Usability Matters: A Human–Computer Interaction Study on Network Management Tools

Fábio Luciano Verdi, Hélio Tibagí de Oliveira, Leobino N. Sampaio, and Luciana A. M. Zaina

***Abstract*—Over the years, network monitoring tools have evolved to keep pace with the advances in networking technologies and paradigms, such as virtualization and softwarization, and more recently, data plane programmability. However, there have been few advances regarding user interfaces and interactions in network management systems. In this paper, an investigation based on the human-computer interaction perspective is presented and improvements on the usability aspects of network monitoring tools are proposed. First, we conducted a survey to capture general information from network administrator professionals. On the basis of these initial data, we identified and specified two personas to represent the groups of users. Thereafter, a study focused on observing users' interaction and experience with a well-known network monitoring tool was conducted. Finally, our qualitative analysis revealed several findings that were used to develop recommendations to aid developers in the (re)design of these tools to enhance end-user interaction.**

***Index Terms*—Network monitoring tools, network management, user-centered design, graphical user interface, usability, verification.**

I. INTRODUCTION

CURRENTLY, network monitoring supports different activities, such as those involving traffic engineering, capacity planning, advanced applications, service level agreement verification, security, and experimental networks [1]. As a result, different applications, tools, and measurement platforms are constantly emerging [2]. Most of these initiatives consist of developing multi-purpose operating and network management tools that frequently do not, however, support specific features for visualizing data (i.e., data view tools) as a means of enabling network administrators to enhance their human perception capacity and, consequently, improve their assertiveness in management and operational decisions. Coupled with that, usability has become even more important by virtue of the constantly increasing amount of information resulting from the rise in new networking trends [3], which leads to a great number of network components that must be manipulated during network operations.

Many of the existing management applications and tools are featured by poor interface resources because of the lack of knowledge regarding users' profiles. Although network administrators are technically trained and have good expertise in analyzing networking data, their work is impeded by a lack of well-designed tools to support the data view, which may cause usability issues. These two issues influence each other. Whereas visualization is the process of representing data, information, and knowledge in a visual form, usability is a factor in the effectiveness, efficiency, and satisfaction with which users achieve specific goals in particular environments [4], [5]. These issues together can trigger problems that go beyond the efficiency and effectiveness with which the users accomplish their tasks and may even lead to misunderstandings at the time they make their decisions [6]. The consideration of usability issues in the design of management applications and tools can affect the user experience (UX) of network experts.

Knowledge about the audience in the domain of network management allows developers to acquire a clear vision of the characteristics of those to whom they dedicate their endeavors in the development of monitoring tools. Consequently, they can focus their efforts on providing a more suitable design in terms of interaction to enhance the experts' work. In this perspective, recent reports in the literature indicate a need for further investigations of the interaction problems involved in human-centered monitoring tools [7]. However, few studies in this direction have been conducted, in particular regarding the identification and exploration of the profiles of network administrators. To address this lack of research, we conducted an investigation of usability aspects of network monitoring tools based on Human-Computer Interaction (HCI) to answer the following research questions (RQs):

- **(RQ1)** *What are the characteristics and interests that distinguish network administrators in the use of monitoring tools?*
- **(RQ2)** *What are the usability issues that arise during the execution of network monitoring management activities?*
- **(RQ3)** *How do network administrators classify UX and the tool's usability?*

To answer the RQs, the investigation described herein involved a study conducted in two phases. First, we performed a study to identify the characteristics of network administrators and their impressions of the systems used in the management and operation of computer networks (Study (i)). Then,

Manuscript received December 30, 2019; revised March 20, 2020 and March 29, 2020; accepted April 7, 2020. Date of publication April 10, 2020; date of current version September 9, 2020. This work was funded in part by the São Paulo Research Foundation (FAPESP) under grant number 2015/19766-9 and in part by the Brazilian National Council for Scientific and Technological Development (CNPq). The associate editor coordinating the review of this article and approving it for publication was D. Pezaros. *(Corresponding author: Leobino N. Sampaio.)*

Fábio Luciano Verdi and Luciana A. M. Zaina are with the Computer Science Department, Federal University of São Carlos, Sorocaba 18052780, Brazil (e-mail: verdi@ufscar.br; lzaina@ufscar.br).

Hélio Tibagí de Oliveira is with Atento Brazil, São Paulo 04726170, Brazil (e-mail: helio.oliveira@atento.com).

Leobino N. Sampaio is with the Computer Science Department, Federal University of Bahia, Salvador 40170-110, Brazil (e-mail: leobino@ufba.br).

Digital Object Identifier 10.1109/TNSM.2020.2987036

1932-4537 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Indian Institute of Space Science And Technology. Downloaded on October 16,2023 at 10:59:23 UTC from IEEE Xplore. Restrictions apply.

we investigated the network administrators' behaviors when, as participants in our study, they performed typical activities using a network monitoring tool. The objective of this investigation was to determine the usability issues that arose during the participants' interactions with the tool and the manner in which these issues could interfere with their UX (Study (ii)).[1] Our usability investigation aims to explore the barriers that end-users face when using the network monitoring tool as well as interacting with data views (e.g., charts, tables).

Hence, Study (i) served as the basis for addressing RQ1. To meet the aim of this part of the study, we invited network administrators to answer a questionnaire, which was available during the months of June and July, 2016. We obtained 70 responses. Subsequently, to identify user profiles more accurately, we used the *personas* technique, which is a very frequently used approach in the HCI research area. This technique can be applied to build a representation of the potential end users of an application, so that a software developer, for example, can establish an understanding of and empathic connection with them [8]. We explain the persona technique in Section IV.

Study (ii) aimed to answer RQ2 and RQ3. For this purpose, we invited nine network administrators who had taken part in Study (i) to participate in this one. They performed typical activities using Nagios, a popular tool used for network monitoring. Our choice was due to the results of Study (i), which indicated this tool as one of the most frequently used among the participants, despite the existence of other well-known and broadly used, as well, e.g., Zabbix. Given that our methodology is based on tasks executed by the network administrators, our results can be generalized to other tools since they use similar interface resources from the perspective of monitoring tasks, as reported in [9]. Additionally, we used HCI techniques and prepared a qualitative data collection process following the rigorous proceedings of previous studies of end users [10]. Subsequently, our findings enabled us to infer conclusions that can be applied to similar tools.

This paper is organized as follows. Section II summarizes related work on usability in network monitoring tools. Section III discusses the survey, revealing the network administrators characteristics and Section IV presents the personas built from the survey data. Sections V and VI discuss the study in which we examined the use of a monitoring tool and the study results respectively. Section VII details the findings and the recommendations we found out from the study and Section VIII presents our final considerations.

II. Related Work

The research presented in this paper is new and there is a lack of comparable work to discuss herein. However, this section shows some initiatives that have investigated the characteristics of the audience for network management and monitoring tools. For instance, the European GÉANT network, through the Special Interest Group on Performance Monitoring and Verification (SIG-PMV), led by the Special Interest Group–Network Operations Centres (SIG-NOC), conducted a survey to gather more information regarding the software tools that network operation centers (NOCs) use to conduct operational activities. It involved operators, network managers, and users of monitoring applications, and its purpose was to identify the profiles of those interested in network performance information [11]. Twenty-five responses to the survey were collected, predominantly from network administrators located in Europe having six years' or more experience. The results revealed that 68% of the respondents chose open-source tools for network monitoring. However, the outcomes showed that, in management activities, traffic monitoring and the use of the simple network management protocol (SNMP) are the most common activities. In addition, the survey also addressed issues related to network performance and the responses revealed that certain types of software, such as Wireshark and Iperf, are used extensively.

In addition to GÉANT's initiative, the growth of the software-defined networking (SDN) paradigm has led to additional initiatives to meet new monitoring requirements. In [12], the authors proposed a management cycle where specific SDN performance metrics can be monitored, processed, and displayed through interactive visualizations so that the network administrator can reach decisions and make the required adjustments. It is noteworthy that user interaction is of great concern and a latent demand exists for the application of optimal data visualization in SDN settings, because the granularity of measurement data may be even finer than in traditional networks [12].

The method presented by Pretorius *et al.* [13] uses eye tracking combined with traditional methodologies for evaluating interfaces. The results of the eye tracking tests revealed that the important region of a monitoring tool's data view is very small, reducing data readability. Additionally, the authors showed that users always prefer to view graphics rather than text. Finally, their study concluded that the eye tracking technique allowed crucial discoveries, which could not be discovered by applying only traditional techniques.

A comprehensive survey related to the usability of firewall configuration was conducted by [14]. After analyzing a very large set of papers, the authors summarized usability issues and recommendations for personal firewalls, listed in Table 9 of their paper. However, the study was focused on the UX only in terms of firewall configuration and did not include experiments involving administrators.

Although the research area of UX in the context of network monitoring tools is not new, to the best of our knowledge, the study presented in this paper is the first that encompasses all the steps necessary to derive a set of recommendations for designing the user interface and interaction of network monitoring tools based on the HCI perspective. The network monitoring tasks were rigorously chosen so that the results of the study with users could be generalized to monitoring tools beyond that used in this paper, i.e., Nagios. We provide an analysis of the usability when the users manipulated network monitoring tools. We also conducted a study with network administrators in which we created a simulated network monitoring environment that enabled us to observe them performing

[1]The materials used in these studies are packaged and can be downloaded at http://networkviz.azurewebsites.net/.

Authorized licensed use limited to: Indian Institute of Space Science And Technology. Downloaded on October 16,2023 at 10:59:23 UTC from IEEE Xplore. Restrictions apply.

typical network monitoring tasks. We recorded the entire study on video, and the recordings were analyzed in-depth with the help of HCI methods.

III. Investigating the Profile of the Network Administrator

In Study (i), we sought to collect data about the users of a network monitoring tool following the principles of the user-centered design (UCD) approach [15], [16]. In a UCD approach, users' goals, behaviors, needs, and tasks are at the core of the system development. By using knowledge of the users' profiles, tools can be constructed that offer features that are more suitable for their daily tasks. Therefore, we conducted a survey to obtain this information, which provided the data required for Study (ii). Additionally, the findings of the first study represent an important contribution to the network monitoring area. To the best of our knowledge, there exist no other studies in which the profile of network administrators was explored in detail. Our survey was supported by the application of HCI techniques recognized for their rigor in scientific research [10].

A. The Survey

We followed a well-known methodology to conduct the survey. We first obtained certain information from specialists and researchers in the computer networks area. To accomplish this task, an online form was designed,[2] consisting of 22 objective questions and 5 open questions, where users could write their own point of view. The form was divided into five categories: (i) demographic information, (ii) monitoring tool used and its functionalities, (iii) visual relevance, (iv) SMS notifications,[3] and (v) other information.

In particular, one of the open questions was aimed to identify the functions of the monitoring tool that the respondents used most frequently. This question was included because of the large number of features in current network monitoring tools and the difficulty in enumerating them accurately. Thus, to evaluate the number of existing functions, the authors of this paper carefully read and interpreted each user's responses. Functionalities were grouped based on the number of times they were mentioned in different responses.

We recruited network managers as our target participants to fill the online form. Invitations to participate in the study were sent through two communication channels: computer network discussion groups and professionals known to our research group. These channels allowed a diverse sample of researchers and professionals in the network monitoring area to be reached, and consequently, we could obtain a heterogeneous sample of individuals [15]. No incentive or reward was offered to respondents for their participation.

[2]https://networkviz.azurewebsites.net/IEEE-TNSM/survey-form/Survey.pdf

[3]At the beginning of our project, we talked with network administrators, who informed us that SMS is one of the main message systems used for notification. The explanation they gave refers to the fact that mobile phones operate well with old cellular technologies, such as 2G, and 3G, which provide connectivity even in remote areas. Therefore, SMS remains a lightweight, well-accepted, and frequently used message system.

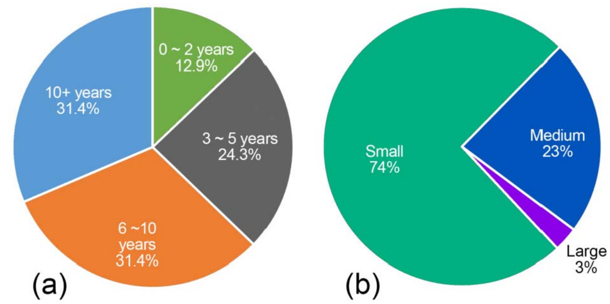

Fig. 1. (a) Amount of experience in years and (b) size of the company in terms of number of network devices.

The online form was open to respondents in June and July, 2016; we obtained 70 responses in total. Before taking part in the survey, participants were asked to agree to the terms of consent regarding the use of data and images for academic purposes and they were assured that their identities would be known only to the authors of this study.

To handle the data practically, the responses were imported into a relational database. Then, each answer was analyzed to identify problems and inconsistencies, thus ensuring that only meaningful data were considered. To support this analysis, respondents' explanations and additional information filled in an open free text field helped us to understand the answers of the participants. Invalid inputs, such as "-" or "00000," were not considered. Finally, several SQL queries were performed to find crossover of information, to discover correlations among the data, and for constructing graphs to obtain a better interpretation of the data.

After exploring and analyzing the data, we were able to identify the personas, which are described in detail in Section IV.

B. Exploring the Results

Fig. 1(a) shows the experience levels of the network administrators who took part in the survey. The size of the company where each respondent worked is displayed in Fig. 1(b). We clustered the participants according to the size of their company to observe the data grouped into a desired number of clusters. For the analysis, we used three clusters and categorized the clusters by the size of the companies in terms of the amount of their network equipment [17]: (i) small (0–200 network devices), (ii) medium (201–1000), and (iii) large (more than 1000 network devices).

We noticed that the respondents were quite heterogeneous in terms of their practical experience, with a very even distribution of years of experience. An analysis of the graphs in Fig. 1 shows that more than 62% of the respondents had more than six years' experience in the field, indicating that the survey participants could be considered experts in the network administration role. In addition, 74% of respondents belonged to small companies with a maximum of 200 network devices. Only 3% of the participants worked in a company with more than 1000 network devices.

For the question about their certification in the areas of monitoring or management of networks, the vast majority of the

Authorized licensed use limited to: Indian Institute of Space Science And Technology. Downloaded on October 16,2023 at 10:59:23 UTC from IEEE Xplore. Restrictions apply.

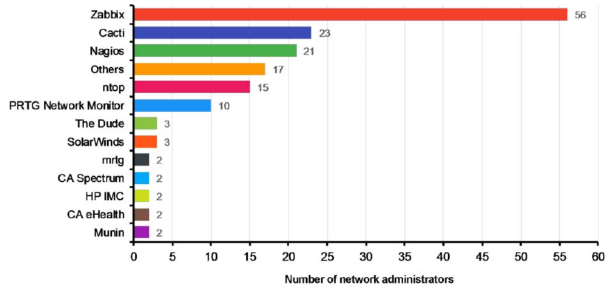

Fig. 2. Network management tools used by the respondents.

participants (81.4%) responded that they had no certification. Those who reported otherwise mentioned the following certifications: 11.5% Zabbix, 2.9% CISCO, 1.4% Linux LPI, 1.4% Microsoft MCP, and 1.4% Oracle.

Fig. 2 shows the number of mentions of a particular monitoring tool. Zabbix was the tool most frequently used by the participants, followed by Cacti and Nagios. It is noteworthy that in their response to this question, the respondents could indicate more than one tool that they used in their work.

Fig. 3 presents the network monitoring and management features most frequently used by the respondent network administrators. The relationship between the experience of the network administrators and the tool features is represented by the segments in the bars showing their distributions. The network administrators indicated that the Internet Control Message Protocol (ICMP ping) and the Simple Network Management Protocol (SNMP) for traffic measurement are the functions with the most frequently features.

The analysis of the features they used revealed that network administrators with up to two years' experience predominantly use the basic functions of a monitoring tool, frequently those that are most easily configured. ICMP (ping) and traffic measurements are prominent with four and five citations, respectively.

However, participants with experience in the 3–5 year range are used to dealing with more features, and some of these features are more complex in terms of the levels of interpretation or configuration. Instead of ICMP checking, which is widely used by professionals with up to two years' experience, this group employed SNMP most frequently, which delivers detailed information about the host or network equipment. It is important to mention that traffic measurements that were highlighted by participants whose experience was up to two years can also be performed by means of data arriving via SNMP. The use of alerts to take actions according to certain triggers previously defined by the network administrator is also highlighted in the data.

Users with 6 to 10 years' experience cited 20 functionalities, which is a reflection of the maturity and diversity of these network administrators. The conclusion is that some network administrators, in addition to practicing comprehensive monitoring, use more specific resources.

Finally, network administrators with more than 10 years' experience continue to use the same monitoring features as those with less experience, although with a slight difference: ICMP (ping) is again highlighted, showing that, although it is simple and frequently used by those with less experience, it is also widely employed by experienced users.

IV. Personas: Who Are the Network Administrators?

In the persona technique, a group of users is described by using fictional persons [8]. A persona is derived by analyzing the data collected from surveys conducted among target user groups. The technique is widely used in software development to gain a better understanding of the users who will be using a system. It also enables developers to visualize clearly the relevant aspects of the application's target user group during software development [8]. Primarily, developers use the persona technique to enable them to be aware of the general characteristics and needs of a target group rather than focusing on one single person's characteristics.

To design the personas, we followed five of the six steps proposed by Pruitt and Adlin [8]. The final step was not included because it consists of the validation of personas from the perspective of domain specialists, which is not mandatory when the raw data are available for the researcher. In our study, a questionnaire was administered, as explained in Section III-A.

In the first step, the researcher defines the domain from which the data of users will be collected. In this step, user roles, user goals, and user segments can be referenced. In our work, we used the user role "network administrator" to define the domain. Taking into account the domain features, the instruments to collect the data are prepared. In the second step, the collected data are imported to a relational database. The researcher then applies filters to the data, removes inconsistent data, and finds the crossover of different information. This step results in an overview of the data that the researcher has in hand to support the elaboration of the personas. These two first steps were performed during the survey (see Section III-A).

Then, in the third and fourth steps, we proceeded to the identification and categorization of users by applying the Automa-Persona process [18]. It consists of running an algorithm to separate the personas into different clusters according to their similar characteristics. The idea is focused on the design of each persona based on a single set of information obtained through the segmentation of very large amounts of data. To find each persona, features that were found repeatedly were grouped by running database queries and by crosschecking data.

The groups of users were identified through a systematic process that relied on a k-means clustering algorithm [19]. To meet this aim, we conducted a data crossing process supported by the Weka framework,[4] considering two features relevant to the context of network management in each execution. Examples of pairs of data were: (i) "*the amount of network equipment available in the company*" vs. "*years of experience in network management*" and (ii) "*years of experience in network management*" vs. "*user's opinion on using SMS as notification channel*." Through this process, we examined

[4]https://www.cs.waikato.ac.nz/ml/weka/

Authorized licensed use limited to: Indian Institute of Space Science And Technology. Downloaded on October 16,2023 at 10:59:23 UTC from IEEE Xplore. Restrictions apply.

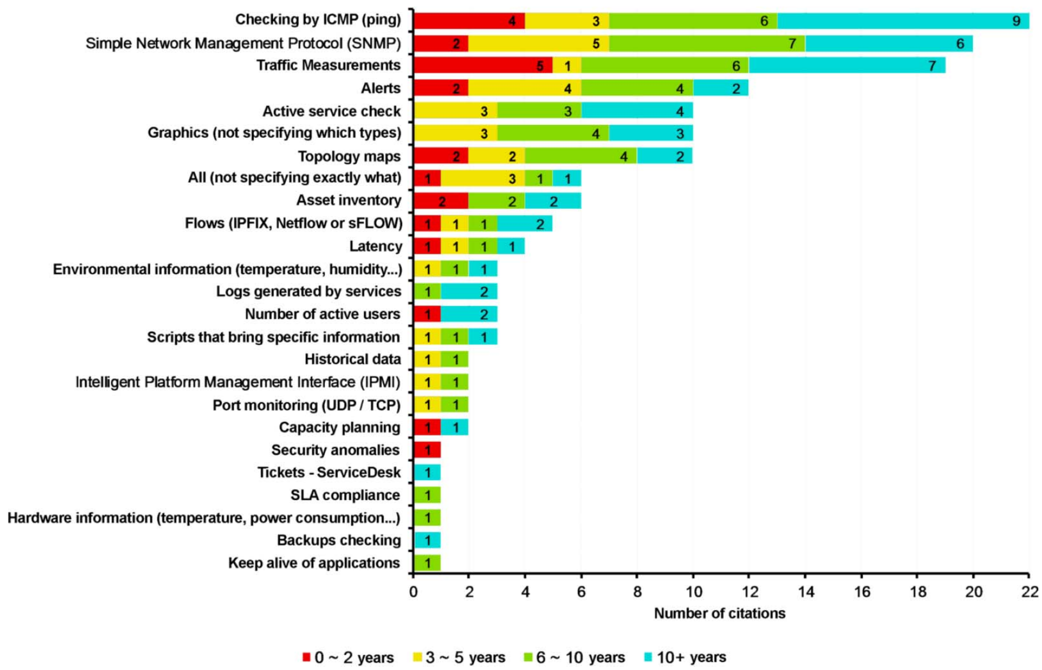

Fig. 3. Functions most frequently cited by the network administrators, segmented by years of experience.

all the data provided by respondents to execute an exhaustive data analysis.

Initially, we supposed that three groups of managers would exist, those working in small, medium, and large companies. After running the algorithm, some mismatches among managers were found. For instance, users who managed 600 devices were clustered with those who managed 4000 devices. However, in a previous work [17], we classified the size of the companies as small (0 to 200 networking devices), medium (201 to 1000 networking devices) and large (more than 1000 devices), and as such, we understand that the complexity for managing a network with 4000 devices brings more difficulties (in terms of quantity of data to be monitored, failures, capacity planning, logs, addons, protocols, services, etc.) than managing a network with 600. So, by adding the size of the company information to the algorithm, the clustering grouped small size companies in a group (74% of the respondents) and medium and large size companies in another group (26% of the respondents).

The discussion above shows that we relied on the observation of all the data and the k-means algorithm. Because we had access to all the users' data, we could identify the best set for our study.

In the fifth step, the creation of personas is supported by the findings of the previous steps. There are several means of constructing a persona. In this research, we modeled the personas according to the method proposed by Gothelf and Seiden [20], as this model presents the relevant characteristics in a lean and direct format by dividing its information into four quadrants. In Quadrant I, the personalization is realized by the presentation of a fictitious name and person's photo, providing the user with a means of referencing the character, as well as of bearing in mind the character's visual representation. The demographic information that illustrates the relevant factors in the user group profile is described in Quadrant II. The main behaviors that describe the user group, such as ordinary and recurring activities, are entered in Quadrant III. Finally, Quadrant IV presents the main needs and objectives of users, that is, the goals they want to achieve and the manner in which they intend to fulfill the tasks. After mapping the data into the quadrants of the persona model described above, we designed the personas of our domain (see Table I). The design of the personas was guided by the answers gathered from the online survey and the characteristics of the groups found in the clustering process.

TABLE I
MAPPING OF THE QUESTIONNAIRE QUESTIONS FOR THE PROTO-PERSONA TEMPLATE

Quadrant	Questions used
I	Respondent's name.
II	Years of experience; company; certification; number of network devices; number of end users in the company.
III	Network monitoring tools being used; how the tool is used; opinion about SMS notifications; features used in the monitoring tool.
IV	Level of importance of the visual interface; features used; open question allowing the respondent to express his/her own opinion.

The first persona represents a senior network administrator and the second a novice network administrator, i.e., a junior in the employment market. In contrast to our persona results, personas that do not match a traditional market segmentation could sometimes be found [15], mainly because the personas were based on the behaviors and attitudes related to the individuals' interactions on a system and not on ordinary features used daily.

Through the mapping shown in Table I, we could answer RQ1 based on the two personas that we derived: Robert Brewer (see Fig. 4) and Josh Baker (see Fig. 5). These personas cover the characteristics of the different profiles of the participants. Persona Robert represents the characteristics of

Authorized licensed use limited to: Indian Institute of Space Science And Technology. Downloaded on October 16,2023 at 10:59:23 UTC from IEEE Xplore. Restrictions apply.

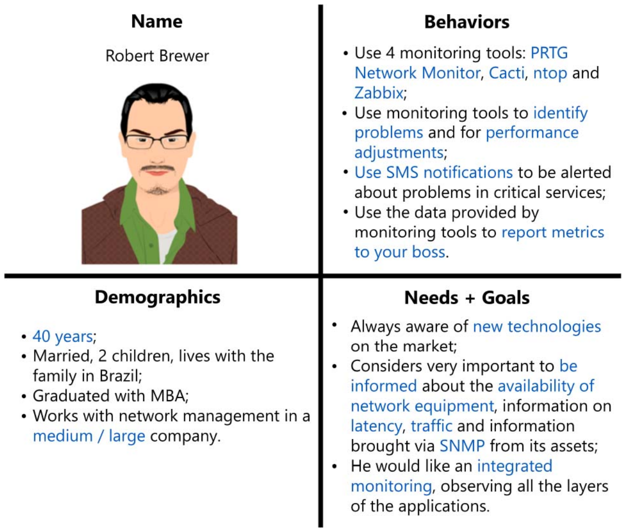

Fig. 4. Persona Robert Brewer, representing a professional with more expertise.

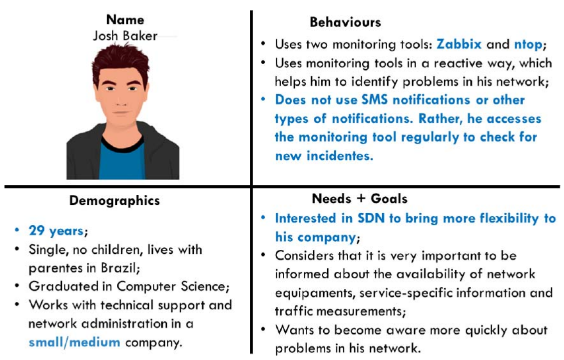

Fig. 5. Persona Josh Baker, representing a network administrator with less experience.

professionals with more expertise, whereas Josh represents those of network administrators with less experience. The relevant persona's characteristics are highlighted in blue in Fig. 4 and 5.

V. Studying End Users' Interactions: Network Monitoring Tool and User Experience

When the relevant features of the network administrators had been identified, we conducted a study based on the observation of network administrators' interactions with a well-known network monitoring tool. In this study, we followed the guidelines of Lazar *et al.* [10] regarding the observations of end users' interactions. Because the study was focused on collecting and analyzing qualitative data, we did not employ inferential statistical analysis or validation of hypotheses in our data analysis [10]. The descriptive statistics were used only to support the discussion of the qualitative results. The following subsections describe in detail each step (planning, execution, and analysis) of the study.

A. Planning

To observe the interactions of the users and generate the required artifacts, we prepared a network environment as

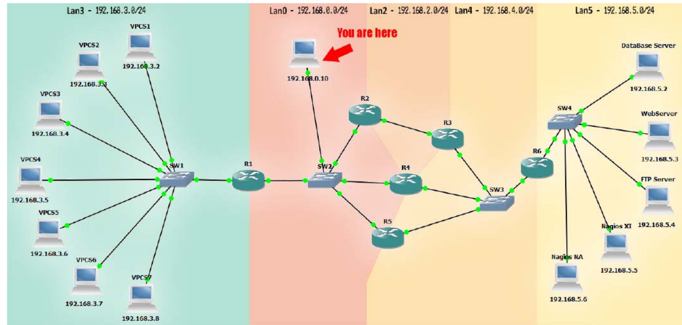

Fig. 6. Topology of the simulated network. The red arrow indicates the machine used by the participant.

follows. (i) A set of tasks to be performed by the participants of the study, i.e., network administrators, was prepared. (ii) A network monitoring environment in which the participants would run the tasks, i.e., the simulation of a network to be monitored by a network monitoring tool, was created. (iii) A script to support the study was developed. (iv) Various instruments for collecting the data of the participants' interactions were prepared. The following subsections present the artifacts that were prepared to support our study.

1) Tasks: With the objective of tracking network administrators' interactions with the tool, we designed tasks, identified from the results of the survey, that encompass the main features of monitoring tools. Some of these features are ICMP checking, SNMP reading, and traffic measurement. The tasks were designed to avoid obvious or non-guided use of a feature. We avoided tasks such as: "Tell me how much traffic in bytes is passing through port 1 of the router." This type of question would be very synthetic and would not provide a full and extensive sample of the troubleshooting that network administrators face daily. Table II presents the four tasks of this study.

2) Setup of the Network Environment: Our objective was to create a network environment setup to reproduce all the adverse situations that may appear during a task. Thus, this environment was required not only to allow the injection of failures, but also to generate network traffic. To this end, we used the network emulator GNS3 [21] with the VMWare Workstation hypervisor to create the network setup. In addition, 13 virtual machines (VMs) were created to act as desktop computers and servers, and 6 CISCO 7200 series routers and switches were instantiated as network elements. A topology (see Fig. 6) was built to provide multiple routes between the desktop computers and servers to increase the degree of difficulty of the tasks. The entire environment ran on a PC Workstation with an Intel Core i7-6700 CPU and 16 GB of RAM. The simulations were performed through scripts and the traffic was generated using Iperf.

To meet the study's goals, we needed to choose a network monitoring tool. We used as criteria that the tool should be commercially well-known and established and could also provide all the necessary functionalities for fulfilling the tasks. Consequently, we chose the NagiosXI [22] tool coupled with

Authorized licensed use limited to: Indian Institute of Space Science And Technology. Downloaded on October 16,2023 at 10:59:23 UTC from IEEE Xplore. Restrictions apply.

TABLE II
LIST OF TASKS THAT A PARTICIPANT PERFORMED. LEVEL OF COMPLEXITY: THE NUMBER OF STEPS REQUIRED TO ACCOMPLISH A TASK

Id	Title	Description (scenario of use)	Complexity	Task's goal
I	Conges-tion Analysis	There is a network congestion consuming almost the entire available bandwidth and causing high latency on some nodes. Identify the highest traffic generators and consumers and the location of the bottlenecks.	High	To take measurements of traffic and latency and identify sources and destinations of flows in a congestion scenario. The network administrator could make the decision to divert traffic and/or relieve the equipment causing the bottleneck, thus returning services to normal.
II	UDP/TCP Traffic	Identify the total traffic in bytes/s related to the HTTP or HTTPS protocol (TCP ports 80 and 443) in Network 3. The computers in Network 3 are accessing Websites from a server dedicated to this function in Network 5.	Medium	To filter the traffic in the granularity of TCP streams, and at the same time, to summarize everything from a specific network. Supported by the information, the network administrator can scale the load of a Web server and plan when to scale his infrastructure.
III	Router Flow Dump	Show by either a table or a graph, using any view you find, all the flows that have passed through the R6 router in the last 24 hours, regardless of the source or destination of these flows.	Medium	To extract a sample with all the flows of a router. The data obtained at the end of this task could be used for an audit, identifying security breaches and anomalies with the help of an adjacent system.
IV	Interface Down	An interface of a network's equipment will fail purposely, simulating a physical failure. Indicate which interface has failed, and what you visualized to reach that conclusion.	Low	To detect a serious fault quickly, thus verifying that the network administrator will respond to an alert. In response to this notification, the network administrator may act on the physical fault by reestablishing the connection of the link.

NagiosNA [23], which is the commercial version of the traditional Nagios open-source project.[5]

We cannot guarantee that user interface elements of each network monitoring tool are the same, present the same positioning of the objects on the screen, and adopt the same colors. However, the literature shows that they have similarities and support features for monitoring tasks [9]. Therefore, the recommendations presented herein are generic for the network monitoring domain so that any software developer can apply or somewhat extend them to (re)design any network monitoring tool interface.

To ensure the homogeneity of access to the monitoring tool, we built a remotely accessed VM through which a Google Chrome browser was available. The Google Chrome browser was considered the most widely used browser on desktop computers worldwide [24] at the time when the study was conducted. The screen resolution for the remote access device was set at 1366×768 pixels. This was an important variable in the context of this study, because the study required the interface of all the users to be the same, with objects in the same location and of the same size. The configuration values adopted for the screen resolution were also those most frequently applied for desktop computers worldwide [25] at the time of the study.

3) Scripts to Support the Study: So that the conditions and guidance during the study would be the same for all participants, a script was developed to guide the researcher who ran the study. The requirements requested for performing the study were: (I) an Internet connection of at least 1 Mbps; (II) a computer with a minimum specification compatible with the Intel Core2Duo 2.4 GHz processor, 4 GB of RAM, and a monitor with a resolution of at least 1366×768 pixels; (III) the use of a physical keyboard and mouse; (IV) Mac OS 10.10 or higher, Windows XP or higher, or Linux with a desktop environment; (V) an instant messenger application for making a voice call to receive directions and to allow screen sharing; (VI) a silent environment without noise or distractions.

The script also contained how and when interventions should take place in the environment. Typical interventions consisted of executing a command batch file to generate a certain type of traffic, to cause a failure, or even to start or shutdown a VM. Two running scripts were designed to randomize the order of task execution [10]. One script (Tasks I to IV) started with the tasks that were considered more complex and the second (Tasks IV to I) started with the simpler tasks to evaluate how the order could affect the execution of the tasks. The participants were divided proportionally between the two scripts, according to their belonging to the senior and junior persona, respectively. A warm-up activity was planned and performed to leverage each participant's experience with the tool.

4) Instruments for Gathering Data From Participants' Interaction: Two well-known instruments were applied to collect the UX and usability perception. Both reported the perspective of the end users, i.e., the participants, about their interaction [26], [27]. We chose the Self-Assessment Manikin (SAM) instrument to collect the participants' UX. SAM is a pictograph evaluation method for measuring emotional responses to a certain type of stimulus. Three dimensions are considered by this technique: pleasure (whether the reaction of the participant was positive or negative), arousal (body stimulation level from an event or object), and dominance (feeling in control of the situation or controlled by it). The user chooses a value on a scale of 1 (worse) to 9 (better) on each dimension, using images, to represent their emotions after interactions [28]. In our study, we did not use the arousal dimension, because we in fact asked users to execute the tasks. Thus, no self-motivation measurement was considered.

The System Usability Scale (SUS) [26] was adopted to evaluate the usability perspective. SUS consists of 10 item-questions that allow the measurement of a system's usability. It supports the calculation of a numerical score from 0 to 100 (this is not a percentage) from three criteria: (i) Effectiveness, which verifies whether the users can complete their goal;

[5]Nagios Enterprises is aware of this research and formally allowed the submission of this paper. Examples of Nagios' screens can be found at this link: https://networkviz.azurewebsites.net/IEEE-TNSM/nagios-images/.

Authorized licensed use limited to: Indian Institute of Space Science And Technology. Downloaded on October 16,2023 at 10:59:23 UTC from IEEE Xplore. Restrictions apply.

(ii) efficiency, which defines the amount of effort and resources users need to achieve their goal; and (iii) satisfaction, which reflects the user's feeling regarding the experience. The score represents the level of usability. Scores under 70 indicate that it is likely that the system presents usability problems.

Additionally, we planned the collection of qualitative data using which the analysis on the researchers' perspective could be performed. The interactions of the participants with the tool were recorded on video by using the Open Broadcaster Software (OBS-Studio) tool.[6] To obtain an alternative source of data, we asked the participants to speak their thoughts aloud while interacting with the tool by applying the think aloud technique [10].

B. Execution

All the network administrators who participated in Study (i) were invited by email to join Study (ii). In this invitation, we did not provide details of the study. We mentioned only its purpose, its duration, and that participants would perform an activity using a network monitoring tool. No additional information was provided and thus we avoided a situation where the participants would have prepared themselves in advance, which could consequently cause unevenness in knowledge among the chosen participants.

A total of nine network administrators participated in Study (ii), four representing the senior persona, and five representing the junior persona. Based on the HCI literature, nine participants are within the range of acceptability and significance.[7] It also states that the quality of the group is so important for its representativeness as the number of participants [6], [30]. So, we carefully selected nine users from 70 network administrators who answered the Survey to guarantee such quality.

All participants accepted the terms of consent concerning the use of data and images for academic purposes. First, we informed the participants that they could request a break, as well as choose not to perform an activity. However, no participant took advantage of either of these options. We also established 1 hr as the time limit for accomplishing a task, and the participant being analyzed was told this. This duration was compatible with the effort demanded for each task. However, this information was conveyed only for motivational purposes, so that the user would not lose focus and thus not fulfill the proposed objective of the task. We started the time measurement when the user clicked on the "login" button to enter the tool and ended it when the participant verbally manifested "I finished," indicating that he/she had achieved the task goal. All the instructions, links to remote connections to the environment, and the SAM and SUS questionnaires were hosted on a Website. The image of the topology was available at all times, so that the participant could refer to it during the accomplishment of the tasks.

[6]https://obsproject.com/

[7]The literature reports that a group of 3-20 participants is considered valid [29].

We observed each participant individually. The study was conducted during October and November, 2016. First, the participants took part in a warm-up session that provided them an opportunity to explore the environment. They were informed that their interactions with the tool and their speech would be recorded. In total, 4 hr and 17 min of video recordings were collected. At the end, the participants filled the SUS and SAM questionnaires.

C. Analysis

We performed two analyses. First, we consolidated the SAM and SUS responses. Then, we examined the videos containing the participants' interaction and their speech to conduct a qualitative analysis. This exploration was done for each participant individually.

A first exploration was done by the second author of this paper. He watched all the videos and annotated relevant parts. In total, 137 annotations were carefully written based on behaviors observed during the interaction with the tool and on the transcriptions of each participant's verbal communication. Basically, he observed usability problems when questions regarding the navigation and understanding of interface elements indicated obstacles in the participant's interaction including difficulty to analyse information presented as a large set of data or graphics.

Then, the second author shared these results with the three other authors, who reviewed, discussed, and refined all the annotations. Subsequently, we conducted the open coding activity to uncover the issues that arose during the participants' interaction. Open coding is a technique that consists of examining data and labeling chunks with codes that give them meaning. It does not use a predefined set of codes; on the contrary, the codes emerge from the data [10]. To achieve this, we examined all the annotations together and not individually for each participant. In our data exploration, we took into account that a code should occur at least twice to be considered as such. We ran the coding iteratively: one author suggested a coding and the others discussed and refined it. The results are explored in the following section.

VI. Exploring the Results

We first present an overview of our findings and subsequently answer RQ2 and RQ3.

Table III shows the coding extracted from the data, including the polarity and description. The descriptions give the explanation of each coding meaning. Polarity analysis is used in a variety of application domains. In the financial domain, for instance, it facilitates understanding by means of social network posts whether a particular company has a good business reputation, allowing companies to observe whether their customers are satisfied with a particular product. Many scholars have used artificial intelligence techniques to interpret text through its conversion into its meaning (positive, negative, or neutral). In this study, given the small amount of data, the polarity was obtained manually and classified only as positive or negative. Thus, all the extracted coding was divided according to polarity, explicitly indicating whether it reflected

Authorized licensed use limited to: Indian Institute of Space Science And Technology. Downloaded on October 16,2023 at 10:59:23 UTC from IEEE Xplore. Restrictions apply.

TABLE III
CODING EMERGING FROM THE DATA

Coding	Polarity	Description
Correctly guided	Positive	The user followed the correct direction by using the comprehensive mechanisms of the tool, which supported his successful completion of the task.
I got a barrier	Negative	The user had to overcome an obstacle to continue the task.
I need help	Negative	The user requested or demonstrated a need for help to proceed further.
I found help	Positive	The user was helped by an automation or tool (such as a search box, a sort button, or a magnifier to see a graph in detail) inside the view that allowed him to progress in his task.
Noise	Negative	The user was faced with inaccurate or incorrect information that impeded the execution of a task.
I am lost	Negative	The user was eventually led down the wrong path (which would not lead to the solution of the task) or demonstrated confusion in the use of the system.
I cannot do this	Negative	The user demonstrated or reported that he did not know or could not use the interface visualization or its interactive mechanism.
Not compatible with my previous knowledge	Negative	The user looked at or interacted with something that did not match his previous knowledge, and therefore, proved ineffective.
I do not usually look at this	Negative	The way the view was presented was not compatible with the user's culture.
I got mad/furious	Negative	The user manifested repudiation of the situation.
Praise of a characteristic	Positive	The user expressed satisfaction with the data visualization.
I gave up	Negative	The user no longer showed interest in continuing the task because of the difficulties he faced.

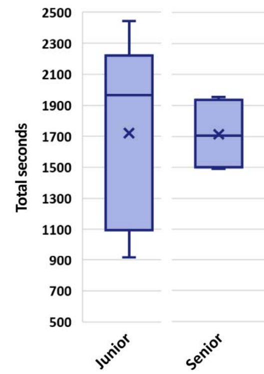

Fig. 7. Time taken to accomplish the tasks: seniors × juniors.

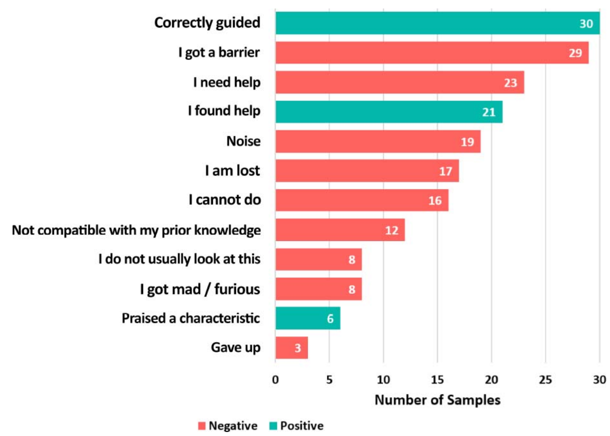

Fig. 8. Coding frequency: green is positive and red is negative.

a positive or negative opinion regarding the interaction with the system. The results presented in the table will be used in Fig. 8 to count the frequency of every code.

The time that each participant took to accomplish the tasks varied significantly between the two personas, as can be seen in Fig. 7. Juniors exhibited a wide variation, from a minimum of 917 sec to a maximum of 2443 sec. Although the seniors on average did not complete the tasks equally quickly, their variation was lower than that of the juniors, with a minimum of 1491 sec and a maximum of 1880 sec. These numbers together with other evidences allow us to observe that although all the users had no previous knowledge about the monitoring tool, the experience of the seniors are homogeneously consolidated making the time variation lower.

To address **RQ2**, we analyzed the number of occurrences of each code listed in Tab. III. In the graph in Fig. 8, it can be seen that the monitoring tool guided the user correctly to achieve the solution of the problem in 30 occurrences. However, several problems in the accomplishment of the tasks arose during the use of the tool. The coding "I got a barrier" was found in 29 interactions. Evidence such as "did not understand the meaning of "*" in this visualization $\cdots$" or "a pie chart with very thin slices which overlapped labels that could not be read" were reported by the users. The "I need help" coding arose 23 times. These evidences thus demonstrate that many visualizations and interactive mechanisms were not sufficiently self-explanatory to allow the user to proceed without help.

In Fig. 9, we show a polarity comparison of the issues reported by the users to identify the positive and negative aspects that, respectively, helped or hindered them in performing the tasks. This analysis revealed that the functionality of *Queries*, with 40 occurrences, was the issue that caused the greatest number of negative perceptions in the users' interactions. Although it was observed that the degree of dominance of the query language can affect the perception of the user, administrators who had good knowledge faced difficulties similar to those of the users with less dominance of the query language. These results show that visualization is an issue present in user interaction regardless of the experience of each individual user.

To answer **RQ3**, we explored the data in the perspective of the participants. We analyzed the data collected from the SUS and SAM questionnaires. Table IV shows the relations among the SUS questions, the responses of users (Un), and the matching of users to the personas. The last row of Table IV shows the SUS score. According to these results, we can conclude that the usability of the tool is graded between "poor" and "acceptable," regardless of the persona group, i.e., senior or junior.

In the analysis of the SAM responses, we observe in Fig. 10 that, in general, the indexes in the dominance dimension of SAM of the senior group persona were higher than those of the junior group persona. This result means that these participants felt in control of the interaction with the tool. However,

Authorized licensed use limited to: Indian Institute of Space Science And Technology. Downloaded on October 16,2023 at 10:59:23 UTC from IEEE Xplore. Restrictions apply.

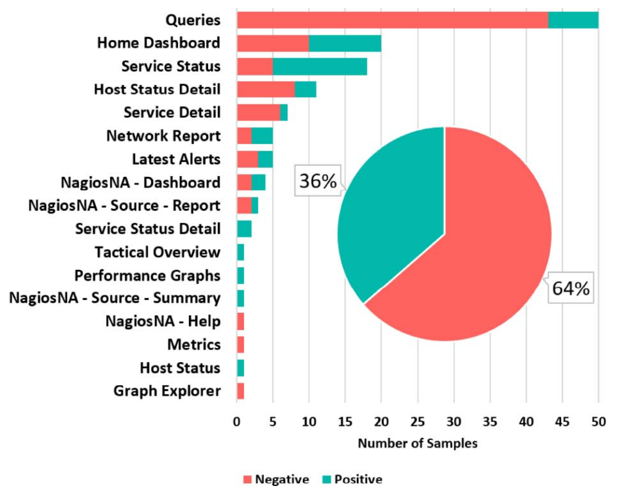

Fig. 9. Polarity per data visualization issue of NagiosXI. Green signifies positive and red negative.

TABLE IV
USERS' RESPONSES ON SYSTEM USABILITY SCALE DISTRIBUTED ACCORDING TO PERSONAS

	Senior persona group				Junior persona group				
SUS question	U1	U2	U3	U4	U5	U6	U7	U8	U9
Q1	4	3	5	2	5	2	2	4	4
Q2	3	2	4	2	2	3	5	2	1
Q3	3	4	4	4	4	3	1	3	5
Q4	2	2	1	2	4	2	4	4	1
Q5	4	3	3	3	5	3	2	5	5
Q6	3	2	3	2	2	1	3	1	1
Q7	4	4	2	4	4	2	1	4	4
Q8	3	1	2	2	1	3	4	3	2
Q9	3	3	3	3	4	3	2	3	5
Q10	4	2	2	3	4	4	2	4	2
SUS Score	42	50	52	57	47	50	50	42	50

SUS statements: Q1 - I think that I would like to use this system frequently; **Q2 -** I found the system unnecessarily complex; **Q3 -** I thought the system was easy to use; **Q4 -** I think that I would need the support of a technical person to be able to use this system; **Q5 -** I found that the various functions in this system were well integrated; **Q6 -** I thought there was too much inconsistency in this system; **Q7 -** I would imagine that most people would learn to use this system very quickly; **Q8 -** I found the system very cumbersome to use; **Q9 -** I felt very confident using the system; and **Q10 -** I needed to learn many things before I could start working with this system.

Response scale: 1 - Totally disagree; 2 - Disagree; 3 -Indifferent; 4 - Agree; 5 - Totally agree.
SUS score: Less than 40 indicates *poor* usability; between 40 and 50 is considered *acceptable*; greater than 70 is considered *good*; and greater that 85 is considered *excellent*.

for certain tasks, we see that the medians for both groups range from 4 to 6, which means the participants perceived that they did not have control in some tasks. The results of the pleasure dimension show that the feedback of both personas was positive (see Fig. 11). In particular, the simplest task, "Interface Down," received higher values of pleasure from junior persona (ranking from level 8 to 9). However, the values of the senior group for the same task were more widely distributed. This could be a consequence of the fact that the senior professionals demanded more elaborate features through which their decisions to handle network problems and improvements can be supported.

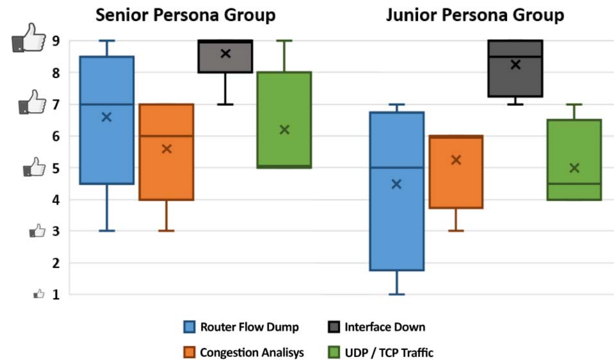

Fig. 10. Self-Assessment Manikin: Dominance dimension results.

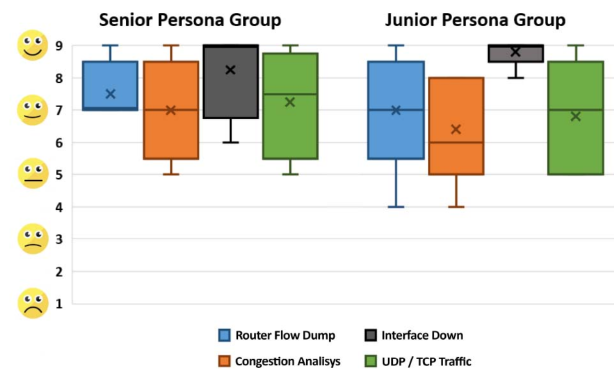

Fig. 11. Self-Assessment Manikin: Pleasure dimension results.

VII. FINDINGS AND RECOMMENDATIONS

The participants' perceptions discovered during the interactions were compiled and classified. The process for organizing the findings followed these systematic steps. (i) The chunks of participants' verbal communications collected from audio and video recordings were examined, and notes were generated. (ii) By analyzing the notes one by one, we determined the aspects that were mentioned at least twice. As a result, we identified certain usability issues. These usability issues are in line with the codings presented in Tab. III which were our guidelines to observe the audio and video recordings. Every time that a coding appeared, we mapped it to an issue.

This process was conducted by the first and second authors for the first interaction. Then, the rest of the authors discussed and revised the results. Considering these results, we created Table V, in which the findings are arranged in four columns: issue identification (id), issue name, description of the cause, and finally, possible solutions to the issue. For every cause, the evidence collected, labeled ECn,[8] is shown, where n is the number of the evidence in the external table.[9] The items in Table V are sorted out by the issue similarities.

After we had examined in detail the findings of the study (see Table V), our next step was to formulate the recommendations that can contribute to the design or redesign of network monitoring tools by considering the usability issues.

[8]Label ECnRO means that the evidence emerged through the researcher observation using in-depth analysis of the video recordings of every user.

[9]https://networkviz.azurewebsites.net/IEEE-TNSM/AuxiliaryTable.pdf

Authorized licensed use limited to: Indian Institute of Space Science And Technology. Downloaded on October 16,2023 at 10:59:23 UTC from IEEE Xplore. Restrictions apply.

TABLE V
FINDINGS

Id	Issue	Description of the cause	General comment
1	Fill-in help	(i) Looked for fill-in features during the activities [EC1]; (ii) requested confirmation that s/he had built the filters correctly [EC2]; (iii) attempted to use wildcard characters when attempting to set a filter for the entire network [EC3].	A user interface that guides users during the filling of fields and consequently prevents user errors.
2	Logical expres-sions	(i) Did not realize how to apply the truth table [EC19]; (ii) had doubts about how to construct filters using "and" and "or" because of a misunderstanding during the application of the truth table [EC19, EC20RO].	Some network administrators are frequently confused in the construction of logical conditions (truth tables) when making filters. Several errors were made when it was necessary to represent this type of filter using text.
3	No filter visualiza-tion	(i) Could obtain a general data overview by not filling in the filter fields [EC45RO]; (ii) made a mistake when attempting to run the query without any fields filled in, because his/her intention was to display an overview (without filters) [EC46RO].	The option of not creating filters should be considered in the system so that users can first explore the big data picture and then filter the data to obtain a detailed, in-depth visualization.
4	Percep-tion of updating	Automatic refreshing of the interface for visualization aided him/her to find the item s/he was looking for quickly [EC4].	Provide support of the perception of movements and transitions of the interface to guide the attention of users by delivering the notion of "before-after" updating.
5	Lack of visibility of inter-action trace	(i) Appeared to be lost because of the loading of a new interface caused by his/her actions (e.g., the click on a link that provided more detailed visualization on data of the R5 router) [EC16RO, EC17]; (ii) became confused because of the change in the level of detail of the visualization presented [EC17, EC18RO]; (iii) became somewhat lost upon changing the data filter topic [EC18RO].	Because the feeling of being lost on the user interface can negatively affect interactions and data visualization, users should always be kept informed about what is happening when they navigate the system.
6	Informa-tion filtering	(i) Reacted positively after finding a router through a searching feature [EC9]; (ii) filtered the data visualization by selecting a group of equipment [EC9]; (iii) struggled to find and select the target equipment because a large list was returned [EC11RO].	Provision of a searching tool to support the finding of items which is made difficult by a large amount of information.
7	Detailed informa-tion	With the mouse, s/he was able to view the value of the X-axis and Y-axis, the curve together with a gauge visualization, and pie chart values [EC12, EC13].	User analysis of information can be improved by adding and providing more detailed information such that it is accessible.
8	Irregular in-depth naviga-tion	(i) Realized that even when clicking on port 0 or 1 of the router, the same visualization appeared [EC24]; (ii) was informed that, depending on the link accessed, the same page is reached; however, the data that were visualized changed [EC25]; (iii) complained about incorrect information presented by the tool [EC26].	The user expects to obtain more detailed information when his/her navigation flows to in-depth queries. However, when this does not happen, the user's reasoning is broken, which can create some confusion.
9	Sorting of data	Reported the existence of a feature to sort information of hosts and classified traffic information by clicking the table headers [EC14, EC15RO].	Functions for sorting the data help users find the target data quickly through a table visualization.
10	Spy before going deeper	(i) Attempted to find more details about the network, but an overview of the data had been obtained previously through the available charts [EC27RO]; (ii) discovered the IPs of the relevant hosts so that s/he could complete the task by examining the top five generators of traffic in the network [EC28]; (iii) confirmed traffic congestion between generators and consumer by checking IPs of the top 10 traffic generators [EC29]; (iv) based on a preview of data summarization, "peeked" at the data of the previous 30 minutes of traffic before reaching a hypothesis regarding which routers were causing the congestion [EC30RO].	Provision of visualizations containing summaries of the whole or partial data would help guide the user to accomplish a task and make appropriate decisions.
11	Starting point	(i) Identified problems needing attention, immediately upon entering into the tool [EC40]; (ii) accessed the link for the "untreated" problems and went straight to a problem from the dashboard [EC41RO]; (iii) was able to reach the cause of the problem directly by using the block of anomalies presented in the dashboard [EC42RO].	Dashboards are a good starting point for obtaining an overview of the network, incident counters, problem counters, and other items previously configured to raise alerts.
12	Suitable arrange-ment of data	The user (i) noticed that some thin slices of a pie chart were overlapping so that labels became unreadable [EC34RO] and (ii) accessed the dashboard and complained that the interface elements were cluttered and were overlapping each other [EC35].	Dynamic information that does not fit elements appropriately into the available space frequently becomes incomprehensible.
13	Percep-tion of colors	(i) Reacted according to information noticed in highlighted colors [EC5]; (ii) easily found the services in critical situations through highlighted information [EC6]; (iii) clicked on the hyperlink appearing in red; reported a critical problem by virtue of information highlighted in red [EC6, EC7, EC8]; (iv) perceived a number of items marked in red, and reported a change in the status [EC7, EC8].	Meaningful colors, e.g., "red = problem" and "yellow = alert," are useful for drawing the attention of users for the identification of abnormal situations in the network.
14	Unsuit-able symbol metaphors	(i) Was not able to determine what the asterisk ("*") meant in the table view [EC36RO]; (ii) moved the mouse over each icon looking for captions or tips in an attempt to discover icon meanings [EC37RO]; (iii) complained when s/he could not recognize which icon would lead to the desired visualization [EC38]; (iv) was confused concerning the asterisk symbol ("*") [EC39].	Some symbols may cause misinterpretation because of the different meanings they communicate, and therefore, the symbol metaphor should be closely related to the context of use and the user's previous knowledge.
15	Symbols represent-ing status	(i) Noticed an alert icon (yellow triangle with an exclamation mark) and succeed in reaching a conclusion [EC43]; (ii) with the help of icons, was able to identify alerts and critical situations [EC44RO].	Symbols that represent the status of some metrics were well understood. Therefore, symbols should be used to improve system-to-user communications, as opposed to only text.
16	Pop-ups	(i) Asked about intention to disable a pop-up that appears on the Nagios main page [EC47]; (ii) when the tool presented a pop-up automatically, s/he completely ignored and simply closed it [EC48RO].	In consideration of the widespread use of pop-ups in Web advertisements, pop-ups are generally labeled as something that is not important, and users usually ignore any information presented in this manner. Thus, pop-ups for communicating information should be avoided.
17	Memory overload	(i) Made notes on the Notepad tool regarding the hosts experiencing a higher traffic load and the routers through which the traffic was passing [EC31RO]; (ii) opened a new visualization in another browser tab to maintain the current visualization of routers and hosts [EC32RO]; (iii) switched between the topology provided to support the task (to see the router) and the topology presented in Nagios (to check the bandwidth consumption) [EC33RO].	Lack of information in a visualization can overwhelm a user's memory, and consequently, users can lose self-confidence when they need to make decisions.
18	Giving up	The user gave up when (i) exploring the interface because of the absence of information about traffic measurement [EC21], (ii) performing a filter setup task that failed because of negligence in fulfillment of a required field [EC22RO], and (iii) experienced a frustrating attempt to access assistance in the help page support [EC23RO].	Many users give up on interacting with the system when the interaction design does not clearly communicate its requirements or the interface adds barriers that prevent them from accomplishing their tasks.

The purpose of these recommendations is to guide developers and designers in the construction of software. They operate as a memory aid concerning the essential aspects that can improve the user interaction features and support users in accomplishing their tasks while expending minimum effort [31].

Authorized licensed use limited to: Indian Institute of Space Science And Technology. Downloaded on October 16,2023 at 10:59:23 UTC from IEEE Xplore. Restrictions apply.

TABLE VI
RECOMMENDATIONS

Id	Ref	General recommendations	Network monitoring tool recommendations
i (1, 2, 3)	Help filling in fields	Use the form filling feature where possible [33].	Provide guides to assist the user in filling out fields to be applied when the set of options to be chosen is large, such as the names of protocols, types of configurations, lists of routers, and lists of interfaces.
ii (4, 5)	Perception of updating	Show the transition of information between states through tracks on the screen [34].	Display important events to the administrator on the monitoring tool screen in almost real time.
iii (6)	Finding specific information in a large set of data	Use mechanisms to accelerate the search for data, because they represent a very useful means of finding items in a large set of information [35].	Adopt acceleration mechanisms for finding information in big tables resulting from data collection from the network.
iv (7, 8)	Obtaining more detailed information	Use historical records to generate complex graphs. Utilize labels and tips that appear when the mouse passes over the graphic to allow users to obtain more information about the network processes [35]	Use the hovering mouse interaction method to reveal details about the protocols, IP addresses, etc., associated with links, routes, or routers.
v (9)	Sorting information	Sort data shown in tables to help the user find the highest and lowest values in a given set of records [36].	Sort TCP/IP flows based on throughput, type of protocol, packets dropped, memory usage, etc.
vi (10)	Spying before going deeper	Summarize information about the environment as a whole to help the user find important information faster [36].	Guide the administrator to the right path for solving a given problem through the summary of records and graphs obtained from historical data of the network topology.
vii (11)	Starting point	Show an overview of the environment when logging to help the user obtain an overview of what is occurring in the network [37].	Use dashboards in the first interface screen of the monitoring tool to show general graphs, tables, alarms, and notifications regarding the network's state. Present information related to congestion points, overloaded services, and service and/or device failures also in the first interface.
viii (8)	Gradual display of information	Adopt small windows to show less information [36], [35].	Instead of showing a large amount of data on the screen, present a subset of the information in a smaller floating window that can be moved around the screen.
ix (12)	Suitably arranged data	Adapt the visualization to different screen sizes and formats.	Implement tools that are aware of the screen size so that adaptation can occur. They should consider the management of dynamic data, and, depending on the frequency of data monitoring, the information should be summarized before being shown in a graph or table.
x (13)	Perception of colors	Use colors to help highlight the more important information in the user interface. Adopt colors to add vocabulary to the dialog of users and systems [34]. Give special attention to visual impairment issues when using colors to transmit a message to users [38].	Use red and yellow to provide a better understanding because they represent alerts and warnings, respectively. Use different shapes to assist color-blind people in identifying different warnings and alerts.
xi (14, 15)	Use of metaphors to inform about status and incidents	Use well-known symbols to communicate important information [39], [34].	Use symbols with meanings such as ones that represent normal states, warnings, and failures in network monitoring tools. Coupled with these, colors such as red, yellow, or green facilitate fast comprehension of what is happening in the network.
xii (16)	Notifications	Inform the user about a situation that deserves attention. Notifications are useful for preventing further problems [34].	Avoid the use of pop-ups for notification. Instead, use a smooth non-intrusive notification mechanism at the corner of the screen or at the top of a menu.

In recent years, many guidelines regarding usability for specific domains have been developed as a means of complementing generic guidelines [32]. According to the authors of [32], general guidelines do not cover the particulars of a given domain, which demands greater effort on the part of the developers and designers to remember all the critical usability aspects. Hence, the contribution of this paper is that we focus on issues relevant to improving the usability aspects of network monitoring tools, and consequently, aiding the experts in carrying out their network management activities.

To formulate the recommendations shown in Table VI, first we analyzed the findings, seeking similarities in relation to their effect on the users' interactions when using this type of tool. The recommendations introduce a referenced association with the findings. However, as some of the findings were combined under one recommendation, we renamed the recommendations. Findings 17 and 18 were not mapped to specific recommendations as these affect the overall experience of users. In the first column of Table VI (id), inside the parentheses we include the findings that generated the recommendation.

As our final and most important contribution, we present the recommendations in Table VI, considering the domain of network monitoring tools. We grouped the recommendations according to their similar roles as follow described: help in filling information (i.e., recommendation i), support to information filtering (i.e., ii, iii, iv), information arrangement (i.e., v, vi, vii), fitting visualization to different screens (i.e., viii, ix), and information updating (i.e., x, xi, xii). The table presents the details of the recommendations divided into four columns: (i) recommendation identification, i.e., id, (ii) recommendation short reference, (iii) explanation of effect on usability in general, and (iv) effect on network monitoring tools. We include references that report the recommendation effects on usability in general.

A. Threats to Validity

To reduce/eliminate possible threats to the study, strategies were adopted considering four levels of treatment: internal (i), external (ii), construction (iii), and conclusion (iv) [10].

(i) Internal threats are the measures adopted to avoid issues that affect the performance of the participants in the activity. The tasks of the study were distributed in two sequences, "Tasks I to IV" and "Tasks IV to I," so that the order of execution of the tasks did not generate any influence by making a subsequent task easier because of the experience the participant gained from the previous task. None of the participants had previously used

Authorized licensed use limited to: Indian Institute of Space Science And Technology. Downloaded on October 16,2023 at 10:59:23 UTC from IEEE Xplore. Restrictions apply.

the monitoring tool, i.e., NagiosXI. We invited all the network administrators who had participated in Study (i); coincidentally, the nine individuals who agreed to take part in the study were not Nagios' users. This is not a problem, because, from the HCI point of view, it is always necessary to use a warm-up session, regardless of the experience of the user. However, the manner in which the warm-up session is conducted does change: the session is tuned according to the user's level of experience. In our case, it was prepared for non Nagios users. Thus, we could smooth the knowledge differences among the participants, thereby preventing any bias in the study.

(ii) Although the network administrators were all residents of Brazil, the network monitoring tool that was chosen (Nagios) is used and well-known worldwide, and therefore, the validity of the study is not threatened by the nationality of the participants.

(iii) All the tasks used in the study were designed based on the survey responses. We built a true-to-life environment, emulating a network with VMs running real-world services. All the participants faced the same conditions and adverse situations.

(iv) All the participant actions were recorded (in audio and video) for careful subsequent analysis. In the case of the video recordings, we accurately collected task execution times and reports and made notes on all occurrences during each session. Our conclusions were based exclusively on the data obtained from the study.

VIII. Final Considerations

This study revealed several important aspects of the monitoring tool, which was evaluated in terms of usability and features related to the perception of the users. According to the results, we classified the perceptions in the findings, which gave rise to the recommendations for developing better network monitoring tools.

To the best of our knowledge, this is the only study in the network monitoring area that has incorporated such a full analysis that started from the collection of survey data and continued to the identification of the personas and the performance of a study with end-users. Finally, we formulated 12 recommendations for producing better network management tools. They are not final, but rather a starting point for taking into account simple rules for building network monitoring tools that are friendlier and more oriented to fast decision making. They also can support the formulation of other guidelines for the development of better analysis tools.

As future works, we include the implementation of a network monitoring tool prototype through which we can evaluate and verify the effectiveness of our recommendations. Besides that, although this type of research, based on the use of a single tool, is well accepted by the HCI community, we suggest an in-depth study on the relation between monitoring tasks and tools used substantiated by numbers.

Acknowledgment

The authors thank all the respondents who participated in their study.

References

[1] M. Swany and P. Calyam, "Second NSF workshop on perfSONAR-based multi-domain network performance measurement and monitoring," NSF, Arlington, VA, USA, Rep. pSW 2014, Feb. 2014. [Online]. Available: https://psw2014.wordpress.com/

[2] V. Bajpai and J. Schönwälder, "A survey on Internet performance measurement platforms and related standardization efforts," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1313–1341, 3rd Quart., 2015.

[3] L. Falschlunger, O. Lehner, and H. Treiblmaier, "InfoVis: The impact of information overload on decision making outcome in high complexity settings," in *Proc. SIGHCI*, Dublin, Ireland, 2016, pp. 1–5.

[4] M. Ward, G. Grinstein, and D. Keim, *Interactive Data Visualization: Foundations, Techniques, and Applications*. Natick, MA, USA: A K Peters, Ltd., 2010.

[5] *Ergonomics of Human System Interaction-Part 210: Human-Centred Design for Interactive Systems*, Standard ISO 9241-210, 2010.

[6] J. Sauro and J. Lewis, *Quantifying the User Experience: Practical Statistics for User Research*, 2nd ed. San Diego, CA, USA: Elsevier Sci., 2016.

[7] V. T. Guimarães, C. M. D. S. Freitas, R. Sadre, L. M. R. Tarouco, and L. Z. Granville, "A survey on information visualization for network and service management," *IEEE Commun. Surveys Tuts.* vol. 18, no. 1, pp. 285–323, 1st Quart., 2016.

[8] J. Pruitt and T. Adlin, *The Persona Lifecycle: Keeping People in Mind Throughout Product Design*. Boston, MA, USA: Morgan Kaufmann, 2006.

[9] Pandora FMS Monitoring Blog. *Zabbix vs Nagios vs Pandora FMS: An in Depth Comparison*. Accessed: Jan. 17, 2016. [Online]. Available: https://pandorafms.com/blog/zabbix-vs-nagios-vs-pandorafms-an-in-depth-comparison/

[10] J. Lazar, J. H. Feng, and H. Hochheiser, *Research Methods in Human Computer Interaction*, 2nd ed. Cambridge, MA, USA: Morgan Kaufmann, 2017.

[11] T. Chown. (2016). *Sig-PMV Survey Overview*. [Online]. Available: https://wiki.geant.org/download/attachments/59933695/sig-pmv-survey-overview.pdf

[12] P. H. Isolani, J. A. Wickboldt, C. B. Both, J. Rochol, and L. Z. Granville, "Interactive monitoring, visualization, and configuration of openflow-based sdn," in *Proc. IFIP/IEEE Int. Symp. Integr. Netw. Manag. (IM)*, Ottawa, ON, Canada, 2015, pp. 207–215.

[13] M. C. Pretorius, A. P. Calitz, and D. van Greunen, "The added value of eye tracking in the usability evaluation of a network management tool," in *Proc. Annu. Res. Conf. South Afr. Inst. Comput. Sci. Inf. Technol. IT Res. Develop. Countries*, 2005, pp. 1–10.

[14] A. Voronkov, L. H. Iwaya, L. A. Martucci, and S. Lindskog, "Systematic literature review on usability of firewall configuration," *ACM Comput. Surveys*, vol. 50, no. 6, pp. 1–35, 2018.

[15] F. Guo, S. Shamdasani, and B. Randall, "Creating effective personas for product design: Insights from a case study," in *Proc. Int. Design Global Develop.*, 2011, pp. 37–46.

[16] J. Billestrup, J. Stage, L. Nielsen, and K. Hansen, *Persona Usage in Software Development: Advantages and Obstacles*, IARIA XPS Press, Denmark, SC, USA, 2014.

[17] H. Tibagí, L. Zaina, L. Sampaio, and F. L. Verdi, "Estudo sobre características de administradores de redes de computadores no brasil para identificação e elaboração de personas," in *Proc. Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC)*, Belém, 2017.

[18] A. A. Masiero, R. de Carvalho Destro, O. A. Curioni, and P. T. A. Junior, *Automa-Persona: A Process to Extract Knowledge Automatic for Improving Personas*. Heidelberg, Germany: Springer, 2013, pp. 61–64.

[19] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A *K*-means clustering algorithm," *J. Roy. Stat. Soc. C* , vol. 28, no. 1, pp. 100–108, 1979.

[20] J. Gothelf and J. Seiden, *Lean UX: Applying Lean Principles to Improve User Experience*. Sebastopol, CA, USA: O'Reilly Media, 2013.

[21] *Gns3—The Software That Empowers Network Professionals*. Accessed: Dec. 13, 2017. [Online]. Available: https://www.gns3.com/

[22] *Nagios XI—Easy Network, Server Monitoring and Alerting*. Accessed: Dec. 13, 2017. [Online]. Available: https://www.nagios.com/products/nagios-xi/

[23] *Nagios Network Analyzer. Netflow Analysis and Monitoring*. Accessed: Dec. 13, 2017. [Online]. Available: https://www.nagios.com/products/nagios-network-analyzer/

[24] *Desktop Browser Market Share Worldwide*. Accessed: Dec. 13, 2017. [Online]. Available: https://gs.statcounter.com/browser-market-share/desktop/worldwide

Authorized licensed use limited to: Indian Institute of Space Science And Technology. Downloaded on October 16,2023 at 10:59:23 UTC from IEEE Xplore. Restrictions apply.

[25] *Desktop Screen Resolution Stats Worldwide*. Accessed: Dec. 13, 2017. [Online]. Available: https://gs.statcounter.com/screen-resolution-stats
[26] P. Kortum and S. C. Peres, "The relationship between system effectiveness and subjective usability scores using the system usability scale," *Int. J. Human Comput. Interaction*, vol. 30, no. 7, pp. 575–584, 2014.
[27] D. Graziotin, X. Wang, and P. Abrahamsson, "The affect of software developers: Common misconceptions and measurements," in *Proc. IEEE/ACM 8th Int. Workshop Cooperative Human Aspects Softw. Eng. (CHASE '15)*, Florence, Italy, 2015, pp. 123–124.
[28] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *J. Behav. Therapy Exp. Psy.*, vol. 25, no. 1, pp. 49–59, 1994.
[29] R. Macefield, "How to specify the participant group size for usability studies: A practitioner's guide," *J. Usability Stud.* vol. 5, no. 1, pp. 34–45, 2009.
[30] J. Nielsen, "Usability 101: Introduction to usability," *J. Usability Stud.*, Jan. 2012. [Online]. Available: https://www.nngroup.com/articles/usability-101-introduction-to-usability/
[31] J. Nielsen and H. Loranger, *Prioritizing Web Usability*. Berkeley, CA, USA: New Riders, 2006.
[32] Q. Daniela and R. Cristian, "How to develop usability heuristics: A systematic literature review," *Comput. Stand. Interfaces*, vol. 53, pp. 89–122, Aug. 2017.
[33] P. Laubheimer. (2015). *Preventing User Errors: Avoiding Unconscious Slips*. [Online]. Available: https://www.nngroup.com/articles/slips
[34] A. Harley. (2018). *Visibility of System Status*. [Online]. Available: https://www.nngroup.com/articles/visibility-system-status
[35] T. Munzner, *Visualization Analysis and Design*. Boca Raton, FL, USA: CRC Press, 2014.
[36] L. Rosenfeld, P. Morville, and J. Arango, *Information Architecture: For the Web and Beyond*. Beijing, China: O'Reilly Media, 2015.
[37] M. Elias and A. Bezerianos, "Exploration views: Understanding dashboard creation and customization for visualization novices," in *Human–Computer Interaction* (Lecture Notes in Computer Science), vol. 6949, P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, and M. Winckler, Eds. Heidelberg, Germay: Springer, 2011, pp. 274–291.
[38] *Use of Color: Understanding SC 1.4.1*. (2016). [Online]. Available: https://www.w3.org/TR/UNDERSTANDING-WCAG20/visual-audio-contrast-without-color.html
[39] A. Kaley. (2018). *Match Between the System and the Real World: The 2nd Usability Heuristic Explained*. [Online]. Available: https://www.nngroup.com/articles/match-system-real-world

Hélio Tibagí de Oliveira received the master's degree in computer science from the Federal University of São Carlos working with information visualization for network monitoring tools. He is currently a Digital Developer with Atento Brazil. He has wide experience with the development of business applications, using mainly Microsoft technologies.

Leobino N. Sampaio is an Associate Professor of computer science with the Federal University of Bahia. His current research interests include future Internet architectures, software-defined networks, information-centric networks, and network performance evaluation.

Fábio Luciano Verdi is an Associate Professor with the Computer Science Department, Federal University of São Carlos (Sorocaba campus). He has been working with data centers, cloud computing, SDN, and data plane programmability.

Luciana A. M. Zaina is an Associate Professor with the Computing Department, Federal University of São Carlos, Brazil. Her research interests have been mostly in requirement engineering, agile practices, user experience (Ux), empirical software engineering, Ux in startups, and human aspects of software engineering.

Authorized licensed use limited to: Indian Institute of Space Science And Technology. Downloaded on October 16,2023 at 10:59:23 UTC from IEEE Xplore. Restrictions apply.