende Mehrheit der Systeme zum Question Answering unterscheidet sich von herkömmlichen Retrievalverfahren durch das sog. Passagenretrieval (im engl. Passage-Retrieval). In Ansätzen, die sich stärker dem Text Mining verbunden fühlen, wird von daher Question Answering häufig auch als Answer Mining bezeichnet, ohne dass sich die dabei verwendeten Verfahren grundlegend voneinander unterscheiden. In sämtlichen Fällen werden Retrieval- und Mining-Verfahren auf Textpassagen der unterschiedlichsten Länge, die von „Textschnipseln“ variierender Länge bis hin zu Sätzen reichen, angewendet.

Im vorliegenden Fall wurde ein gemischter Ansatz aus Text- und Passagenretrieval unter Verwendung der open-source Suchmaschine Lucene[8] sowie einer eigens entwickelten Formel zum Passagenretrieval entwikkelt. Dies bedingt, dass nach Indexierung der Dokumente zunächst ein herkömmliches Dokumentenretrieval durchgeführt wurde, welche die aus der Frage verbliebenen Schlüsselwörter mit den Dokumenten matcht. Auf den ausgewählten Dokumenten wurde mittels nachstehender Formel anschließend ein Passagenretrieval auf Passagen einer Länge von 200 Zeichen plus den darauf folgenden Zeichen bis zum nächsten Punkt durchgeführt. Das Verfahren zum Passagenretrieval ist eng an den von MITRE (Light et al. 2001) entwickelten Algorithmus angelehnt und berechnet für jede Passage eines Dokuments, in dem ein oder mehrere Schlüsselwörter auftreten, einen Wert, der sich aus Auftreten und Anzahl aller Schlüsselwörter berechnet.

6. http://wordnet.princeton.edu/doc
7. http://www.sfs.nphil.uni-tuebingen.de/lsd/
8. http://lucene.apache.org/

$$Score = \sum_{t=0}^{n} (k_1 \times b_t \times \frac{1}{A_t} + k_2 \times \frac{f_{p,t}}{A_p})$$

Relevanzwert 1 Relevanzwert 2

Formel 1: Passagenretrieval

Relevanzwert 1 berechnet sich aus dem booleschen Wert b_t für die einzelnen Terme t. Dieser Wert hat dementsprechend entweder die Ausprägung 0 für kein Auftreten sowie 1 für ein oder mehrere Auftreten von t. Dieser Wert wird mit A_t, d.h. der absoluten Anzahl der Frageterme t sowie einer heuristisch ermittelten Konstante k_1 berechnet.

Relevanzwert 2 trägt der Häufigkeit der einzelnen Terme t in der Passage p Rechnung und ergibt sich aus dem Quotienten $f_{p,t}$, d.h. der Häufigkeit von t in Passage p und der Anzahl der Wörter in Passage p, der wiederum mit einer heuristischen Konstante k_2 multipliziert wird. Die Summe beider Relevanzwerte wird über sämtliche Terme der Suchanfrage aufsummiert.

Dieses Verfahren wurde aufgrund folgender Überlegungen entwickelt: So soll das mehrmalige Auftreten eines Terms nicht so stark gewichtet werden wie das Auftreten mehrerer verschiedener Terme. Dieser Anforderung wird Relevanzwert 1 gerecht. Um ein häufiges Auftreten eines Suchterms jedoch nicht gänzlich unberücksichtigt zu lassen, wird für jedes Auftreten eines Suchterms ein von der Anzahl der Wörter in der Passage abhängiger Wert addiert. Das Verfahren berücksichtigt bislang weder die Synomyme noch die Dichte der Schlüsselwörter. Die ersten Ergebnisse zeigen allerdings, dass der Algorithmus ausreichend gute Ergebnisse liefert.

2.8 Antwort-Ranking

Im Gegensatz zu herkömmlichen Verfahren des Information Retrieval bestehen die Anforderungen des Question Answering darin, eine, d.h. genau eine Antwort auf eine (Suchan-) Frage zu liefern, die nicht aus einem Dokument, sondern aus einer Textsequenz oder idealerweise aus einer korrekt formulierten Antwort besteht. Die Menge der im Passage-Retrieval gelieferten Antworten, müssen demnach auf ein oder kein Element reduziert werden. Letzterer Fall wäre dann gegeben, wenn die zur Verfügung stehende Textmenge die Antwort auf die gegebene Frage nicht enthält.

Die Auswahl bzw. Bewertung der durch das Passagenretrieval gelieferten Antworten wird demnach von zwei Faktoren bestimmten: a) dem maximalen Wert des Passagenretrievals, b) einer Menge von Regeln, welche die Antworten mit der Vorgabe der Antwort-Taxonomie vergleicht (siehe Abschnitt 2.2). So sinkt die Wahrscheinlichkeit, dass es sich um eine korrekte Antwort handelt trotz des höchsten Werts für das Auftreten der Schlüsselwörter, wenn bspw. kein Wort der Antwort mit dem Antworttyp korrespondiert, bspw. wenn nach einer Stadt gesucht wird, die Antwort jedoch keinen Städtenamen enthält. In diesem Fall müsste einer weiteren Antwort, die zwar einen niedrigeren Wert, dafür jedoch einen Ortsnamen enthält, der Vorzug gegeben werden.

Aus diesem Grund wurde nach obigem Muster eine Reihe von Regeln spezifiziert, welche die hoch bewerteten Passagen noch einmal auf ihre Verwertbarkeit überprüfen. Nach Abar-

beitung dieser Regeln wird dann der verbleibende optimale Antwortkandidat vom System als Antwort auf die zu bewertende Frage ausgegeben, andernfalls (d.h. wenn kein Antwortkandidat verblieb) wurde als Antwort die Zeichenkette NIL (für Not In List) ausgegeben.

3 Evaluierung

Im Rahmen der Systemevaluierung wird eine Teilnahme an der seit 2003 stattfindenden Evaluierungsinitiative CLEF angestrebt. Die Ergebnisse der Teilnahme lagen zum Zeitpunkt der Veröffentlichung des Aufsatz noch nicht vor und sind nach erfolgreicher Teilnahme in den jährlichen Konferenzveröffentlichungen sowie der Homepage von CLEF nachzulesen.

Für die Evaluierung wurde eine eher einfach gehaltene Benutzerschnittstelle erstellt.

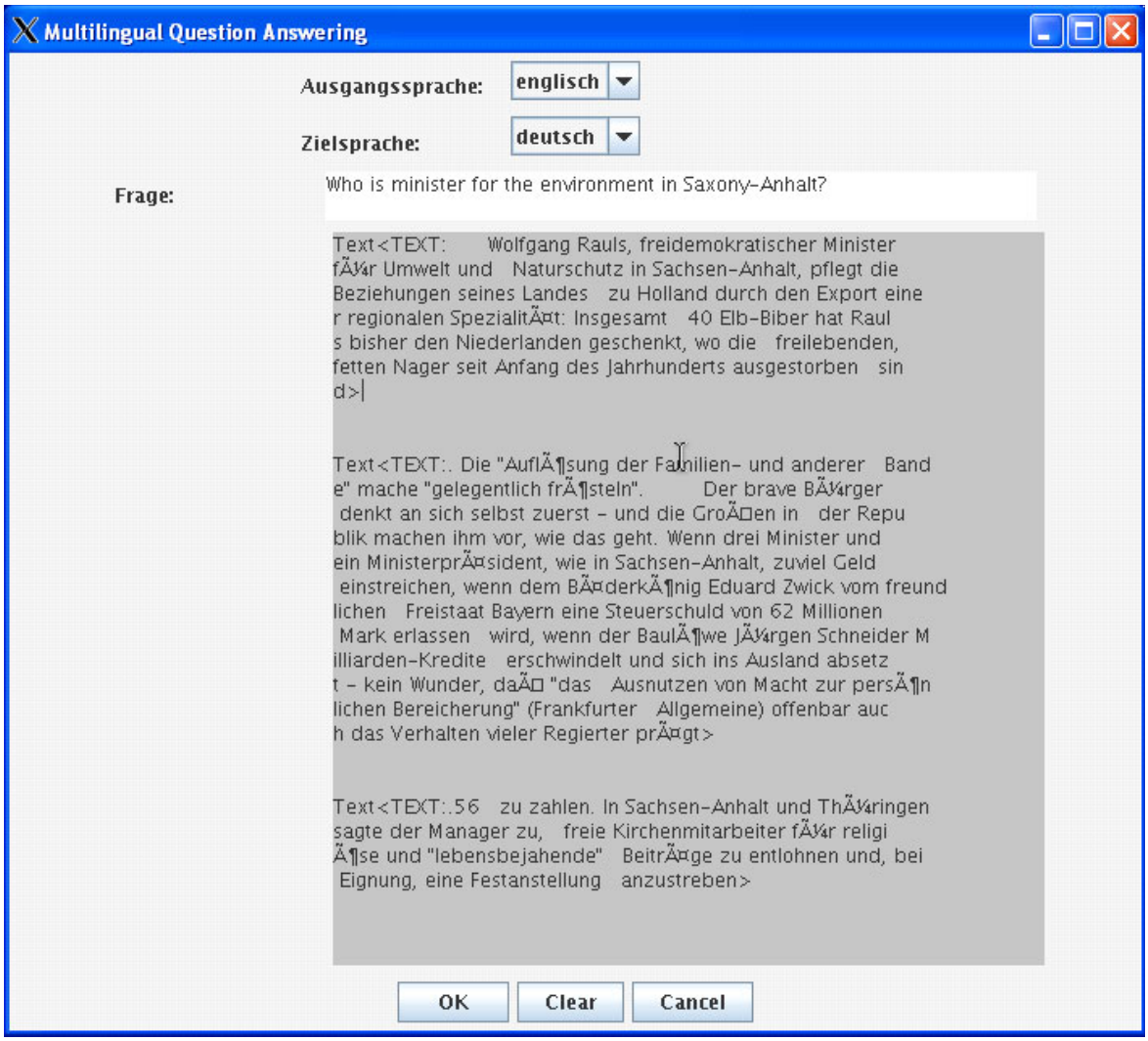

Abbildung 4: Eingabemaske

Für die Evaluierung steht zusätzlich eine Kommandozeilen-Variante der Anwendung zur Verfügung, um im Batch-Betrieb automatisiert eine Vielzahl von Anfragen an das System zu übergeben.

4 Arbeiten und Entwickeln im Team

Das hier beschriebene System wurde im Wintersemester 2004/05 im Rahmen eines Projektseminars im Rahmen der Studiengänge *Internationales Informationsmanagement* (IIM)

und *Informationsmanagement/Informationstechnologie* (IMIT) an der Universität Hildesheim entwickelt (Bach et al. 2005).

Ein wichtiges Lernziel dieses und aller Projektseminare im Rahmen des informationswissenschaftlichen Hauptstudiums ist die gemeinsame und eigenständige Arbeit im Team. Für den Austausch von Nachrichten und Materialien sowie Diskussion anstehender Fragen und Probleme wurde wieder der BSCW (Basic Support for Cooperative Work)-Server der Fraunhofer-Gesellschaft[9] genutzt.

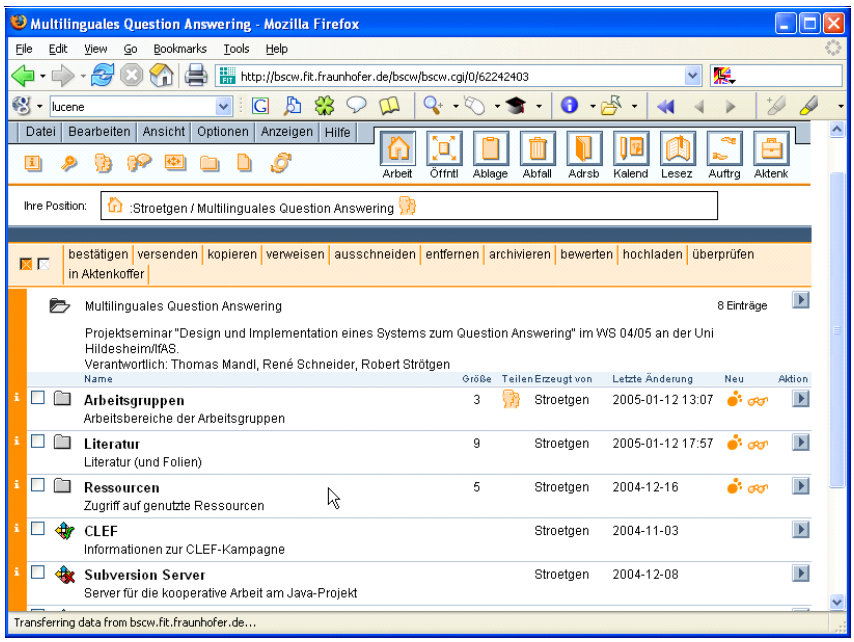

Abbildung 5: Arbeitsbereich auf dem BSCW-Serve

Diese Plattform für kooperatives Arbeiten wurde von den Studierenden angenommen und intensiv genutzt. Allerdings wirkte sich die Speicherplatz-Beschränkung des öffentlichen Servers sehr einengend aus, so dass zukünftig eine eigene BSCW-Installation genutzt werden wird.

Für die Software-Entwicklung in Java wurde die frei verfügbare Eclipse-Plattform[10] eingesetzt. Für das kooperative Arbeiten wurde hier das Versionsverwaltungssystem Subversion[11] eingesetzt, dass sich über das Plugin Subclipse[12] einfach in Eclipse einbinden lässt. Als Vorteil gegenüber Alternativen wie z.B. CVS[13] sprachen vor allem die einfache Administration und unkomplizierte Zugriff für Subversion. Es zeigte sich, dass auch Studierende mit fortgeschrittenen Java-Kenntnissen zunächst Schwierigkeiten mit den Werkzeugen zur Entwicklung im Team hatten, sich aber sehr schnell einarbeiteten und diese Werkzeuge nutzbringend einsetzen konnten.

9. http://bscw.fit.fraunhofer.de/
10. http://www.eclipse.org/
11. http://subversion.tigris.org/
12. http://subclipse.tigris.org/
13. https://www.cvshome.org/

5 Schlussbemerkungen

Der Aufsatz beschreibt die Erstellung eines Systems zum multilingualen Question Answering, das sowohl auf bestehende Systeme zum Information Retrieval und zur Maschinellen Übersetzung als auch auf eigene Überlegungen zur Fragetaxonomie und zur Antwortextraktion gestützt ist. Erste Evaluierungen lassen erkennen, dass die bislang erstellten und integrierten Funktionalitäten eine vielversprechende Basis für Weiterentwicklungen sein können. Dies betrifft weniger die Auswahl neuer Module als vielmehr deren Modifikation und Kombination mit weiteren Ansätzen. Die Weiterentwicklung wird insbesondere von dem Abschneiden bei einer der internationalen Evaluierungsansätze zum Question Answering abhängig gemacht.

Literatur

Abney, S.; *Collins*, M.; *Singhal*, A. (2000)
Answer extraction. In Proceedings of the Applied Natural Language Processing Conference (ANLP), Seattle, WA, pages 296-301.

Bach, Kerstin; Marco *Fischer*; Tanja *Mackensen*; Ulrike *Maichel*; Curt *Nowak*; Carina *Völpel*; Kathrin *Wünnemann* (2005)
Projektdokumentation Question Answering. Arbeitsmaterialie. Informationswissenschaft, Universität Hildesheim. Erscheint.

Bagga, Amit; *Breck* Baldwin (1998)
Algorithms for scoring coreference chains. Proceedings of the Seventh Message Understanding Conference (MUC).

Braschler, Martin; *Ripplinger*, Bärbel (2004)
„How Effective is Stemming and Decompounding for German Text Retrieval?“, In: *I*nformation Retrieval. Volume 7, (3-4) S. 291-316.

Fellbaum, Christiane (1996)
WordNet: Ein semantisches Netz als Bedeutungstheorie. In: Grabowski, Joachim; Herrmann, T.; Harras, G. (Hrsg): Bedeutung, Konzepte, Bedeutungskonzepte, Opladen, Westdeutscher Verlag, 1996, pp. 211 – 230.

Greenwood, Mark, *Gaizauskas*, Robert (2003)
Using a Named Entity Tagger to Generalise Surface Matching Text Patterns for Question Answering. In. Proceedings of the Workshop on Natural Language Processing for Question Answering (EACL03), Budapest, Hungary, April 14, S. 29–34

Greenwood, Mark; *Roberts*, Ian; *Gaizauskas*, Robert (2002)
The University of Sheffield TREC 2002 Q&A System. In Proceedings of the 11th Text REtrieval Conference, 2002.

Hamp, Birgit; *Feldwig*, Helmut (1997)
GermaNet – A Lexical-Semantic Net for German. In: Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications. Association for Computational Linguistics. S. 9-15.

Hackl, René; *Kölle*, Ralph; *Mandl*, Thomas; *Ploedt*, Alexandra; *Scheufen*, Jan-Hendrik; *Womser-Hacker*, Christa (2004)
Multilingual Retrieval Experiments with MIMOR at the University of Hildesheim. In: *Peters*, C.; *Braschler*, M.; *Gonzalo*, J.; *Kluck*, M. (Hrsg.): *C*omparative Evaluation of Multilingual Information Access Systems: 4th Workshop of the Cross-Language Evaluation Forum, CLEF 2003, Revised Selected Papers. Trondheim, Norway, August 21-22, Springer [LNCS 3237] Vorab in: Working Notes http://clef.iei.pi.cnr.it:2002/2003/WN_web/10.pdf

Harabagiu, Sanda; *Moldovan*, Dan (2003)
Question Answering. In: *Mitkov*, Ruslan: The Oxford Handbook of Computational Linguistics. Oxford, University Press.

Krauss, Sebastian; *Petzold*, Jörg (2004)
Babylon: Implementierung von Wrappern für Übersetzungsdienste. Arbeitsmaterialie. Informationswissenschaft, Universität Hildesheim.

Lehnert, Wendy (1978)
The Process of Question Answering: a computer simulation of cognition. Lawrence Erlbaum.
Light, Marc; *Mann*, Gideon S.; Riloff, Ellen; *Breck*, Eric (2001)
Analyses for elucidating current question answering technology. Journal of Natural Language Engineering, Special Issue on Question Answering Fall-Winter 2001.
Magnini, Bernardo; Simone *Romagnoli*, Alessandro *Vallin*, Jesús *Herrera*, Anselmo *Peñas*, Vìctor *Peinado*, Felisa *Verdejo*; Maarten *de Rijke* (2004):
The Multiple Language Question Answering Track at CLEF 2003. In: *Peters*, C.; *Braschler*, M.; *Gonzalo*, J.; *Kluck*, M. (Hrsg.): Comparative Evaluation of Multilingual Information Access Systems: 4th Workshop of the Cross-Language Evaluation Forum, CLEF 2003, Revised Selected Papers. Trondheim, Norway, August 21-22, Springer [LNCS 3237] Vorab in: Working Notes http://clef.isti.cnr.it/2003/WN_web/36.pdf
Magnini, Bernardo; Alessandro *Vallin*, Christelle *Ayache*, Gregor *Erbach*, Anselmo *Peñas*, Maarten *de Rijke*, Paulo *Rocha*, Kiril *Simov* and Richard *Sutcliffe* (2005): Multiple Language Question Answering (QA@CLEF). Overview of the CLEF 2004 Multilingual Question Answering Track.
In: Working Notes 5th Workshop of the Cross-Language Evaluation Forum, CLEF 2004. Bath, England, http://clef.isti.cnr.it/2004/working_notes/WorkingNotes2004/35.pdf
Mandl, Thomas; *Schneider*, René; *Schnetzler*, Pia; *Womser-Hacker*, Christa (2005)
Evaluierung von Systemen für die Eigennamenerkennung im cross-lingualen Information Retrieval. In: Gesellschaft für linguistische Datenverarbeitung. Beiträge der GLDV-Frühjahrstagung. Bonn, 30.3. – 01.04. [Sprache, Sprechen und Computer/Computer Studies in Language and Speech] Frankfurt a. M. et al. Peter-Lang.
Mandl, Thomas; *Womser-Hacker*, Christa (2005)
The Effect of Named Entities on Effectiveness in Cross-Language Information Retrieval Evaluation. In: Proceedings of 2005 ACM SAC Symposium on Applied Computing (SAC). Information Access and Retrieval (IAR) Track. Santa Fe, New Mexico, USA. March 13.-17. 2005. S. 1059-1064.
Moldovan, Dan; Sanda *Harabagiu*; Marius Pa*º*sca; Rada *Mihalcea*; Richard *Goodrum*; Roxana *Gîrju*; Vasile *Rus* (1999)
LASSO: A tool for Surfing the Answer Net. In: Proceedings of the Eighth Text Retrieval Conference (TREC-8) S. 175-184. http://trec.nist.gov/pubs/trec8/papers/smu.pdf
Nebel, Bernhard; *Marburger*, Heinz (1982)
Das natürlichsprachliche System HAM-ANS: Intelligenter Zugriff auf heterogene Wissens- und Datenbanken. In: GI Jahrestagung 1982, Gesellschaft für Informatik. S. 392-402.
Voorhees, Ellen (2002)
Overview of the TREC 2002 Question Answering Track. In: Proceedings of the 11th Text REtrieval Conference (TREC 2002). http://trec.nist.gov/pubs/trec11/papers/QA11.pdf