

WEB ANALYTICS VISUALIZATION

William Rojas Martínez, Emilio Alvarado Santis, Adriana López Santa, John Alexis Guerra Gómez,
José Tiberio Hernández P.

Universidad de los Andes

Carlos Rodríguez, Abelardo Duarte

VoiceBunny Inc.

ABSTRACT

This document describes the process of designing and implementing a visualization for analysis of users' navigation on a web page, based on the data collected on the navigation on the site www.voicebunny.com, who is our customer, and whose Design was made based on the business questions that the company wanted to solve, but was designed in open source, so the tool can be reused by any company or person interested in performing a web page analysis. The proposed visualization allows the user to interact in order to discover new insights and allows great scalability.

INTRODUCTION

With the rise of electronic commerce, through the supply and purchase of goods and services online, which has annual growth of 18% worldwide, and by 2020 it is expected that consumers will prefer to make their purchases online [1] , And "... these **new consumers** demand from the brands a dialogue, an exchange of ideas that feedback them to achieve fidelity" ... [2], it is for this reason that the analysis of navigation data from its users, in order to identify the interests of its visitors, to be able to segment them and thus to offer services tailored to the needs of its users or attract new customers, track the volume of

purchases made by individual customers or by specific groups of customers, observing the geographic regions where more and fewer customers visit the site, buying specific products, and predicting which services are more and less likely to be purchased in the future, which Provides entities with savings in advertising costs, optimize their marketing campaigns and increase the volume of their customers.

The main limitations of the current Web analytics services correspond to the previous parameterization of the data to be analyzed and the inability to reprocess the freight information to a failure in such parameterization, the use of sampling in the analysis of large volumes of information which can alter the results of the analysis and in other cases the high costs of those tools that are not open source, and on the other hand in most cases its use is not so intuitive for the end user in charge of the data analysis [3]. Although the data that is stored are public the information is frequently stored in the servers of the companies that offer such services. [4]

The visualization tool has been built to help the user easily identify the path followed by visitors once they enter the web page, which will allow analysts to identify the conversion of purchase, ie which pages induce the Users to

purchase a service, in order to measure the performance of the page and thus achieve improvement, identify the days of increased traffic within the period of time evaluated and determine exactly what type of user visited the page in a determined period.

The main advantage of the visualization tool is that the company decides what period of time it wants to analyze, by selecting the corresponding data, without having to make previous parameterizations that can delay and make the analysis tedious and can replicate the analysis How many times it requires it.

RELATED WORK

The main objective of the Web Analytics Visualization project is to make an interactive visualization that allows to increase the analysis capabilities of the client against the behavior of the users in its web page of the company, based on the navigation data. For this purpose, a visualization is proposed that tries to solve three main tasks: to present the distribution of the pages visited by the users in the website of the company in a given period of time, to discover characteristics of interest as the volume of users with navigation behaviors Similar per session or the different routes and sales funnels that were found and finally compare the behavior of sales funnels of site visitors without a previously established goal.

In the first part a sunburst visualization is presented (Figure 1), this technique uses a radial design [5] to present the information of the most

frequent navigation routes. Here it is possible to select the frequency range of the sessions that presented the same sequence of pages visited.

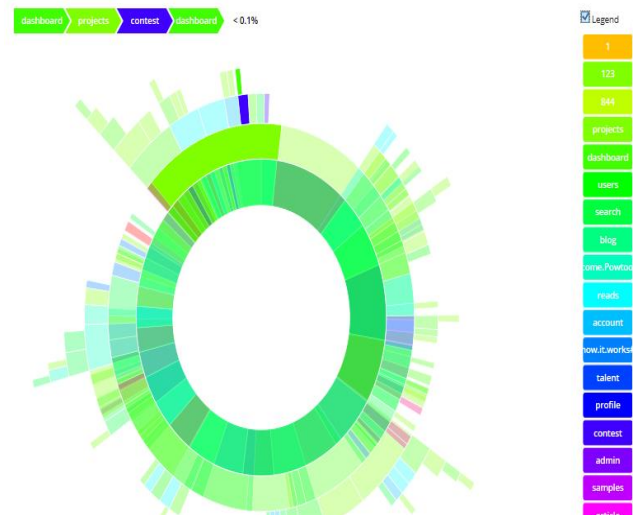


Figure 1: Sunburst that shows each of the visits made by visitors to the website during the period evaluated and shows the percentage of visitors who made the same sequence of pages.

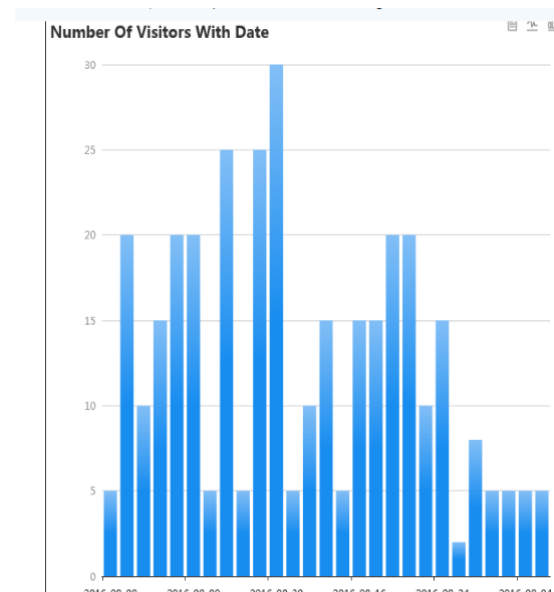


Figure 2: Bar chart shows the distribution of visits per day of the period evaluated, activated when selecting a path in the Sunburst.

Secondly, there is an auxiliary display that follows the bar diagram (Figure 2), in order to present the distribution in time of the visits made to a specific route selected in the first display, when the user selects A path in the display activates in the graph of Figure 2 the specific information for that selection.

User Id	Domain SessionId	Domain UserId	NavPath	Page Url
NULL	7136d236-b62b-4391-9e62-9863ac339f50	7778a2c2-bc79-496f-8c70-88286297eedd	h1.Welcome.Pontoon.users-search-h1.Welcome.Pontoon.users-search-projects	https://es.voicebunny.com
NULL	7136d236-b62b-4391-9e62-9863ac339f50	7778a2c2-bc79-496f-8c70-88286297eedd	h1.Welcome.Pontoon.users-search-h1.Welcome.Pontoon.users-search-projects	https://es.voicebunny.com
NULL	7136d236-b62b-4391-9e62-9863ac339f50	7778a2c2-bc79-496f-8c70-88286297eedd	h1.Welcome.Pontoon.users-search-h1.Welcome.Pontoon.users-search-projects	https://es.voicebunny.com/first-project?lang=spa&g
NULL	7136d236-b62b-4391-9e62-9863ac339f50	7778a2c2-bc79-496f-8c70-88286297eedd	h1.Welcome.Pontoon.users-search-h1.Welcome.Pontoon.users-search-projects	https://es.voicebunny.com
NULL	7136d236-b62b-4391-9e62-9863ac339f50	7778a2c2-bc79-496f-8c70-88286297eedd	h1.Welcome.Pontoon.users-search-h1.Welcome.Pontoon.users-search-projects	https://es.voicebunny.com/first-project?lang=spa&g
NULL	064e3679-2c08-4989-b96c-7108f563509e	77423165-fad9-49cb-bee9-bf88218eaf0	h1.Welcome.Pontoon.users-search-h1.Welcome.Pontoon.users-search-projects	https://voicebunny.com/p
NULL	064e3679-2c08-4989-b96c-7108f563509e	77423165-fad9-49cb-bee9-bf88218eaf0	h1.Welcome.Pontoon.users-search-h1.Welcome.Pontoon.users-search-projects	https://voicebunny.com/s
NULL	064e3679-2c08-4989-b96c-7108f563509e	77423165-fad9-49cb-bee9-bf88218eaf0	h1.Welcome.Pontoon.users-search-h1.Welcome.Pontoon.users-search-projects	https://voicebunny.com/h/project?gfrman-acq-aff
NULL	064e3679-2c08-4989-b96c-7108f563509e	77423165-fad9-49cb-bee9-bf88218eaf0	h1.Welcome.Pontoon.users-search-h1.Welcome.Pontoon.users-search-projects	https://voicebunny.com/s
NULL	064e3679-2c08-4989-b96c-7108f563509e	77423165-fad9-49cb-bee9-bf88218eaf0	h1.Welcome.Pontoon.users-search-h1.Welcome.Pontoon.users-search-projects	https://voicebunny.com/h/project?gfrman-acq-aff

Figure 3: Table showing the detail of users who visited the page, user id if it was authenticated on the page, session, date and time

Finally, a data visualization in table format (Figure 3) with all the necessary information to know the detail of previously selected data, such as the user, the session, the date, time and the pages visited, is available.

With this visualization it is possible to detect how many users arrive and leave the website. If search is recategorized using the search parameters and pre-processing the data, you can take advantage of the visualization to identify the search types that a user performs, for example, in which languages or which type of voice are most wanted, Where the user is delayed more in a search and how long those sessions are, a situation that is perfectly adaptable to obtain the same information regarding a product in other types of industries.

CONCLUSIONS

This visualization tool will not only contribute to our client, it will also contribute to the general community interested in doing the analysis of the behavior of the users of its Web pages, allowing them to discover insights oriented to improve the usefulness of the Web site, to understand and to improve The interrelation with its customers and users, without the limitations that present similar tools that currently exist in the market, both open source and paid, as the previous parameterization, the inability to reprocess the data the use of sampling since the user is who Defines the volume, period, and number of records you want to scan.

REFERENCES

- [1] <http://pueyrredonline.com/blog>
- [2] <https://www.entrepreneur.com/article/267933>
- [3] <http://www.devservice.es/blog>
- [4] <http://www.overalia.com/blog/06/2010/google-analytics>
- [5] [Visualization Analysis and Design. Mundzer. Ak Peters / CRC Press, Oct 2014. Chap 9: Arrange Networks and Trees](#)