

Final Report: Projekt "Experimentelle Knowledge-Engine der österreichischen Zivilgesellschaft im Informationszeitalter"

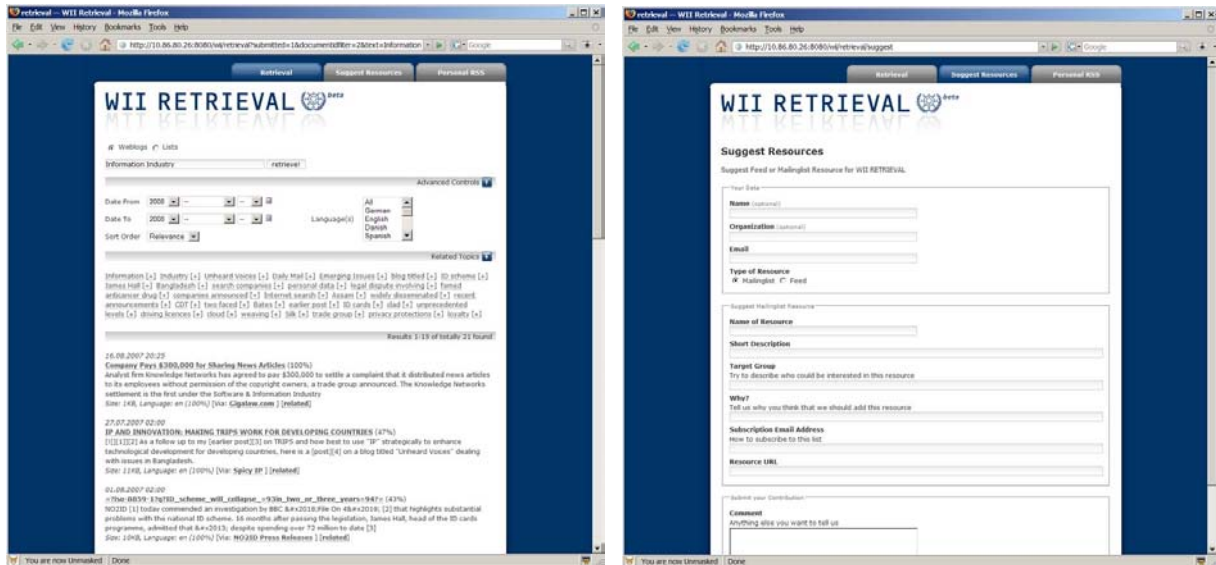
Wien, 21. Jänner 2008

Einleitung

Das Projekt "Experimentelle Knowledge-Engine der österreichischen Zivilgesellschaft im Informationszeitalter" fokussierte auf experimentellen Technologien, mit denen von traditionellen Knowledge Engines schlecht erfassten Datensätze, etwa Blogs und Email Listen, besser erfasst und genutzt werden können.

Die erste Phase des Projektes, die sich von Anfang Dezember 2006 bis Ende Januar 2007 erstreckte, war der Evaluation und dem Aufbau der technischen Basis-Infrastruktur gewidmet. In der zweiten Phase, die sich von Anfang Februar bis Mitte März erstreckte, wurden Informationen mittels standardisiertem Fragebogen und ExpertInneninterviews eingeholt, wie und in welchem Umfang "nicht-traditionelle" Medien von den AkteurInnen der Zivilgesellschaft genutzt werden. In der dritten Phase des Projekts, wurde die Search Engine eingerichtet, in eine Serverumgebung eingebaut, mit Backend-Schnittstellen zu Zope und mit Testdaten gefüttert. Die Search Engine ging in einer Alpha Version online (Url: <http://wii.t0.or.at/wii/retrieval>). In der vierten Phase ging es darum, die Stabilität der Engine zu erhöhen, die Breite der verfügbaren Datensätze zu steigern, in dem ein öffentliches Interface entwickelt wurde, damit User relevante Datensätze (Emaillist Archive, Blogs, etc) selbst zur Indizierung anmelden können und via RSS Feeds sich personalisierte News Seiten zusammenstellen können.

Die vierte Phase hat sich leider als sehr viel komplexer herausgestellt, als das abzusehen war, und konnte nicht planmäßig abgeschlossen werden. Die Engine bleibt deshalb im Rahmen dieses Projektes auf dem Niveau des Alpha Stadiums, das sich noch nicht an End-Nutzer richtet. Wir müssen damit leider hinter unseren ambitionierten Projektzielen zurückbleiben. Gleichzeitig ist es uns aber gelungen, mit dem internationalen Forschungsnetzwerk "Information Retrieval Facility" (IRF, www.ir-facility.org) einen neuen Partner für unser Projekt zu interessieren. Damit besteht die Chance, auf dem bisher Erreichten aufzubauen, das Projekt aber auf eine Größenordnung und technische Basis zu setzen, die Erfolg versprechender ist, als die jetzige Situation. In diesem Sinne müssen wir das Projekt im Rahmen der Netidee mit weniger vorzeigbaren Resultaten beenden, als wir das selbst gerne hätten. Mit der Möglichkeit einer Partnerschaft mit dem IRF sehen wir aber viel Grund für Optimismus, was die Weiterführung betrifft.



Knowledge-Engine Frontend: Eingabe und Ergebnisse mit „Releated Topics“, Ressourcemanager

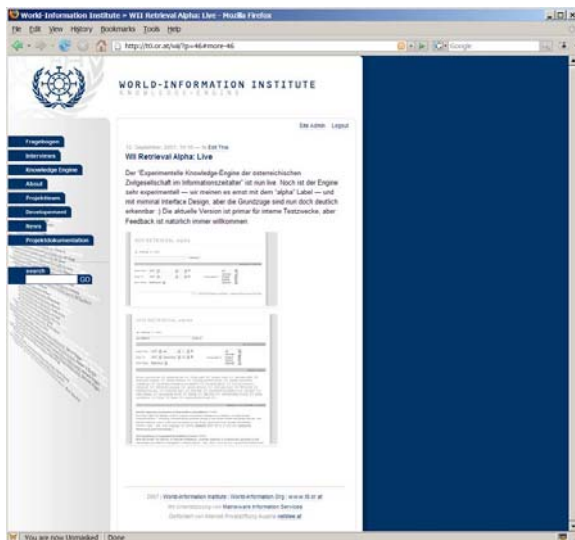
Zentrale Technische Herausforderungen

Während wir am Anfang recht rasche Fortschritte machten, was den Aufbau der Infrastruktur betraf, so haben sich die Grundsatzentscheidungen, die wir getroffen haben, für den Ausbau des Projektes als problematisch herausgestellt. Wir haben uns etwa für eine Lösung entschlossen, die freie Software rund um ein uns zugängliches proprietäres Programm (conceptSearching) gebaut hat, das einige erweiterte Features (semantic indexing, probabilistic relevance ranking) und Suchfunktionalitäten anbietet. Wir mussten diesen Ansatz verfolgen, weil einerseits die freien Software Lösungen, insbesondere im Bereich Indexing, vor einem Jahr noch nicht weit und stabil genug waren (insbesondere Lucene (<http://lucene.apache.org>), das sich im letzten Jahr sehr positiv entwickelt hat, war damals für unsere Zwecke zu Beginn des Projektes noch nicht ausgereift genug) und andererseits die rein-propietären Lösungen nicht nur viel zu teuer gewesen wären, sondern weder die Flexibilität boten, die wir benötigten noch die Offenheit und Transparenz ermöglicht hätte. Mit dem Tool gab es nun das Problem, dass die neue Version, die im Herbst veröffentlicht wurde, nicht nur nicht mehr rückwärtskompatibel war, sondern auch die API veränderte, die die Interaktion über den von uns eigens entwickelten SOAP-Adapter für Zope/Plone mit den freien Tools definierten.

Die freien Tools, auf der anderen Seite, waren selbst oftmals noch im einem frühen Beta Stadium, so dass deren Stabilität nicht ausreichend gegeben war. Die Verknüpfung von mehreren Beta Versionen mit der notwendigen Eigenentwicklung in weiten Bereichen resultierte in zahlreichen Problemen, die oft nur mit extremem Zeitaufwand in den Griff zu bekommen waren. Das Downscaling des Projekts und die damit verbundene Skalierung der Programmumgebung auf einige wenige Features

haben sich rückblickend gesehen überproportional auf die erforderliche technische Grundlagenarbeit ausgewirkt.

Die Verzögerung, die in der 4. Phase unseres Projekts auftrat, lässt sich damit erklären, dass wir einerseits versuchten, die aktuelle Version des Indexing Tools in unser Retrieval Framework zu integrieren, andererseits die vorhandenen Open Source Tools soweit zu verbessern, dass sie stabil laufen würden. Dies ist uns nicht gelungen, und das hatte auch zur Folge, dass für jeden Schritt, etwa der Einpflegung neuer Datenquellen, aufwendige manuelle Arbeiten und Tests notwendig waren. Vor allem Userinput und Interaktionen produzierten wiederholt instabile Situationen und Prozessabstürze sowie neue Fehlerquellen. Dies ist natürlich keinem Enduser zu zumuten und erlaubte uns auch nicht, die Datenquellen so weit zu verbreitern, dass der Search Engine über das Proof of Concept hinausgekommen wäre.



Knowledge-Engine Online: Blog, Dokumentation und Interviews

Wissensgewinn

Als sehr positiv an unserem Ansatz, uns auf bestimmte Datenquellen zu konzentrieren, hat sich erwiesen, dass es uns sogar im Rahmen dieses engen Projektrahmens gelungen ist, wesentliche Verbesserungen des Standard Indexing und Retrieval einzubauen, insbesondere die Möglichkeit, die Datumseingrenzungen sauber zu implementieren.

Auch die Interfaceentwicklung, der Einsatz von Frontends, hat sich im experimentellen Einsatz bewährt. Darüber hinaus ist auch die Möglichkeit, live searches via RSS Feeds anzufordern und in 'third party applications' einzubauen, als sehr viel versprechend zu beurteilen, auch wenn es uns aufgrund der oben genannten Probleme nicht gelang, dies bis einer publizierbaren Stufe zu entwickeln. Aber diese beiden Elemente sowie die klare Ausrichtung auf eine Endnutzer Zielgruppe haben das Interesse der ExpertInnengruppe von "Information Retrieval Facility" geweckt, und wir werden in den

nächsten Monaten die Möglichkeiten einer Kooperation erkunden. Der beigelegte Brief zeigt den Beginn dieses Prozesses. Ein Forum für die Auseinandersetzung mit den Themen des Projekts (und den bis dahin zu Verfügung stehende Tools) soll im Herbst 2008 in Verbindung mit dem internationalen IRF Symposium in Wien stattfinden.

Die Auseinandersetzung mit dem Thema in einer allgemein verständlichen und auch für Nicht-TechnikerInnen leicht zugänglichen Sprache hat ihren Niederschlag in zahlreichen Diskussionsbeiträgen und Weblog-Einträgen gefunden. (URL: <http://t0.or.at/wii/>)

Abschluss

Es ist keine Frage, die hoch gesteckten Ziele, die wir in unserem ursprünglichen Ansuchen formuliert haben, konnten in der verkleinerten Version, die letztendlich bewilligt wurde, trotz effizienten Einsatzes der Mittel (siehe Abrechnung) nicht erreicht werden. Dies ist bedauerlich. Dass nun aber die Möglichkeit besteht, auf internationalem Niveau eine Kooperation aufzubauen, in die wir wesentlich die Erkenntnisse, die wir in diesem Projekt gewonnen haben, einbringen können, gibt dem Projekt doch einen alles in allem positiven Charakter.

Es ergaben sich auch wesentliche Anknüpfungspunkte für unsere weitere Arbeit. Dazu zählt auch die Möglichkeit einer Förderung im Rahmen des ZIT (Zentrum für Innovation und Technologie Wien), von dem eine Unterstützung für eine diskursive Auseinandersetzung und Vermittlung mit diesem gesellschaftlich relevanten Thema in Aussicht gestellt wurde. In diesem Zusammenhang werden wir auch bei der künftigen Publikation von Ergebnissen und Veranstaltungen auf den Beitrag der "Internet Privatstiftung Austria" hinweisen.

Kontakt:

Institut für Neue Kulturtechnologien / t0

World-Information Institute

Lehargasse 11, A-1060 Wien

+43 1 522 1834

<http://world-information.org/wii>

office@t0.or.at