

NATURAL LANGUAGE ACCES TO REGIONAL INFORMATION SOURCES: *THE PORT OF ROTTERDAM CASE*

R. PETERS

*Framfab BV, Amsterdam, The Netherlands
E-mail: Rob.Peters@framfab.nl*

F. WILSON

*Interaction Design, Ltd., Welwyn Garden City, UK
E-mail: fwilson@i-d.co.uk*

Natural Language Interface (NLI) for search queries is regarded as one of the key technologies to support meaningful information retrieval. The combination with statistical Bayesian matching algorithms should create the ideal tool for an international economic region like the Rotterdam Port Area. In this paper we describe the tests performed to support or counter this notion using market leader Autonomy. It was known that the combination did work for knowledge workers at Heineken, Organon and Philips. The tests were carried out in an environment with 11000 Small and Medium Enterprises (SMEs). A regional environment with thousands of websites provides an unusual challenge for a technology normally used for Intranets. The public visitor is not so easy to instruct. Human factors analyses in combination with technical research shows a positive outcome for the Bayesian algorithm but counterproductive results with Natural Language Interfacing. In consequence other Interface support mechanisms need to be developed to combine the power of matching a certain information source somewhere in the region with the low quantity of data in the user query articulation. The solution chosen to tackle this problem was to encourage visitors of the portal to use an email look-a-like query window.

1. The Case

1.1. *The Port of Rotterdam as a regional information source*

The strategy of the 'PortofRotterdam.com' portal was to investigate how a 'regional search' could be used to improve business. While the commercial market was making considerable investment in knowledge management tools such as Verity and Autonomy, it was not clear if these would fit the demands for the navigation of information related to an 'economic region' (distributed data). Green and Novick¹ do articulate the need for collaboration between the research field of Human Computer Interaction and Natural Language, but the application area seems to be focused on speech recognition rather than the use of NLP on the web, and so we had to develop a suitable approach.

1.2. *Autonomy as a natural language search tool*

Geunba Lee² points out in 1997 the gap between considerable funds being poured into NLP research (DARPA, '97) and the small amount of products selling successfully in the market. One reason given is the linguistic preoccupation of the researchers, and so performance increases depend on partial

and incremental parsing, template matching and statistical vector-space pattern matching techniques. In the case of the present market leader, Autonomy, implementation for an Intranet meant the adjustment of the Dynamic Reasoning Engine (DRE) to certain databases and tuning the so-called “Fetch” to ignore irrelevant types of data. The special characteristic of Autonomy is the Bayesian algorithm, which processes the data in the DRE for the index mechanism that creates textual “concepts”. The concepts are clusters of related words where the relation is derived by means of the Bayesian calculus (statistical relations).

1.3. *Input/output issues*

Autonomy requires substantial ‘tuning’ material relevant to the domain of interest, and relevancy is determined by expert judgment of the outcome of search entries. The tuning feed is adjusted to refine outcome. In our ‘Regional’ scenario, the tuning material was derived firstly from generic material (e.g. around shipping) for creating the ‘statistical landscape’ or frame of reference, then from the specific SMEs with their services and products (a third variable is the actual visitor query). Over 500 SME websites were spidered to index content, and the SMEs provided additional material (for the DRE) via an upload tool (free texts and any attachment). For exceptional cases paper was accepted (we digitized), and we attempted to reduce bias where companies had content only vaguely related to their core business. Visitors could then enter a search query to be matched against SME’s “concepts”. NLP is often criticised due to ambiguity of language or anthropomorphism by the user, but this was not the case in The Port environment.

2. Test Methodology

A ‘broad brush’ approach was derived from the STEPS³ approach, and included focus group meetings, direct observation of user behavior, indirect observation via activity logs, user reports (web feed back forms), a ‘comparative output test’ and interviews with users around a targeted enquiry (semi-structured interviews plus questionnaires).

The focus groups involved 22 companies, including both managers and employees. Yong⁴, et al describe the significance of involving users in this way, but their approach requires observation of experienced users and takes an hour for a query response. Our visitors can be anywhere and may be infrequent users, so required a different approach. Our focus groups included stevedores, freight forwarders, shipping agents, warehousing, bunkering, legal services and banking. Participants were asked to fill in a questionnaire of 13 multiple choice

questions regarding the search facilities. A group of 15 SME's was also asked to participate by taking part in actual usage (observed example tasks) and follow-on discussions with testers. Further interviews were performed at the offices of participating SME's.

During the 2-month trial, 1100 Queries were performed in a total of 281 sessions, and we received 50 responses in online questionnaires where they compared the Business search with Google and Yahoo, gave their view on usability, and made comparisons between the 'category navigation' to SME's and exact ratings on a 1-10 scale. We were not the first, to compare our result set with those generated by Google, and for example, Kushniruk⁵ et al applied the same method to determine the power of a health care interface among friends of patients picked out of a waiting room.

The comparative output test (comparison with Google) used 13 pre-defined scenario's and was carried out by a human factors expert and a domain expert grading results as "direct hit", "relevant" link and "not relevant".

3. Key Finding and Proposed Solution

Our system was well received and performed well against measures, but query articulation by visitors proved to be a problem. When input is reduced to keywords the match is reduced to exact or linguistic overlap, and in our tests we saw that the Keyword usage was dominant. Finding a balance between simplicity and sophistication at the input side has been discussed by Huang⁶ et al, while looking at a range of search engines, and mixed interfaces to access heterogeneous sources. They formulate two problems: 1) increasing the cognitive load and making the system hard for novice users; 2) increased difficulty of maintaining mixed interfaces to unstable Internet information sources. In an intranet with knowledge workers, it takes time, education and experimentation to get the best results from tools like Autonomy. Users are 'frequent' and 'known', defined by login/password, and only require training to teach them NOT to use keywords (algorithms require enough input for processing). Earlier research by Jansen⁷ et al, and Bradshaw⁸ et al, confirms that search engine users hardly ever use more than three words. In the case of PortofRotterdam.com, the visitor may visit only once or twice with no link to the management of the site like a password or an email address, so they cannot be trained and must be encouraged to enter more text by some other means.

We identify other solutions that might help the visitor. For example, Baclawski⁹ et al, argue a case for a "breadcrumb path" to support the visitor in his quest, while Google "sets" provides a graphical tool displaying the

interlinking network relationships in a powerful way. Similarly, the aqua browser creates a graphical knowledge representation. We considered that for our users these would be too sophisticated, and so we have developed the concept of 'Email screens' for query articulation, whereby the visitor is encouraged to use more words by writing an 'email' enquiry. This also fits our local program whereby methods are sought to increase the service of the PORT call center. There is an intuitive connection between Call centers and Email entries and we have seen earlier natural mergers of ECRM technologies, Email handling applications and Artificial Intelligence tools at the Dutch National Information office PB 51. Future publications will report on the experience of adopting this method.

References

¹ Green, N., Novick, D., Natural Language in Computer Human Interaction, CHI 99 Special interest Group 1999

² Lee, Geunbae, What makes NLP technologies practical: towards futuristic marketable products. 1997

³ Wilson, F. and Selby, C., Evaluation approaches for Periphera- STEPS. Project Periphera, CEC Telematics Programme, Project Report No. UR1022/WP7/Int7.1/v3a, 1997.

⁴ Yong, L.T., Kong, T.E. Usability study on Soft Computing Tool: Internet Search Tool, 2000

⁵ Kushniruk, A., Kan, M., McKeown, K., Klavans, J., Jordan, D., LaFlamme, M., Patel, V. Usability Evaluation of an Experimental Text Summarization System and Three Search Engines, Implications for the Reengineering of Health Care Interfaces, AMIA 2002.

⁶ Huang, L., Ulrich, T., Hemmje, M., Neuhold, E., Adaptively constructing the Query interface for meta Search Engines. Intelligent User Interface conf., 2001.

⁷ Jansen, B.J., Spink, A., Bateman, J., Saracevich, Tefko. "Searchers, the subjects they search and Sufficiency: A study of a large Sample of Excite Searches. Proceedings of Webnet 1998, 472-477.

⁸ Bradshaw, S., Scheinkman, A., Hammond, K., Guiding people to information: providing an interface to a Digital library Using reference as a basis for indexing" UIU, 2000

⁹ Baclawski, K., Cigna, J., Kokar, M., Major, P., Indurkha, B., Knowledge Representation and Indexing Using the Unified Medical Language System, PSB 2000

This research was supported by the European Commission Information Access and Filtering Program; Key action line III of the fifth framework, 2001-2002