Jorge Serrano Cobos  
Planeta de Agostini, Barcelona, Spain

Ana Quintero Orta  
University of Huelva Library, Huelva, Spain

Design, Development and Management of an Information Recovery System for an Internet Website: from Documentary Theory to Practice

Abstract: A real case study is shown, explaining in a timeline the whole process of design, development and evaluation of a search engine used as a navigational help tool for end users and clients on a content website, e-commerce driven. The nature of the website is a community website, which will determine the core design of the information service. This study will involve several steps, such as information recovery system analysis, comparative analysis of other commercial search engines, service design, functionalities and scope; software selection, design of the project, project management, future service administration and conclusions.

1. Introduction

Information retrieval systems tell us generally about new investigations, even new retrieval models, but it is not frequent enough to find in the literature explanations on concrete experiences, or case studies.

This paper intends to explain, based on a real case study, the whole process of design, development and exploitation of an internal information retrieval system, within a content website, oriented specially on vertical communities and e-commerce. The system will give service both users/clients and the enterprise. Thus, we will follow the basic steps needed to transform theoretical approaches onto praxis, and how to apply the classical documentary techniques to online selling.

The website Temalia.com is the example used in this case. It is a web portal with more than a hundred vertical communities hobbies driven, with such diverse thematics as hunting, literature, biology or action cinema. The commercial nature of the website holds the key to understand the nature of the project and its scope.

The main objective of this service is to supply the bare information necessities of multicultural users, with different geographical and linguistical realities, even inside the Spanish language as a "lingua franca", with others true intentions:

- To serve as a bridge or an initial door to many different contents and the communities themselves.
- To answer the questions of the users.
- To show at once the whole world of services which Temalia.com has.
- To help users to become clients.

Let's see the process of making this real step by step:
2. Information recovery system analysis

Every project has to be developed following this first question: “what this will be useful for?”

The information professional tends to try to develop a “perfect” system, or, better said, a perfect system for him. This would be great if we were talking of a typical inner service, designed for information professional exclusive use. In that case, would be easier for the professional to better performance search queries, gaining feedback and expertise on using it and learning the best way to exploit the tool. But this is not the case.

In a web information recovery system, the mediation between text language and users, natural query language will be at least transparent for the user. So we have to focus on user. Usability will be a prority, and free-text based search will be the way we let users search.

Plus, we have to answer other basic questions to determine how we will design the system: Why use a search engine? What for? For whom is designed? Which implications has it for the whole enterprise, the content management or the business processes?

3. Comparative analysis of other commercial search engines

Before deciding the tool we will use, or the front page design, we should take care of what is doing some other websites and what type of capabilities have some typical search engines, the ones who users all around the world prefer. Benchmarking is required.

We will obtain some interesting comparative indicators. We will be able to analyze their main focus, their strategy on serving public, their own original characteristics and how they affect the performance of the whole system.

We will find coincidences and differences between them, learning from them, in such a way that we will be able to gain competitiveness by means not only of quality, but matching specific necessities matching.

4. Service design. Functionalities and scope

We will first think about how we are going to answer the problem of the cultural diversity of our potential users. How we can show the contents, not only the information, to the final user and which kind?

We decide to offer on the results of a search query a little bit of the different types of content and services that we have, bringing to the user the first navigational option. This idea will affect the rest of the project.

We will specify then which kind of service we need for our project, and how the system architecture is going to be organized, depending on the performance necessities.

Finally, and not before, we will start to search for commercial search tools, thinking of which technical requisites we need to look for. This is crucial for the success of the project. Normally enterprises look first at the tool, and then at the project. To the contrary, what we need is first determine why and what for we are going to use a tool, and then look for the tool.
5. Software selection

Several commercial and non-commercial tools are studied. In the business world, it is very important to distribute efforts, so if we don’t have a great technical
force, we should think of using other human professionals to help us in the
development of the service.

Another study is made, comparing the capabilities of every tool, and how they can match with our necessities. We generate more comparative indicators, focused on:
- their technical characteristics
- flexibility to adapt to our main architecture system
- integration with our management content system and content formats
- the company that will develop the project for us, and the number and expertise of their professionals involved. The success and variety of other similar experiences is very important when we consider the tool and the people which we are going to work with.
- Relation price / quality of service
- The priority is not to select the best software itself, the ultimate solution on the market, but the one which best covers our necessities.

6. Design of the project

Once we know which tool will be used, and what company will develop the project with us, we study all details around the development process, such as technical requirements, programming effort, and material and human resources involved.

The point is to find the best way to reduce the time to market, because in e-business this factor is critical. The user needs to find what we have got, and the company needs to offer the products which best match the user necessities, and time is money.

7. Project management

A schedule of actions and people involved is built-up, and shared between them, in order to have everybody aware of the others’ responsibilities.

The steps will include installation of the software, adaptation to the architecture system, adaptation to the content management system, integration with diverse software and creation of front-end templates.

Several professionals collaborate in this interdisciplinary exercise: web designers, usability experts, information professionals, technicians, programmers, content managers, marketing and project directing. Periodical briefing meetings are scheduled, in a state-of-the-art way of working till the pre-fixed objective.

8. Service administration

After being first used in an beta mode and its kick off, a constant period of maintenance is commenced. If worthwhile, a professional should feed the inside thesauri tool with new keywords, building up semantic relations which better help the information recovery continuous feedback.

Example: We have a section of cinema. The film originally entitled in English “The Sound of Music”, in Spain was translated into ‘Sonrisas y Lágrimas’, but in Chile "La novicia novata". This tells us about the immense variety of misspellings, cultural and geographical variations, etc. The service should try to
fight this kind of problem, and one of the better ways is enabling and integrating a thesaurus tool to build all kind of semantic or commercial relationships.

Search logs are one of the best and least tools used to study CBA (Customer Behavior Analytics) since search behaviour offers an unique opportunity to analyze proactive queries about contents, hobbies and products, which will help us to coordinate the products placement between the contents, and the selection of new offers with more probabilities of selling success.

Plus, we will take care of the efficiency and effectiveness of the system. The system will be designed to enhance user satisfaction, helping users to navigate, even if we do not have what user is looking for.

Conclusions

After a certain period of time, the whole project will have to be re-evaluated, because in e-businesses, mobility and adaptivity are common coin. With user feedback and changes of designs or necessities of the product (defining 'product' as the whole website itself) we could think again about the rate of success of the service.

Learning that nothing on Internet is static, we will have the ability to change before, better and rapidly. A system flexible then, with a developing application, will be needed, in order to quickly change what any new project design demands.

References


