Historical query data as business intelligence tool on an internationalization context[1]

Carro-Rodríguez, J.M; Lorenzo-Romero, C.; Gómez-Borja, M.A.
Business Administration, University of Castilla-La Mancha, Spain

Abstract

This article reports theory concerning the key strategies of information search behaviour on an international market-orientation context, proposing a common framework to identify search goals using data generated from the keyword planner from google covering the period 2014/17 applied to the clothing retail sector.

A conceptual framework of user goal identification from search data within clothing retail sector is presented and discussed in light of existing empirical studies. For that, a case study of an important fashion retailer has been analyzed: Zara. This firm is situated in the second position within the ranking of most valuable brands of clothes 2015 around the world (Kantar, 2016). A comparative analysis of search patters of this company, between United Kingdom and Spain, has been developed in order to offer the possible internationalization strategies in the online retail sector.

User goals are identifyed and are stable over the period of study, a framework that covers main clothing consumer search goals have been identifyed.

Keywords: Search behaviour, internationalization strategy, Zara as study case, search patterns by query data framework with KeywordPlanner, comparative analysis with Google Trends.

---

1. Introduction

Historical search data is a great source of business intelligence about consumer search in order to describe the number of searches related to a specific query or groups of queries.Therefore search data offers a volume of queries that might convey different types of goals. In order to understand the underling goals from keyword volume, we have to associate user goals with queries, in other words, the why of search behaviour is actually essential to offer or satisfy the user information need.

Our premise is that search data reflects a diverse set of underlining goals and the knowledge of those goals offers the prospect of future improvements and the degree of interest for product purchasing or information for decision-making. we will analyze the web search patters of Zara, as specific study case of the second most important fashion firm in the world. Moreover, we will compare the obtained query data for two different countries, United Kingdom and Spain, in order to offer different international strategies to fashion retailers respect to positioning and web contents.

2. Literature review

Through literature information, we can find a long history of educational behaviour which allow us to understand the consumer behaviour in an online context, based on how people make their searches and the whole search process. Future use of Big Data with the help of Artificial Intelligence is very bright. The use of artificial intelligence will lead to production of machines and computers, which are much more advanced than what we have today. Researchers are continuously working to handle growth of data as well as to convert it into valuable assets. There has been an increasing interest in the application of Aritificial Intelligence tools to IR in the last few years. Concretely, the machine learning paradigm whose aim is the design of system able to automatically acquire knowledge by them. In the light of requirement of intelligent processing of the big data so as to retrieve the information as per the business requirement, authors have proposed a novel architecture.

Lau & Horvitz (1999) analysed a set of server logs and develop a user behaviour model from log data. They were motivated by the idea that informational retrieval could lead to develop of models to diagnose a user's informational goal, in their model defined a refinement strategy of query sequences', whereas the refinement class of query represents user's intent to his prior query. Concluding that there is a difference between query and informational goals. Broder (2002) was motivated by the idea that the intent behind a web search is not often informational. In his work, he recognises that the informational need is associated with some task, and this "need" generates a human-computer interaction where cognitive aspects play a significant role as the need is verbalized mentally and translated

into a query posed to a search engine, in his research he has come up with the trichotomy of web search "types" navigational, informational and transactional. Rose & Levinson (2004) developed a framework of user goals with a sample of queries from Altavista logs, they reached the conclusion that the goals naturally fell into a Broder´s (2002) trichonomy at the top level and the full framework was a hierarchy structure, the researches believed that in many cases, user goals could be deduced: By looking at the user behaviour, the query itself, the type of results returned in the search list, results click by user further research or other action by the user. Webb (2009) studied the historical weekly search data from the query "Foreclosure" from google trends in the "US" and correlate with statistics of US home foreclosures, finding that search data from the related domain could predict the foreclosures. Search data has been used to predict financial behaviour. Preis, Moat, & Stanley (2013) quantified trading behaviour using data from Google Trends by analysing changes in query volumes for search terms related to finance. They analysed the performance of 98 terms including stock markets with some terms suggested by the Google set service tool that identifies semantically related keywords. Researchers concluded that conveying large behavioural financial trading data sets with data on search query volumes related to the financial might offer insights into different stages of the decision-making. Despite the increasing popularity of predicting consumer behaviour with historical search data with Google Trends, little research has examined the potential of conveyning query formation and user search goals conditions required for successful in the international market-orientation decision-making contex. Not a lot of research has been carried out or examined with precision how search analysis data using a search goal framework can obtain user behaviour to be used for business intelligence insight.

Currently, Sharma & Srivastava (2017) proposed architecture may help in faster information retrieval with better accuracy and recall.

3. Clasifiying user goals for the clothing sector

This study proposes a search pattern analysis to identify the search goals from the clothing sector. We analysed a specific study case characterized for being an e-commerce website from major-clothing retailer: Zara. Racked Fashion (2016) report includes Zara within the ranking of "retailers which offer intuitive site navigation, a smooth checkout, easy shipping and returns, and awesome products (clothes, shoes, bags, and accessories)". In fact, Inditex has hegemony in the textile sector (being Zara the 50% total sales of group). The Zara's customer invests in clothing average of 266 euros per year (Fintonic, 2016). their search processes. We collected 6 sets of queries per each year (2014/2017) , 3 per country that contained the brand name itself and the brand name plus the main category name (Spain: Zara tienda, Zara hombre; UK: Zara shop, Zara man, Zara woman). The purpose of this

examination was to get a wide sample of queries related to the brand name and main product category that will allow us to identify user goals over time, the full download covered 346 groups that contained 4835 suggested queries for Spain and 262 groups and 4,665 queries for the UK. Not related terms to the country in study emerge in the sample, such as "Tienda Zara DF". They were all removed.

3.1. Identifying user goals from search data

We consider that user goals can be inferred from analysing the user interaction with the search engine, these interactions might be: The query itself, results retrieved by the search engine, type of results clicked on by the user and query reformulation. However, this information is not available or to condense to analyse or at least the last two options.

Our first task was to classify the type of search goals that users might have in mind when interacting with a search engine. Those interactions are the result of information needs from the users during. The queries downloaded where organized in columns,"adgroup" had the key tag word for each group of keywords, Then we manually classified the key tag word from the column adgroup into our framework, in order to identify the underlying goal, first, we identify it by the query itself and second by typing the keyword in the browser and check the list results it they matched the query criteria.We proceed in the following way: A tag query or adgroup like "Zara Locations" has several queries in the group: "Zara store locator", "Zara store locations", "Zara men store locator", and "Zara man store locator". Due to the relationship among the tag query and the queries within its group, the suggesting goal for all queries was local. The goal was collaborated after submitting the query in the search engine and analysing that the main result link was a store locator and the google maps feature displayed.

3.2. Taxonomy of queries

In our study, we used the following taxonomy of query goals largely based on Rose & Levinson (2004).

We defined **navigational** goal as a demonstrating desire by the user to be taken to the home page. We differentiate queries consisting on the clothing brand named as "Brand" and "Transactional" queries which are formed by the website name or its reformulation. For it to be considered navigational, the query must have a single authoritarian web site that the user has in mind. Not all queries containing the brand name are considered as navigational, user might type an authoritarian web site name to narrow the results within a specific web page. This is because of the relevance relationship between terms is directed, term "A" may strongly suggest term "B" but not vice-versa according to Joshi & Motwani (2006).

Informational category covers goals for answering questions and learning more about the topic.Queries seeking for reviews, sales instructions, customer service and suggestions, we

have named them **Advice,** this group can be broadly defined as – I need to know what to do. Product gender could lead as to estimate the interest on female or male products, we split queries related to gender to quantify the volume and interest on those products, categories like woman containing female products and its first link in result is the authoritarian Zara female product link "Zara woman" and the same classification method was applied to add "Man" category. The informational goal **Locale** requests information about stores locations, but we also include contact information and opening hours (e.g. "Zara phone numbers", "Zara opening hours store"). **List** can be define as exploratory search queries, the user is not even considering a purchase, queries are less focused, the main goal is to get a suitable list of results suggested for more in-depth information (e.g. "Zara trouser"). New categories emerged, we categorized under **Deals** when users search for information about products or services that have a specific and temporary status where price is a main constrain, queries such as "clothing sale" or "black Friday Zara deals" are included in this category – I am searching on the web to find only bargains. Another contribution to this taxonomy is the category **Trends** - I am searching on the web for X because I want to know what is currently trendy or new-, queries consisting on topics about "new arrivals", "new season products", "spring summer 2016", "shoes trends", "Zara fashion" will be included into this category.

Another discovery was the category **Target Group** that defines searches that have a define target consumer group and they are not releated to gender (e.g. "Zara kids, "Zara pregnancy clothes"). Many keywords have obietives out of product or future purchase, those are categorized under **Enterprise,** we define this objective as searching information about the company itself and not its prodcuts or services (e.g. "Inditex total turover", "Zara number of shops"). **Clothing attribute** goal are described as searches related to product or service with a high level of search domain by the user, (e.g. "Zara military jeans", "lether jackets").

Resource queries achieve something apart from information. If the resource is, something that we install via App, computer or online the goal is **Download**. When the resource is something that I need to use in the real world, such as logos, turnover statistics, corporate brand information and so forth, we call it **Obtain**. If my goal is to enjoy the resource for leisure and for sharing and social profiles, the goal is **Entertainment and social**; the most common examples are social media profiles, blogs, online catalogues and video. Lastly, the **Interact** goal arise when further interaction will happen after landing the desire page, (such as fashion directories and job vacancies opportunities. we don't include map service) to accomplish user search.

4. Results

We observe that for the set of data under study the searches related to navigational goals in the Spanish market are 7% of the total of searches for the 4 years of study, being 5% for transactional, related to online purchases and 2% as brand awareness.

On the other hand, users in the UK generate a total volume of 13% for the same goal category 8% only for transactional searches and 5% for brand awareness, in both cases the search volume is higher than the Spanish market, we discovered that although the volume is higher for the objective in general only differs by 1% in those that refers to transactional objectives which are related to ecommerce mainly.

Regarding the searches generated for the information goal which receive the most attention for consumers in both countries, the values are very similar, being 37% for Spain and 35% for the UK, the new objectives identified in this research, emerge from this category.

Spainish users have more interest in trends issues by marking a volume of the total of 11% whereas the UK only reaches 5%, in the search goal deals the British market reaches 4% compared to the 3% in Spain.

Users in both countries have the same behavior when performing searches related to the goal named"target" with 2% of the total searches, Attribute and Enterprise goal categories have emerged however they are not consolidated since the volume that show is below 1% for both countries so we will have to take it into account for future research, on the contrary we see that the main objective of users is defined by the search for information since the goal "List" reaches the highest percentage in "informational" goal with a 12% for Spain and a 17% for the United Kingdom, this goal is closely linked to the locale with 4% and 3% respectively. As user may perform a searcg before planning to go to the shop.

Consumers in both markets search for "advise" with a volume of 5% for Spain and 4% for the United Kingdom, Regarding the objectives of resources is very different where UK users do not generate searches related tom this goal , according to data, only Spanish users collect 7%, mainly for "obtain" objectives with 3% and "entertainment" with 4%.

As an overall view Spain reaches 51% of the total volume and the united kingdom 49%, so the data are comparable and can be compared for strategic decision making.

5. Conclusion and implications

In this paper, we have studied the potential of historical search data to identify underlying goals of search. We have collect datasets from a given query using the KeywordPlanner tool and develop a framework based on previous studies. To understand user goals in the

fashion sector we established a search goal framework where all queries naturally fell into, analyzing the search patterns for Zara brand in United Kingdom versus Spain.

We observe that queries grouped themselves by common goals from the datasets, we classify query groups into categories for decision making, transactional category was added into the navigational goal, as the proposed of those queries is to reach a site for interact, but we could no conclude whether those transactional queries finally end up into a sale.

From the extracted data we have monitored that user goals identified under this new framework are stable in time as the sample covered 2014 to 2017 therefore we concluide that the current user goals are the most popular for the clothing industry consumer.

As main limitations we have found that datasets were generated by product related queries and brand name, therefore results can be too focused on search goals categories. Moreover, goal classification was carried out by the query itself first and the links generated in the results page, the result list would be organised by the algorithms and might change through time, it can only be taken as a current picture of the current situation.

We suggest for further investigations to extend the framework to analyse user goals not only with historical search data also with internal data from google analytics to categorize inbound traffic queries and search terms into the framework. Another line of interest for investigations should be the data analysis to predict or forecast events in the real world using search data with the search goals framework to compare how the goals correlate with data, and extended to more countries and brands.

References

Broder, A. (2002). A taxonomy of web search. In ACM Sigir forum, Vol. 36, No. 2

Fintonic (2016). El consumo de moda en España [Fashion consumption in Spain]. Available at http://www.elmundo.es/economia/2016/01/12/5694d7feca474159218b45c5.html

Joshi, A. & Motwani, R. (2006). Keyword generation for search engine advertising. Proceedings of Sixth IEEE-ICDM, 123-129. 2. Dongqing Zhu, Ben Carterette, 2010.

Kantar (2016). Las 10 marcas de ropa más valiosas en 2015 [The 10 clothing brands most value in 2015]. Available at http://es.kantar.com/empresas/marcas/2015/mayo-2015-ranking-brandz-de-las-marcas-de-ropa-m%C3%A1s-valiosas-del-mundo-en-2015/

Lau, T., & Horvitz, E. (1999). Patterns of search: analyzing and modeling web query refinement, pp. 119-128. Springer Vienna.

Moat, H. S., Curme, C., Avakian, A., Kenett, D. Y., Stanley, H. E., & Preis, T. (2013). Quantifying Wikipedia usage patterns before stock market moves.Scientific reports, 3.

Preis, T., Moat, H. S., & Stanley, H. E. (2013). Quantifying trading behavior in financial markets using Google Trends. Scientific reports, 3.

Racked Fashion (2016). The 38 Essential Online Shops. Available at http://www.racked.com/2015/7/14/8923189/best- online- shopping- stores

Rose, D. E., & Levinson, D. (2004). Understanding user goals in web search. In Proceedings of the 13th international conference on World Wide Web, ACM, 13-19.

Webb, G. K. (2009). Internet search statistics as a source of business intelligence: Searches on foreclosure as an estimate of actual home foreclosures. Issues in Information Systems, 82.

Sharma, L. & Srivastava, V. (2017). Performance Enhancement of Information Retrieval via Artificial Intelligence, International Journal of Scientific Research in Science, Engineering and Technology, 3(1), 187-192.