Breaking the rules. Ir and Business

Theo Huibers and Bernd Wondergem

ABSTRACT

The emergence of advanced information retrieval technologies has caused a drastic change in business models, in the nature of information intermediaries and in its automated processing. The characteristics of these changes and their economic consequences are analyzed.

Keywords: information retrieval, business, information processing.

Rompiendo las reglas. Ir y los negocios

RESUMEN

La aparición de tecnologías avanzadas de Recuperación de información ha provocado un cambio drástico en los modelos de negocio, en la naturaleza de los intermediarios de la información y en el procesamiento automatizado de la misma. Se analizan las características de estos cambios y sus consecuencias económicas.

Palabras clave: recuperación de información, empresa, procesamiento de la información

INTRODUCTION

In a recent study, the Dutch Social and Cultural Planing Office speaks of "the multiple choice society"[1]. Citizens now have more choice than ever. And this freedom of choice means they require information to make the right choice. If you need a mortgage, you can choose from countless products and suppliers, each with their own specific terms and conditions. And if you're looking for a home, a holiday or a restaurant, you can take your pick from a huge amount of options.

Faced with all that choice we have developed an almost insatiable appetite for information. Choosing is becoming increasingly complex and so we also need more and more (background) information and guidance. In our services economy these developments have sparked an enormous increase in the number of intermediaries active in such fields as insurance, home finance and travel. They serve up the vast array of product offerings in bite-sized chunks so that you can digest all the facts and make your choice. In other words, they give you the information you need to make the best buying decision. Publishers are also well-placed to make hay in this inter-

---

[1] K. Breedveld and A. van den Broek. De meerkeuzemaatschappij. Facetten van de temporele organisatie van verplichtingen en voorzieningen, Social and Cultural Planing Office, The Hague, 2003.

mediary world. Their business is not to select holidays, mortgages or insurance, but information. On the basis of their selection, they provide information to their customers who are thus able to make better decisions.

Such intermediaries have set up a business model around the advisory process. Tasks are performed by individuals[2] with (powerful) support from information systems. The intermediaries exist by virtue of their ability to collect, select, process and enhance information in a specific way for a (specific) target group; information that everyone can in principle access, either free or for a fee. Everyone, of course, can shop around for the best or cheapest mortgage. But the advantage of intermediaries is that they have made this their field of expertise. They keep up with the latest offers and inovations on a daily basis and make this multitude of information digestible for their customers.

Thanks to new technologies, information can now be automatically collected, selected and, to a certain extent, enhanced. These systems are also known as information retrieval systems. Familiar examples are the popular search engines Google, Lycos and Yahoo! In addition, a broad range of vendors supply a wealth of advanced Information Retrieval technology and applications. Using this technology in inovative ways to create customer value, new business services can be developed with a short time to market.

This new way of collecting information is taking a heavy toll on the traditional intermediaries. Now that travel information, for instance, can be easily obtained via the internet, the added value of a travel agent as a source of information has become marginal. Publishers too are said to be losing their unique position in the information society for the simple reason that relevant information is now readily available free of charge.

In the past few years, the application of Information Retrieval technology in businesses has seen a significant broadening. Nowadays, not only basic search functionality is used in workflow and customer interaction, but new business services and portal components are developed using search technology. This anchors search technology in the commercial heart of many information service providers.

This article homes in on all these changes. First, we will analyse the unavoidable rise of new (automated) information intermediaries. Next we will discuss the impact of the existing information-intensive value chains. The changing environment also compels us to alter our view on the existing income model. Finally we will look at some new players who may soon be playing a central role in the information society.

THE RELENTLESS RISE OF NEW INFORMATION INTERMEDIARIES

Information technology, and particularly the internet, has given us easy access to an overwhelming amount of data and information. The floodtide of information is

---

[2] In many cases qualified people.

unstoppable. It has been said that on an average weekday the New York Times contains more information than a person living in Shakespeare's day would have gathered in his entire lifetime[3]. Via the internet we can access not one but thousands of digital newspapers.

In order to keep the rapid proliferation of digital information under control, we need automated systems that are able to search for that information on the basis of relevance. In the past decades many practical information retrieval systems (IR systems) have been developed. Such systems can be used as an information filter (e.g. the popular spam filters) but also as a search engine (e.g. Google, Ilse or Yahoo!). The available systems are becoming more and more sophisticated and are increasingly capable of satisfying the user's expectations.

The information society and commerce have obviously embraced information retrieval in a big way. Google, for instance, provides access to more than three billion web pages and almost the entire internet population makes use of search systems. Another significant fact is that 70% of all eCommerce-transactions are initiated by an online search. Prospects and customers are less likely to find your organisation via *www.yourcompanyname.com* than by using a search engine.

CHANGING VALUE CHAINS

Information intermediaries are now repositioning en masse or are being replaced by other players. In short, the entire value chain of the information process is in a state of flux.

Information chains change and consequently so do the accompanying value chains. This has huge consequences for the traditional intermediaries who must change to survive. Those who remain stuck in the old pattern will inevitably go under. The really successful players in the new chain are usually not the familiar old players in a new outfit, but relatively young companies like Google and Overture. It's not the selling of information but the search for information that has become big business.

The challenge for both old and new information intermediaries lies in building effective filters and search engines to conect supply and demand. But how can you earn just as much doing that as you could in the old paper-based world where the selling price basically consisted of the cost price plus a fat margin. In the digital environment, "relevance" will be the key price-determining factor.

In other words, the value of information is largely related to the relevance of information. So the bottom-line question is: how do you calculate the costs of relevance? The degree of relevance must somehow be expressed in financial terms. This depends on a combination of multiple factors, including up-to-dateness and reliabi-

---

[3] John Seely Brown and Paul Duguid, De waarde van informatie, een holistische benadering voor organisatie en samenleving, Pearson Education Uitgeverij BV, 2000.

lity but also the receiver's needs and existing knowledge. Setting the right price for a service is one of the most challenging issues for the coming years.

THE FUTURE OF THE NEW PLAYERS

This article showed how existing intermediaries who focus on collecting, selecting, processing and enhancing information are being or will be confronted with severe competition from new entrants. The question remains who are the new players and how robust is their future.

At first sight it would seem that owners of internet search systems can look forward to a lucrative life as the intermediaries of the future. You could say that companies are almost compelled to do business with them as their search systems determine whether the consumer will find the company's website. In this way they can create a virtual monopoly in the intermediary market. But the life of an internet search operator is also full of uncertainties.

For it is still impossible to predict which internet search system will prevail in the near future. User convenience, speed and the size of the accessible collection of documents will probably be the deciding factors. But, as Google has shown, the ability to rank hits in order of relevance is also vital. In the space of one year, Google grew from a relatively unknown internet search system into the absolute world leader thanks to their then brand-new method of detecting documents according to relevance. The inovative part was that Google determines relevance partly on the basis of a site's popularity, which can be measured by the number of referral links and the referral ratings. The big question, however, remains: will a new IR system be able to poach Google's entire customer base? A more advanced IR system might be able to pull off such a coup, and so might a system that is integrated into a control system or a standard company application.

The second group of new players are those who offer their services on the basis of existing search machines. Think, for instance, of internet travel providers who are slowly but surely driving travel agencies out of business. The same is happening in the advertising, car, housing, investing, mortgage, airline and insurance markets, to mention but a few examples. The future of these new intermediaries is also far from certain. The customer's loyalty to an internet-based intermediary is minimal. What's more, online customers are always free to shop around for a better offer before deciding to buy. Time will tell whether the customer is also prepared to pay for the quality and relevance of the information supplied on the services offered. If so, we may see the birth of a new generation of agents who collect information which only paying consumers can access. In such cases the intermediary will collect a fee for his services, which brings us back to the traditional business model of publishers.

Finally, new players may also emerge from the old intermediaries of today. Publishers, for instance, have for some time been seeking to secure a position in the new digital information landscape[4]. They are trying to win the loyalty of customers

by offering better access, good-quality information and a select range of products. But can existing intermediaries make this jump into that new role? That is by no means certain. For it means redesigning their entire business model and everything related to it.

IR IN BUSINESS

Advanced IR technology is used to implement an increasing number of business functions. Before, the example of intermiary portals was already mentioned. Their appeal to customers is further enhanced by application of vertical search technology. Vertical portals, also called vortals, use this technology to crawl and subsequently present users with (only) relevant material concerning a specific theme. It also facilitates more powerful search concepts, such as fielded and facetted search with 'slice and dice' functionality as known in the world of business intelligence. Vortals often also provide users with the possibility to receive alerts, for instance when new material corresponding to their interests is available. A vortal thus aims to provide users with a one stop shop for particular fields of interest and uses several types of search technology to reach this goal.

A second example of advanced search technology is Expert Finding. It provides a solution for the search for people in stead of documents, often related to the search for knowledge. Knowledge is in the heads of people. Access to knowledge, thus requires access to individuals. Yet, the classic approach to knowledge management primarily focuses on finding documents, not on finding people. Of course, attempts are made to let people create profiles, using a set of predefined keywords. However, in practice such profiles provide only limited insight in what people really know.

Expert Finding, as applied in organisations and businesses, uses advanced search technology to crawl and index all documents that possibly represent a firms body of knowledge. Using implicit metadata, that exists in any organization, each document is linked to one or more individuals. The combination of search technology and implicit metadata eliminates the need to add any kind of meta data to documents or individuals. Given a textual query describing some topic, Expert Finding automatically returns the top-n of experts on that topic.

Opinion mining is one of the latest trends in academic research into Information Retrieval. It strives to automatically detect subjective statements in texts. These can for instance be opinions, thoughts and meanings. Subjective statements, such as "the new DVD player is hard to operate", are firstly automatically identified. Secondly, the sentiment of the opinion (i.e. positive or negative) is automatically detected, possibly as a degree.

---

[4] See e.g. the KPMG and Thaesis study "De uitgever aan het woord", Theo Huibers and Cathelijn Timmers, 2007.

Using opinion mining technology, a new series of business solutions can be devised. As seen in practice, opinion mining can be used for the automatic evaluation of marketing campaigns and product launches. One can, for instance, crawl all documents (and blogs etc) about a certain company or their latest product, and subsequently detect all opinionated terms about that product. Over time, it can be evaluated if more (or less) positive opinions are found, if the nature of these opinions changes and how they compare to their competitors products. Another example application detects the sentiment in news papers, e.g. to form some kind of automatically generated Good News Index.

IN CONCLUSION

The position of traditional intermediaries is under fire. The rise of advanced IR technology makes it possible to introduce new business models and to automate new business fuctions. Collection, selection, processing and enhancement of information can be taken over by automated systems. This leads to new business opportunities for new players in the information society. It will drastically change the original business models and value chains of existing information intermediaries –and possibly all intermediaries– in our services economy.

Prof. dr. Theo Huibers MMC is Professor in Information Retrieval at the University of Twente and also managing partner at Thaesis, a strategic advisory firm (*www. thaesis.com*; theo.huibers@thaesis.com).

Dr. Bernd Wondergem MMO is partner at Thaesis, (bernd.wondergem@thaesis.com). He also is a professional trainer in the field of Information Retrieval and teaches courses on this subject at the Erasmus University in Rotterdam, The Netherlands.