

Software Analytics for Web Usability: A Systematic Mapping

Lucas Henrique Pellizon¹, Joelma Choma², Tiago Silva da Silva^{1(✉)},
Eduardo Guerra², and Luciana Zaina³

¹ Universidade Federal de São Paulo - UNIFESP, São José dos Campos, SP, Brazil
silvadasilva@unifesp.br

² Instituto Nacional de Pesquisas Espaciais - INPE, São José dos Campos, SP, Brazil

³ Universidade Federal de São Carlos - UFSCar, Sorocaba, SP, Brazil

Abstract. Software usability has become a key factor for the success or failure of web-based systems. However, traditional evaluation methods – user tests and field observations – are expensive and time consuming when applied to a large number of users. In order to deal with these inherent difficulties and costs and to propose a new method to automatically capture and analyze web usage data, we carried out a systematic mapping on web analytics and web usability. A total of 970 studies were identified, of which only 42 studies were selected for this mapping. We found out that most studies are focused on tools for capturing information on user's navigation, however, few tools have presented mechanisms for visualization of these user interaction data.

1 Introduction

In the last years, software usability has been seen as an important software quality attribute. For example, several ISO (International Organization for Standardization) standards [19–21] include usability as a component of the overall quality of a software product [31]. For web-based systems, in which the distance from a competitor is just one click, usability became a key factor in its success or failure.

In general, usability evaluations are carried out through inspection or observation methods. Traditionally, the observation of the user interaction with software systems is based on carrying out user tests and field observations [51]. Although these methods are popular and efficient, they are expensive and time consuming. User tests are considered expensive because of their costs to find users to perform a test, to move the users to a usability laboratory, to set up the infrastructure, conduct the test, collect and analyze the data generated during the tests.

Field observations are also an expensive method due to their necessity to one or more usability analysts observing users while they interact with the software systems in their work environment for long periods of time. Observing users in their real work environment, performing their own tasks, allow us to keep the

context of usage, an important element in a usability analysis. Nevertheless, analyzing video recording or observing users by video-conference are also extremely time consuming activities.

Due to these costs, a common practice is to analyze the behavior of a few users in user tests or field observations. Analyzing the behavior of a few users limits the evaluation to just a qualitative analysis. Moreover, some problems can be just highlighted in a quantitative analysis and also their impact can just be evaluated considering a large number of users [51].

In order to deal with the difficulties, risks and costs of carrying out user tests in laboratories and/or observing users in field studies to analyze the usability of web applications, and aiming at proposing a new method to automatically capture and analyze web usage data, we carried out a systematic mapping about the employment of software analytics to enable web usability.

The remainder of this paper is organized as follows: Sect. 2 presents a short background on software analytics usability. Section 3 introduces the research method and presents our research questions. Section 4 presents our results and the answers for the research questions. Section 5 highlights some limitations of this study, and presents our final considerations and future work.

2 Background

In this Section, we present some concepts related to software analytics and web analytics, particularly focusing on navigation behaviour and usage data.

2.1 Software Analytics

Web systems generate a huge amount of information, which includes not only the sending and receiving data requests, but also actions and decision making processes in order to reach a certain goal. The software analytics area is concerned with collecting, exploring and analyzing this large amount of data aiming to transform it into actions to improve the product and its usability [7].

This study refers to the software analytics geared towards web-based systems – an area also known as web analytics [25]. Within this context, we have mainly focused on the means of capturing navigation and usage data, since these data can provide valuable information on navigation behaviour and how software systems are actually being used. Real usage data analysis collected from web logs, for instance, can lead to better understanding of how users use software in practice.

We list below some of the most common means used to capture navigation behaviour and usage data.

- Server Logger: a web server automatically saves all requests received by users, such as pages, images, files, among others services. Usually, these records contain the client IP address, the time and date of the request, the requested resource, the status of the request, the HTTP method used, the size of the

object returned to the client, and the referring web resource. Hence, of these logs can be extracted detailed information about the browsing behaviour of visitors to a website [47].

- Proxy Logger: a proxy-based system logs all communication between the client browser and the web server. However, the user's browser needs to be set up in order to direct web requests and captured events to the proxy-server. Basically, a proxy logger collects the same information type captured by a Server Log [16].
- Browser Plugin: a specific tool can be created for certain browsers in order to collect not only the requests made to the server, but also some behaviors, such as keyboard and mouse actions [37].
- JavaScript: a code is added to the website in order to inspect certain user actions, such as mouse and keyboard usage, for example. This JavaScript code must be added by the website developer himself, unless he chooses to take a integration with a proxy aiming to automatically add it to the page [44].
- Proxy + JavaScript: as mentioned in the previous item, the JavaScript code is automatically added through a proxy acting as an intermediary between the client and server. This approach has the same characteristics as the previous one and, in addition, it avoids the need to manually add the inspection code to the product. On the other hand, developers may be confronted with security issues [26].

2.2 Usability

Usability is a quality attribute that assesses how easy user interfaces are to use, in which the navigability is an important measure of how easily users can locate and access the necessary information to achieve their goals. According to [36], usability can be defined by five quality components: (i) Learnability: ease of learning; (ii) Efficiency: efficiency of use; (iii) Memorability: ease of remembering how to use; (iv) Errors: low rate of error during use; (v) Satisfaction: a pleasing user experience.

Online usability studies commonly use strategies for collecting both qualitative and quantitative data on either user attitudes or behaviors, and measuring the user experience. Quantitative metrics are useful, for example, to test the awareness of key features, or verify the intuitiveness of the navigation, while qualitative metrics can provide traceability into the nature of the problems that users encounter, or insight on possible design solutions [47].

3 Research Method

3.1 Systematic Mapping

Systematic mapping is a research method that provides an overview of a research area, and allows us to identify the quantity and type of research and results

available within specific area, or a phenomenon of interest. This research method consists of classifying, conducting thematic analysis and identifying publications in order to identify research gaps. Also, it can provide indications for lack of evaluation or validation research in certain areas with limited effort. Usually, systematic mapping studies do not study articles in depth in order to identify best practices based on empirical evidence. Nevertheless, it can be a first step toward a systematic review for further investigation on a specific topic [41].

For this study, we follow a mapping process based on the guidelines provided by [42], in which at least two of the authors participated in each step of the analysis of publications in order to allow us to immediately resolve disagreements as to how publications should be keyworded through discussion. The mapping process consists of four steps: (i) definition of research questions, (ii) search for relevant publications, (iii) screening of papers and (iv) mapping of publications.

3.2 Research Questions

The goal of this systematic mapping study is to get an overview of existing research on software analytics approaches focusing on web-based software usability – encompassing techniques, methods and tools commonly employed to capture and analyze navigation behaviour and usage data. In order to narrow the mapping study scope, we defined six research questions as follows:

- RQ1: How are the studies on web analytics and usability distributed over time?
- RQ2: What are the most common types of studies, and in what environments are they conducted?
- RQ3: What is the research focus of these studies?
- RQ4: What types of technologies have been used to capture navigation behaviour and usage data?
- RQ5: How can the captured information be visualized?
- RQ6: What is the relationship between capture technologies and data visualization?

Having defined the research questions, we selected some search keywords related to the two interest subjects: analytics and usability. Table 1 presents the keywords used in the search. Once we are focusing our study on web-based software, the term “web” is common for both areas.

3.3 Search Strategy

Taking into account that this is a preliminary study on the two aforementioned areas of interest, we chose to conduct a search for publications relevant by performing an automated search from a single database. Thus, we chose to use the Scopus digital library¹ because this library is an online abstract and indexing

¹ <http://www.scopus.com/>.

Table 1. Keywords related to interest areas.

Search terms	
Subject	Keywords
Analytics	capture
	collect
	analyze
	web
Usability	user behavior
	usability
	user experience
	web

service provided through Elsevier², which also indexes content provides by other digital libraries such as IEEE Xplore³, ACM⁴ and Springer⁵, as content coverage guide [45].

The search in the digital library was conducted in October 2016. According to our search terms, we used the following query expression in order to search in the electronic database: (capture **OR** collect **OR** analyze) **AND** (user behavior **OR** usability **OR** user experience) **AND** (web).

Furthermore, the Scopus search engine allowed us to specify the subject area (*Engineering or Computing*) and the document type (*Conference Paper, Article, Chapter or Article in Press*). This search resulted in 970 documents.

3.4 Screening of Papers

For the purpose of selecting only relevant studies aiming to answer our research questions, we applied the inclusion and exclusion criteria present in Table 2 to titles and abstracts.

Table 2. Inclusion and exclusion criteria

Screening criteria	
Inclusion	Exclusion
-Peer reviewed studies	-Non-English publications
-Addresses web analytics and usability	-Panel summaries, keynotes and posters
-Means for capture navigation behaviour	-Non-accessible in full-text
-Technologies for capture usage data	-Capture technologies based on Server Logs

² <https://www.elsevier.com.br/>.

³ <http://www.ieee.org/>.

⁴ <http://dl.acm.org/>.

⁵ <https://link.springer.com/>.

Table 3. Distribution of publication per digital libraries.

Papers source	
Database Name	# Relevant Papers
ACM	14–33.33%
Springer	9–21.43%
IEEE Xplore	9–21.43%
Other	10–24.81%

Based on the screening criteria, we selected 42 relevant papers, which are listed in the Table 4. Table 3 presents the distribution of these publications per digital libraries. ACM, Springer and IEEE Xplore stand out as the main sources of 76.1% of the publications (32 of 42). Despite we having selected only documents with title and abstract in English, we found out later a paper written in Portuguese by Brazilian researchers. However, due to the relevance of the study, we decided to keep this paper in our collection.

Table 4. Selected studies

ID	Paper	Ref.	ID	Paper	Ref.
1	Hong et al.	[16]	22	Peska	[40]
2	Paternó and Paganelli	[39]	23	Huang et al.	[17]
3	Waterson et al.	[53]	24	Deufemia et al.	[10]
4	Shahabi and Banaei-Kashani	[46]	25	Premchaiswadi and Romsaiyud	[43]
5	Niño et al.	[37]	26	De Vasconcelos et al.	[15]
6	Chui and Li	[8]	27	Li et al.	[27]
7	Cuddihy et al.	[9]	28	Apaolaza et al.	[2]
8	Arroyo et al.	[4]	29	Apaolaza	[1]
9	Atterer et al.	[5]	30	Nakano et al.	[34]
10	Ignatova and Brinkman	[18]	31	Uchida et al.	[48]
11	Mor et al.	[33]	32	Mao et al.	[30]
12	Zhu et al.	[55]	33	De Vasconcelos and Santos	[52]
13	Rivolli et al.	[44]	34	Breslav et al.	[6]
14	Kiura et al.	[24]	35	Papatheocharous et al.	[12]
15	Ding et al.	[11]	36	Zahoor et al.	[54]
16	Luna et al.	[28]	37	Apaolaza et al.	[3]
17	Vargas et al.	[49]	38	Arbelaitz et al.	[32]
18	Majji and Singh	[29]	39	Kalavri et al.	[23]
19	Vargas et al.	[50]	40	Gkantouna	[14]
20	Leal and Dias	[26]	41	Garcia and Paiva	[13]
21	Mayz et al.	[31]	42	Neelima and Rodda	[35]

4 Results

In this section, we present the results of the analysis and mapping of the selected papers answering our research questions.

4.1 Publications Distribution Over Time (RQ1)

In the set of relevant papers, the first papers to be published date from 2001 [16,39]. Nonetheless, no paper was selected in 2003 or 2004. The number of publications per year in the studied field increased moderately from 2011 onwards, as can be seen in Fig. 1. Also, we note that there was a peak in 2014 in which six papers were published, however, the average in the period 2011–2016 is four papers per year.

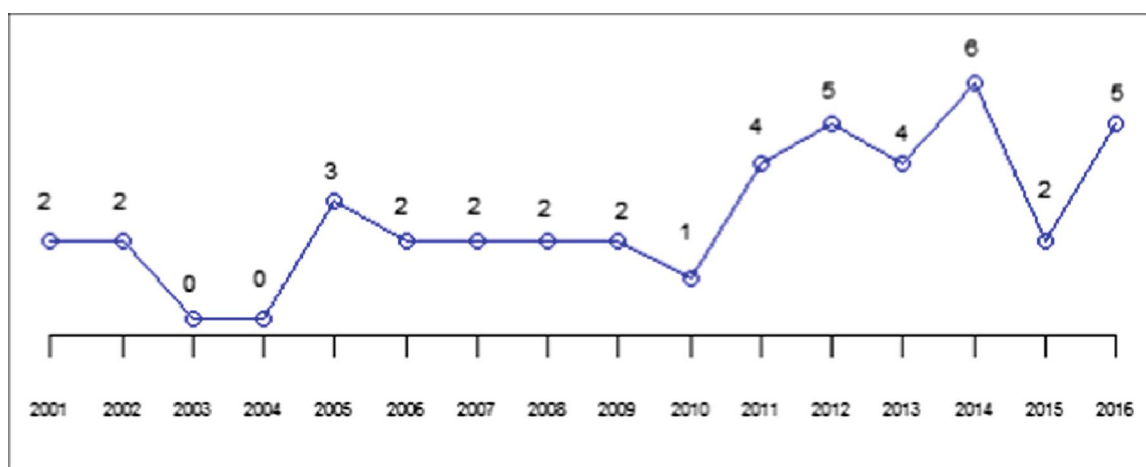


Fig. 1. Number of relevant papers on Web Analytics-Usability per year

4.2 Research Classification (RQ2)

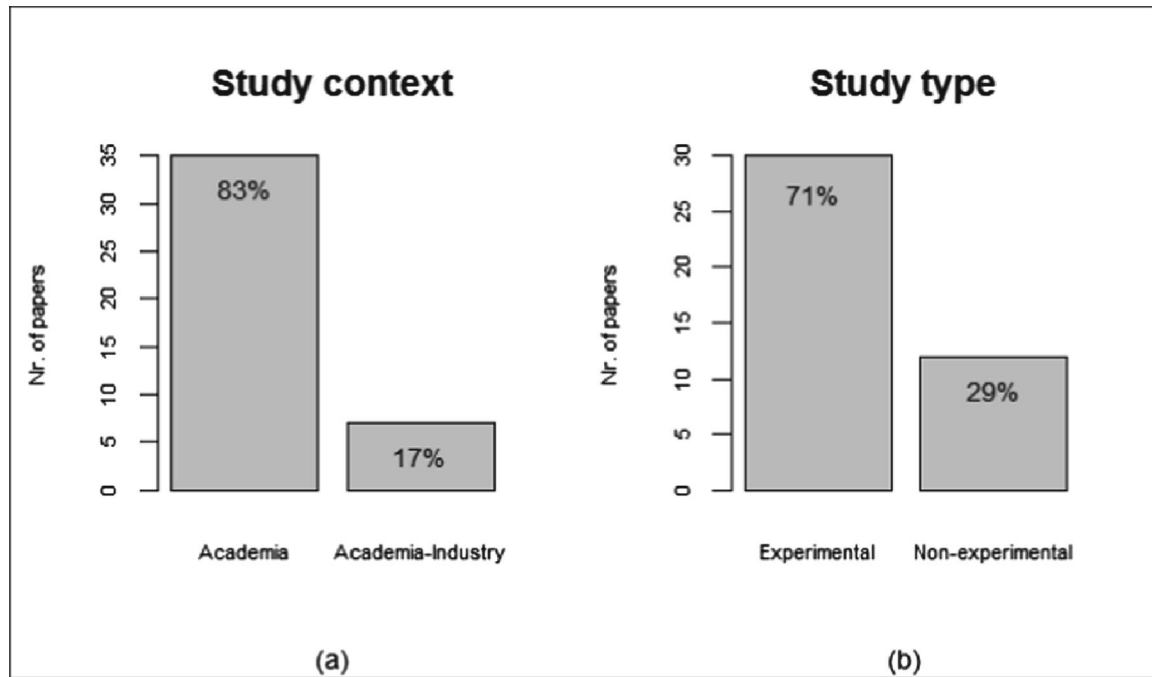
First, we classified the papers into three categories of research according to the context in which it was developed: (i) academic environment, (ii) industrial case study, or (iii) academia-industry partnership. As shown in Fig. 2-a, there are a small share of studies developed in partnership with industry (7 of 42), while we find no case study involving practitioners of the software development industry in this research follow-up. Thus, the greater part of the studies was developed in an academic environment (35 of 42).

Secondly, we classified the studies according to the type of validation into two categories: (i) experimental or (ii) non-experimental. We identified that 30 studies (71%) were experimentally validated, as shown in Fig. 2-b. However we have not yet evaluated the experimental studies in relation to compliance with guidelines proposed for reporting experiments in software engineering [22].

Table 5 summarizes the number of studies classified according to the context and type of research. We noticed that most experimental studies have been performed in academic contexts (24 of 30).

Table 5. Research classification according to context and research type

Context	Type	
	Experimental	Non-experimental
Academia	24–69%	11–31%
Academia-Industry	6–86%	1–14%

**Fig. 2.** Research classification

4.3 Research Focus (RQ3)

As for the research focus, when analyzing the selected set of papers, we identified three predominant type of research. Table 7 presents the papers distributed by these categories. More than half of the papers (52.4%) are focused on (i) “Tools for capturing navigation data” of which one part deals with both capture and data analysis (8 of 22), while another part (7 of 22) describes only about data capture mechanisms on user navigation. Kiura et al. [24], for instance, implements a tool that captures the user’s navigation and reproduces the data captured for further analysis, which enables to product owners to improve the product usability. In another example, Zhu et al. [55] propose an user behavior analyzer, which can help the product owner find faults without having to reproduce the navigation.

Also, we identified that 41% of papers are divided between (ii) “Analysis of navigation from existing tools” and (iii) “Real-time page customization”. With respect to the real-time page customization, we note that tools of capturing are as needed as the contextual analysis. Huang et al. [17] uses these user characteristics, for instance, to improve results of search engines. And finally, we found

Table 6. Research focus of the selected papers

Research focus		
#	Research focus	Papers
22	Tools for capturing navigation data	[1, 2, 4, 6, 8, 15–18, 23, 24, 26, 27, 33, 34, 37, 39, 44, 46, 48, 53, 55]
10	Analysis of navigation from existing tools	[3, 5, 9, 12, 13, 35, 43, 49, 50, 52]
7	Real-time page customization	[10, 28–32, 40]
3	Other matters	[11, 14, 54]

out that the three remaining papers refer to matters that are out of our research scope (Table 6).

4.4 Means for Capturing (RQ4)

As aforementioned, most of the studies are focused specifically on means and tools for capturing usage information. However, some other studies – whose main focus is not the means of capture – mention the type of technology used to capture the information about the behavior of the users. Thus, we also included these studies in our analysis about means of capture, totalling 32 papers examined.

We have identified four types of means for capturing: (i) JavaScript, (ii) Proxy Logger, (iii) Browser Plugin, and (iv) Java Applet. In the Fig. 3, we present the result of the mapping of such technologies distributed over time. Additionally, we listed the papers according means of capture mentioned, as showed in the Table 7.

Table 7. Means for capturing user navigation.

Types of means for capturing		
#	Means	Papers
21	JavaScript	[1–6, 9, 10, 15, 17, 18, 24, 26, 30, 34, 39, 44, 49, 50, 52, 55]
8	Proxy/Server Logger	[8, 16, 29, 32, 33, 35, 43, 53]
3	Browser Plugin	[37, 46, 48]

A large part of the studies (21 of 32–66%) uses the JavaScript technology as the main means to capture of client logs. Of these studies that use JavaScript, 13 studies (62%) insert JavaScript directly into the source code of the application, while the other 8 studies (38%) it is injected through a proxy (Table 4-a).

The second most commonly used means of capture is the Proxy Logger. We found four studies related to Server Logger [32, 33, 35, 43], which added to the Proxy Logger studies. Also, we found three studies about Browser Plugin,

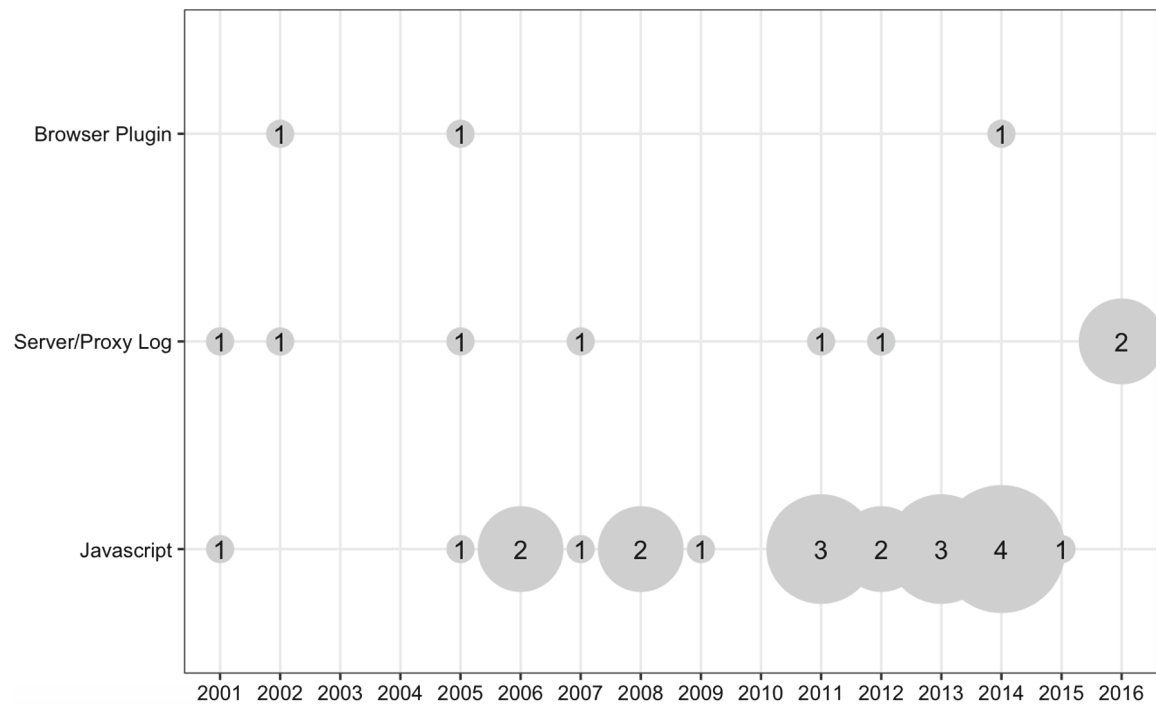


Fig. 3. Means for capturing user navigation over time

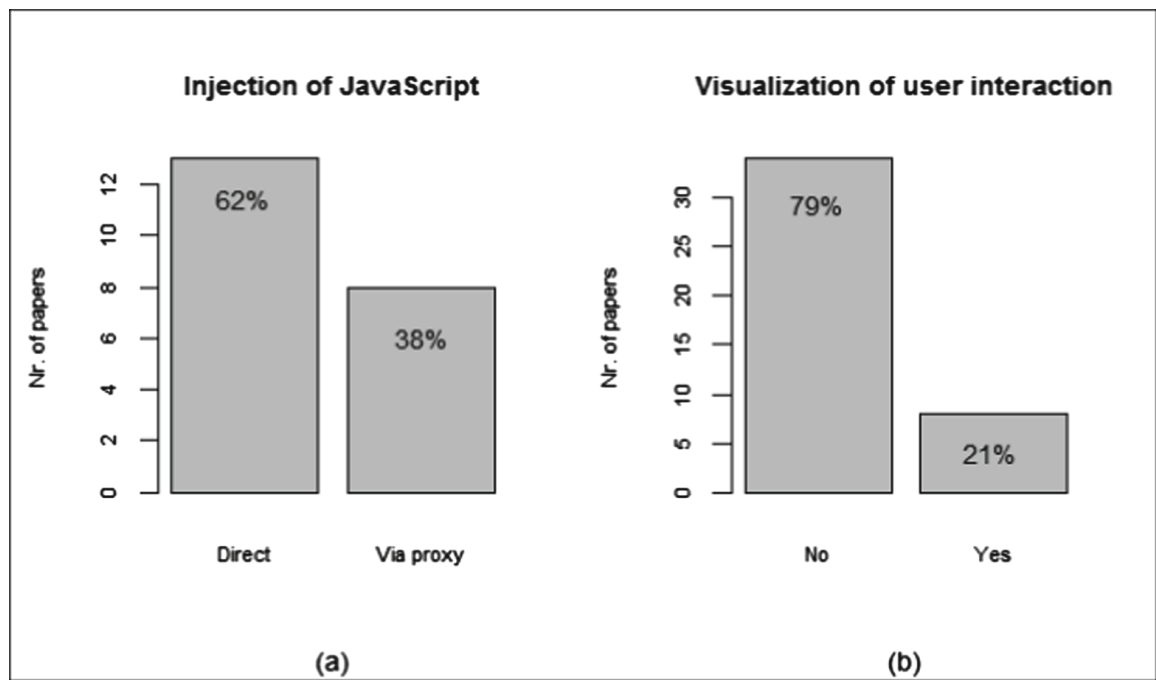


Fig. 4. Means of capture and visualization mechanisms

including a study about Java Applet published in 2001. Shahabi and Banaei-kashani [46] used the Java Applet technology to create a remote agent that transparently runs at the client machine, where is uploaded only once when the user enters the web-site. This agent tracks the user interactions only during a single session and does not store any information at the client machine. But over time, this technology has been replaced by technologies which do not rely on a browser plugin, once modern browser vendors decided restrict and reduce plugin support in their products [38].

4.5 Visualization of User Interaction (RQ5)

Mechanisms for viewing recorded information about the user interaction by tracking of screen activities – e.g., mouse movements, key presses, and log traces – are important to enable a posterior and more in-depth analysis on user experience with the software. However, when analyzing the set of studies, we observed that few papers report such mechanisms (see Table 4-b). We have identified four types of visualization: (i) page navigation, (ii) mouse trail, (iii) heat map graphics, and (iv) page elements.

Page navigation refers to the accessed pages by an user in order to accomplish a goal. *Mouse trail* is the result of the mouse movement of an user in a single page in order to identify it's path to reach a goal. *Heat map graphics* is a visual way to represent where are the users focusing in a certain page, by collecting mouse data. And, *page elements* refers to the HTML elements each user is accessing, which can be used for data mining techniques or others purposes. Table 8 presents the 9 studies that proposed mechanisms for visualizing data on user interaction, which we mapped according to the type of information visualized.

Table 8. Type of user interaction visualization.

Visualization of user interaction		
#	Type of visualization	Papers
3	Page navigation	[16, 37, 53]
1	Mouse trail	[5]
3	Heat map graphics	[4, 6, 24]
3	Page elements	[17, 24, 44]

4.6 Mapping Means of Capture and Types of Visualization (RQ6)

In order to verify the relationship between the types of visualizations and the means of capturing, we have analyzed the 9 studies that present mechanisms to visualize the data about the user interaction, as shown in Fig. 5. According to the means used to capture the information, we can notice that there are restrictions on the types of information visualized. In other words, *proxy logger* and

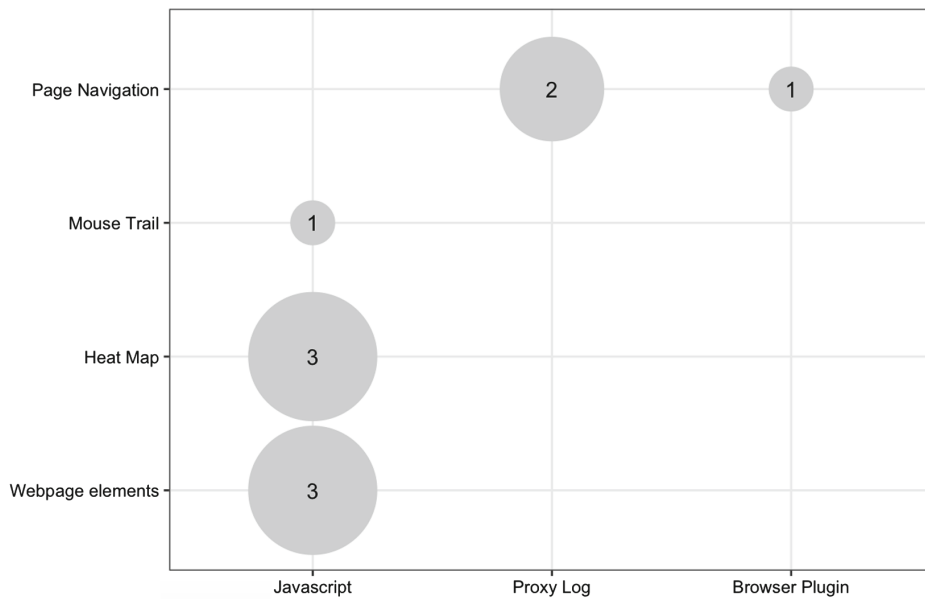


Fig. 5. Mapping means of capture and types of visualization.

browser plugin-based technologies offer limited resources, while *JavaScript* is a technology that allows us to implement more complex data capture functionality.

Niño et al. [37] proposed a client-side browser-embedded tool to capture and replay only user navigation sessions. While, Atterer et al. [5] use JavaScript tracking code in order to capture data about mouse movements, keyboard input and more, and then, mouse trails are combined to a screenshot of the website to show the user navigation.

5 Final Remarks and Limitations

In this paper we present the findings of a systematic mapping carried out to get an overview of existing research on software analytics for web usability, focusing on techniques, methods and tools commonly employed to capture and analyze navigation behaviour and usage data. We have selected and analyzed 42 relevant papers published between 2001 and 2016.

With regards to the threats to validity of our mapping study, we know that the selection of search terms and digital libraries can exclude some relevant studies in our search results. Furthermore, we searched only in a single library, and studies published in proceedings not indexed by this digital library may not appear in our analysis. Nonetheless, in this first stage, this was not our main concern since the intention was to have an overview about the research area.

Another limitation of this work is that we did not define a criteria for quality assessment of included studies based on the aims and research questions of our study. As regards the extraction and categorization process, it was carried out by the first author, while the first coauthor provided input to resolve ambiguities during the process.

Furthermore, a systematic mapping study can be considered less reliable than a systematic review when we usually consider only the abstracts and titles alone rather than the full text of each paper. However, aiming to answer some of our research questions (RQ4, RQ5 and RQ6), we had to analyze a considerable part of the text besides the title and abstract.

In response to RQ1, we found that the number of publications in this field increased moderately from 2011 onwards, and the average of the last six years is four papers per year. Regarding the research classification (RQ2), the majority of experimental studies have been developed in academic context.

To respond to RQ3, we identified three research focus: (i) “tools for capturing navigation data”, (ii) “analysis of navigation from existing tools”, (iii) “real-time page customization”. More than half of the papers are focused on tools for capturing navigation data.

Concerning the means for capturing navigation behaviour and usage data (RQ4), we identified three types of technologies: (i) JavaScript, (ii) Proxy Logger, and (iii) Browser Plugin. However, JavaScript is the most used technology.

With respect to the visualization of user interaction (RQ5), there are few studies that presented some mechanism for visualization of these data. And, analysing the relationship between capture technologies and data visualization (RQ6), we can notice that there are restrictions on the types of information visualized depending on the technology employed.

As future work, we intend to carry out a systematic literature review to answer more specific issues. With this purpose and based on the results of this mapping, we will adjust our search terms as well as consider more research sources, and also look at the references of the relevant papers.

Acknowledgments. This work was supported by FAPESP [grant number 2014/25779-3].

References

1. Apaolaza, A.: Identifying emergent behaviours from longitudinal web use. In: Uist, pp. 53–56 (2013)
2. Apaolaza, A., Harper, S., Jay, C.: Understanding users in the wild, pp. 13:1–13:4 (2013)
3. Apaolaza, A., Harper, S., Jay, C.: Longitudinal analysis of low-level web interaction through micro behaviours. In: Proceedings of the 26th ACM Conference on Hypertext and Social Media - HT 2015, pp. 337–340 (2015)
4. Arroyo, E., Selker, T., Wei, W.: Usability tool for analysis of web designs using mouse tracks. In: CHI 2006 Extended Abstracts on Human Factors in Computing Systems CHI 2006, no. (3), p. 484 (2006)
5. Atterer, R., Wnuk, M., Schmidt, A.: Knowing the user’s every move. In: Proceedings of the 15th International Conference on World Wide Web - WWW 2006, p. 203 (2006)
6. Breslav, S., Khan, A., Hornbæk, K.: Mimic: visual analytics of online micro-interactions. In: Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces - AVI 2014, pp. 245–252 (2014)

7. Buse, R.P.L., Zimmermann, T.: Information needs for software development analytics, pp. 987–996 (2012)
8. Chui, C.K., Li, C.H.: Navigational structure mining for usability analysis, pp. 126–131 (2005)
9. Cuddihy, E., Wei, C., Barrick, J., Maust, B., Bartell, A.L., Spyridakis, J.H.: Methods for assessing web design through the internet. In: Proceedings of ACM CHI 2005 Conference on Human Factors in Computing Systems, vol. 2, pp. 1316–1319 (2005)
10. Deufemia, V., Giordano, M., Polese, G., Tortora, G.: Inferring web page relevance from human-computer interaction logging, pp. 653–662 (2006)
11. Ding, Z., Jiang, M., Pu, G., Sanders, J.W.: Modelling and verification of web navigation. In: Gaedke, M., Grossniklaus, M., Díaz, O. (eds.) ICWE 2009. LNCS, vol. 5648, pp. 181–188. Springer, Heidelberg (2009). doi:[10.1007/978-3-642-02818-2_13](https://doi.org/10.1007/978-3-642-02818-2_13)
12. Papatheocharous, E., Belk, M., Germanakos, P., et al.: Towards implicit user modeling based on artificial intelligence, cognitive styles and web interaction data. *Int. J. Artif. Intell. Tools* **23**(2), 1–21 (2014)
13. Garcia, J.E., Paiva, A.C.R.: An automated approach for requirements specification maintenance. *Adv. Intell. Syst. Comput.* **444**(March), 827–833 (2016)
14. Gkantouna, V.: Mining interaction patterns in the design of web applications for improving user experience, pp. 219–224 (2016)
15. Guarino de Vasconcelos, L., Coelho dos Santos, R.D., Baldochi, L.A.: Exploring client logs towards characterizing the user behavior on web applications. In: Proceedings of SPIE - The International Society for Optical Engineering, p. 8758 (2013)
16. Hong, J.I., Heer, J., Waterson, S., Landay, J.A.: WebQuilt: a proxy-based approach to remote web usability testing. *ACM Trans. Inf. Syst.* **19**(3), 263–285 (2001)
17. Huang, J., White, R.W., Buscher, G., Wang, K.: Improving searcher models using mouse cursor activity. In: Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR 2012, p. 195 (2012)
18. Ignatova, E.D., Brinkman, W.P.: Clever tracking user behaviour over the web: enabling researchers to respect the user. In: BCS-HCI 2007 Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...but not as We Know it - vol. 2, 2 September 2007, pp. 179–182 (2007)
19. ISO/IEC: Iso/iec std. 9241–11: Ergonomic requirements for office work with visual display terminals. Part11: “guidance on usability”. Technical report, International Organization for Standardization (1998)
20. ISO/IEC: Iso/iec std. 9126–1: Software engineering - product quality. Technical report, International Organization for Standardization (2001)
21. ISO/IEC: Iso/iec std. 25010–3: Systems and software engineering: Software product quality and system quality in use models. Technical report, International Organization for Standardization (2009)
22. Jedlitschka, A., Ciolkowski, M., Pfahl, D.: Reporting experiments in software engineering. In: Shull, F., Singer, J., Sjøberg, D.I.K. (eds.) *Guide to Advanced Empirical Software Engineering*, pp. 201–228. Springer, Heidelberg (2008)
23. Kalavri, V., Blackburn, J., Varvello, M., Papagiannaki, K.: Like a pack of wolves: community structure of web trackers. In: Karagiannis, T., Dimitropoulos, X. (eds.) *PAM 2016*. LNCS, vol. 9631, pp. 42–54. Springer, Cham (2016). doi:[10.1007/978-3-319-30505-9_4](https://doi.org/10.1007/978-3-319-30505-9_4)

24. Kiura, M., Ohira, M., Matsumoto, K.: Webjig: an automated user data collection system for website usability evaluation. In: Jacko, J.A. (ed.) HCI 2009. LNCS, vol. 5610, pp. 277–286. Springer, Heidelberg (2009). doi:[10.1007/978-3-642-02574-7_31](https://doi.org/10.1007/978-3-642-02574-7_31)
25. Kumar, L., Singh, H., Kaur, R.: Web analytics and metrics: a survey. In: Proceedings of the International Conference on Advances in Computing, Communications and Informatics, pp. 966–971. ACM (2012)
26. Leal, J.P., Dias, H.: A framework to develop meta web interfaces. In: Proceedings of the IADIS International Conference WWW/Internet 2011, ICWI 2011, pp. 301–308 (2011)
27. Li, W., Cao, G., Qin, T., Cao, P.: A hierarchical method for user's behavior characteristics visualization and special user identification. In: IEEE International Conference on Networks, ICON (2013)
28. Luna, E.R., Garrigós, I., Rossi, G.: Capturing and validating personalization requirements in web applications. In: 2010 1st International Workshop on the Web and Requirements Engineering, WeRE 2010, pp. 13–20 (2010)
29. Majji, S., Singh, S.R.: Proxy server: integrating client side information and query (2011)
30. Mao, J., Liu, Y., Zhang, M., Ma, S.: Estimating credibility of user clicks with mouse movement and eye-tracking information. In: Zong, C., Nie, JY., Zhao, D., Feng, Y. (eds.) Natural Language Processing and Chinese Computing. Communications in Computer and Information Science, vol. 496, pp. 263–274. Springer, Heidelberg (2014)
31. Mayz, M.A., Curtino, D.M., De la Rosa, A.: Avoiding laboratories to collect usability data: two software applications. In: Sistemas Y Tecnologías De Informacion, vols. 1 and 2, pp. 145–149 (2012)
32. McKerlich, R., Ives, C., McGreal, R.: Measuring use and creation of open educational resources in higher education. *Int. Rev. Res. Open Distance Learn.* **14**(4), 90–103 (2013)
33. Mor, E., Minguillon, J., Santanach, F.: Capturing user behavior in e-learning environments. In: 3rd International Conference on Web Information Systems and Technologies (WEBIST 2007), pp. 464–469 (2007)
34. Nakano, A., Tanaka, A., Akiyoshi, M.: A preliminary study of relation induction between HTML tag set and user experience. In: Kurosu, M. (ed.) HCI 2014. LNCS, vol. 8512, pp. 49–56. Springer, Cham (2014). doi:[10.1007/978-3-319-07227-2_6](https://doi.org/10.1007/978-3-319-07227-2_6)
35. Neelima, G., Rodda, S.: Predicting user behavior through sessions using the web log mining. In: 2016 International Conference on Advances in Human Machine Interaction (HMI 2016), pp. 1–5, March 2016
36. Nielsen, J., Loranger, H.: *Prioritizing Web Usability*. New Riders Publishing, Thousand Oaks (2006)
37. Niño, I.J., De La Ossa, B., Gil, J.A., Sahuquillo, J., Pont, A.: CARENA: a tool to capture and replay web navigation sessions. In: 3rd IEEE/IFIP Workshop on End-to-End Monitoring Techniques and Services, E2EMON 2005, pp. 127–141 (2005)
38. Oracle's blogs: Moving to a plugin-free web (2015). https://blogs.oracle.com/java-platform-group/entry/moving_to_a_plugin_free
39. Paternò, F., Paganelli, L.: Remote automatic evaluation of web sites based on task models and browser monitoring. In: CHI 2001 Extended Abstracts on Human Factors in Computing Systems, pp. 283–284 (2001)
40. Peska, L.: User feedback and preferences mining. In: Masthoff, J., Mobasher, B., Desmarais, M.C., Nkambou, R. (eds.) UMAP 2012. LNCS, vol. 7379, pp. 382–386. Springer, Heidelberg (2012). doi:[10.1007/978-3-642-31454-4_41](https://doi.org/10.1007/978-3-642-31454-4_41)

41. Petersen, K., Feldt, R., Mujtaba, S., Mattsson, M.: Systematic mapping studies in software engineering. *EASE* **8**, 68–77 (2008)
42. Petersen, K., Vakkalanka, S., Kuzniarz, L.: Guidelines for conducting systematic mapping studies in software engineering: an update. *Inf. Softw. Technol.* **64**, 1–18 (2015)
43. Premchaiswadi, W., Romsaiyud, W.: Extracting weblog of Siam university for learning user behavior on mapreduce. In: 2012 4th International Conference on Intelligent and Advanced Systems (ICIAS 2012), vol. 1, pp. 149–154, June 2012
44. Rivolli, A., Marinho, D.A., Pansanato, L.T.E.: WAUTT: uma ferramenta para o rastreamento da interação do usuário com aplicações interativas web. In: Companion Proceedings of the XIV Brazilian Symposium on Multimedia and the Web, pp. 179–181 (2008)
45. Scopus, S.: Content coverage guide (2016)
46. Shahabi, C., Banaei-Kashani, F.: A framework for efficient and anonymous web usage mining based on client-side tracking. In: Kohavi, R., Masand, B.M., Spiliopoulou, M., Srivastava, J. (eds.) *WebKDD 2001*. LNCS, vol. 2356, pp. 113–144. Springer, Heidelberg (2002). doi:[10.1007/3-540-45640-6_6](https://doi.org/10.1007/3-540-45640-6_6)
47. Tullis, T., Albert, W.: *Measuring the User Experience, Second Edition: Collecting, Analyzing, and Presenting Usability Metrics*, 2nd edn. Morgan Kaufmann Publishers Inc., San Francisco (2013)
48. Uchida, H., Swick, R., Sambra, A.: The web browser personalization with the client side triplestore. In: Mika, P., et al. (eds.) *ISWC 2014*. LNCS, vol. 8797, pp. 470–485. Springer, Cham (2014). doi:[10.1007/978-3-319-11915-1_30](https://doi.org/10.1007/978-3-319-11915-1_30)
49. Vargas, A., Weffers, H., Da Rocha, H.V.: Discovering and analyzing patterns of usage to detect usability problems in web applications. In: *International Conference on Intelligent Systems Design and Applications, ISDA (Ic)*, pp. 575–580 (2011)
50. Vargas, A., Weffers, H., da Rocha, H.V.: A method for remote and semi-automatic usability evaluation of web-based applications through users behavior analysis. In: *ACM Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research - MB 2010*, pp. 1–5 (2010)
51. Vargas, A., Weffers, H., da Rocha, H.V.: Analyzing user interaction logs to evaluate the usability of web applications. In: *2011 3rd Symposium on Web Society*, pp. 61–67. IEEE (2011)
52. Vasconcelos, L.G.D., Santos, R.D.C.D.: Classifying user experience of web applications in real time using client logs (2014)
53. Waterson, S.J., Hong, J.I., Sohn, T., Landay, J.A., Heer, J., Matthews, T.: What did they do? Understanding clickstreams with the WebQuilt visualization system. In: *Proceedings of the Working Conference on Advanced Visual Interfaces AVI 2002*, p. 94 (2002)
54. Zahoor, S., Rajput, D., Bedekar, M., Kosamkar, P.: Inferring web page relevancy through keyboard and mouse usage. In: *2015 International Conference on Computing Communication Control and Automation*, pp. 474–478 (2015)
55. Zhu, Z., Wang, Q., Xu, H., Wang, H., Bing, L.: Research on a method of Ajax-based web user behavior collection. In: *2008 3rd International Conference on Pervasive Computing and Applications, ICPCA 2008*, vol. 2, pp. 965–969 (2008)