Analyzing the usability of the WYRED Platform with undergraduate students to improve its features

Francisco José García-Peñalvo[1] · Andrea Vázquez-Ingelmo[1] · Alicia García-Holgado[1] · Antonio M. Seoane-Pardo[1]

Abstract

The WYRED ecosystem is a technological ecosystem developed as part of WYRED (netWorked Youth Research for Empowerment in the Digital society), a European Project funded by the Horizon 2020 program. The main aim of the project is to provide a framework for research in which children and young people can express and explore their perspectives and interests concerning digital society. The WYRED ecosystem supports this framework from a technological point of view. The WYRED Platform is one of the main software components of this complex technological solution; it is focused on supporting the social dialogues that take place between children, young people and stakeholders. The ecosystem, and in particular the Platform, are already developed, but it is vital to ensure the acceptance by the final users, the children and young people mainly. This work presents the usability test carried out to evolve the Platform through the System Usability Scale. This usability test allows the identification of the weaknesses of the Platform regarding its characteristics, also allowing the corresponding improvement of the WYRED Platform, and it will serve as a reference for further usability testing.

Keywords Human interaction · Usability analysis · Digital society · Citizen science · Communication networks · Technological ecosystems

1 Introduction

As society changes, there is a need to understand how it is changing, to explore what is going on. The young people have a crucial role to play in society. They are frequently the drivers of new behaviors and understandings, and since they are part of the future of society, their views and perceptions should be taken into account. However, they are not well represented, and their voices are unheard, and this makes it hard for research and policy to identify and understand their needs.

WYRED is a European Project funded by the Horizon 2020 program (Table 1). It is coordinated by the GRIAL Research Group [1, 2] of the University of Salamanca (Spain), and it started at November 2016 and will be developed along 3 years, until October 2019.

The project aims to provide a framework for research in which children and young people can express and explore their perspectives and interests concerning digital society, but also a technological tool from which they can communicate their views to other stakeholders effectively through innovative engagement processes [3]. WYRED is informed by the recognition that young people of all ages have the right to participation and engagement. It has a strong focus on inclusion, diversity and the empowerment of the marginalized. The aim is to replace the disempowering scrutiny of conventional research processes with the empowerment of self-scrutiny and self-organization through the social dialogue and participatory research [4].

To support the framework for research from a technological point of view, a WYRED technological ecosystem has been defined [5, 6], allowing both the interaction

Table 1 WYRED project details

Title	netWorked Youth Research for Empowerment in the Digital society
Acronym	WYRED
Funding entity	European Union
Call	Horizon 2020. Europe in a changing world—inclusive, innovative and reflective Societies (HORIZON 2020: REV-INEQUAL-10-2016: Multi-stakeholder platform for enhancing youth digital opportunities)
Reference	727066
Project leader	Francisco José García-Peñalvo
Coordinator	University of Salamanca (Spain)
Partners	Oxfam Italia (Italy) PYE Global (United Kingdom) Asist Ogretim Kurumlari A.S.—Doga Schools (Turkey) Early Years—The organization for young children LBG (Ireland) Youth for exchange and understanding international (Belgium) MOVES—Zentrum für Gender und Diversität (Austria) Boundaries Observatory CIC Tel Aviv University (UK)
Budget	993.662,50€
Start date	01/11/2016
End date	31/10/2019
Web	https://wyredproject.eu

among the members of the community and managing the generated knowledge [7]. The WYRED Platform is one of the main software components of this complex technological solution; it is focused on supporting the social dialogues that take place between children, young people and stakeholders.

The ecosystem, and in particular the Platform, are already developed, but it is important to ensure the acceptance by the final users, the children and young people mainly. To achieve this goal, a usability test with real users was carried out. A group of undergraduate students between 18 and 25 years old used the WYRED Platform for 3 weeks with different roles. During the pilot, the interaction of the students was collected using Google Analytics. After the pilot, the participants filled the System Usability Scale (SUS) questionnaire anonymously.

This work describes, in the first place, the technological framework in which the WYRED Platform is implemented, with a particular focus on its software architecture, and in the second place, the usability study and the analysis of the SUS test results.

Finally, this work is set out in seven sections. The second and third sections provide a brief description of the WYRED ecosystem and the WYRED Platform. The fourth section outlines the research methodology; it explains the materials and methods used to evaluate the WYRED Platform, the evaluation procedures and the involved participants. The fifth section presents the results obtained after the analysis of the responses. The sixth section is the discussion, and the final section concludes the work with its more significant contributions.

2 Technological framework

The increase in concern for successful knowledge management in any kind of organization or institution is part of the current Knowledge Society. The adequate management of knowledge [7, 8] in a different context has led to the evolution of traditional information systems in what are nowadays called technological or software ecosystems. Information systems emerged in the Information Society to cover the information management needs, whereas technological ecosystems are focused on supporting the knowledge management typical of Knowledge Society.

These technological solutions use the metaphor from nature to represent an intrinsic property of the knowledge, the evolution. One of the main characteristics of a technological ecosystem is its ability to evolve in different dimensions [9, 10]. Several authors use the definition of natural ecosystem to provide a definition of technological ecosystem [11–18]. In particular, García-Holgado and García-Peñalvo translate the main elements in a natural ecosystem to the technological point of view [19–22]: the biotic factors are the users and the software components; the relationships are the information flows that establish the communication between them; and the abiotic factors are the physical environment—hardware, network, etc., that provide support to those flows. Thereby, a technological ecosystem is a set of users and software components related to each other through information flows in a physical environment that provides the support for those flows.

In addition to the strong evolutionary component, there is another important difference between a technological

ecosystem and an information system. People are part of the ecosystem in the same way that the software components, people are not only users of the technology, are part of the technology [23, 24].

The WYRED ecosystem is based on this natural ecosystem approach. It is composed of a set of open source tools and the people involved in the project, not only the partners, but also the stakeholders, children and young people. Figure 1 shows the architecture of the ecosystem, the connections among the software components and the human factor. This architecture is not a final solution, like other ecosystems, the evolution is part of the solution, the WYRED ecosystem will evolve over time to cover the new requirements provided by the project partners and solve the problems detected during the usability studies, for instance, the results of the study described in this work.

The architecture of the WYRED ecosystem has been defined using the architectural pattern for technological ecosystems defined by García-Holgado [9, 20, 25]. The WYRED ecosystem is organized in four layers—presentation, services, static management data and infrastructure—and two input streams which introduce the human factor as another element of the technological ecosystem.

The top layer, presentation, is focused on providing a uniform interface to all the components of the ecosystem. The WYRED ecosystem has an identity that is applied to the whole ecosystem. Moreover, this layer also ensures the ecosystem accessibility from every kind of device.

The services layer provides the software components with the main user-level services and the connection with the social networks to give visibility to the knowledge generated inside the ecosystem. In the WYRED ecosystem, the leading software component is the WYRED Platform. It is connected to other tools to support the social dialogues carried out by young people.

The static data management layer does not represent a common layer because it does not separate the services and the infrastructure completely [19]. In the WYRED ecosystem, the objective of the social dialogues is to generate individual and collaborative research projects. Text files, presentations and multimedia material, among other types of documents, are used to develop the researches. The main software component in this layer is a platform to share documents and provide simple version control for files.

The last layer, infrastructure, provides all necessary services at an internal level to make the system work properly. In particular, the mail server, the user management tool based on CAS (Central Authentication Service) and the indexing service to improve the search tool in other software components.

Finally, the two input streams represent the methodology and the management for supporting the methodology and the ecosystem evolution. The methodology is provided by the WYRED Consortium composed by the different institutions and entities involved in the project, and the management is carried out by the project coordinator.

Currently, some components of the architecture are in the development phase, in particular, the document management system based on OwnCloud and the indexing service based on ApacheSolr.

Fig. 1 Architecture of the WYRED ecosystem
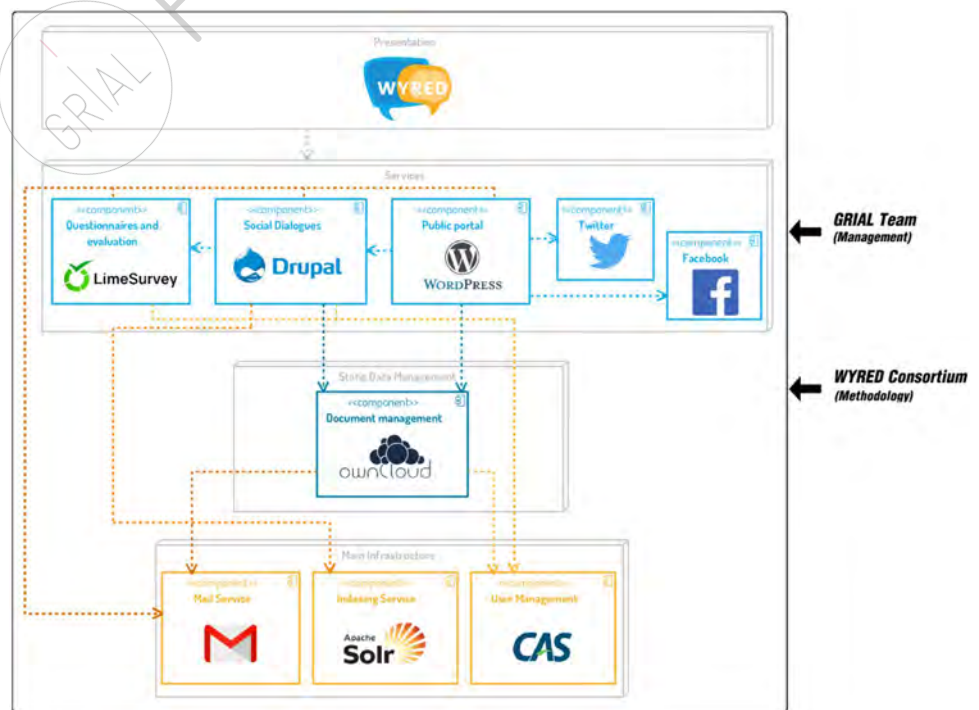
3 The WYRED Platform

The WYRED Platform (https://platform.wyredproject.eu) is one of the primary services included in the WYRED ecosystem. A principal aim of this project is to engage children and young people in the process of social dialogue giving them a voice to share their thoughts, fears and feelings about digital society. The WYRED Platform aims to provide technical support to these social dialogues, and its design is centered around and driven by children and young people.

The WYRED Platform allows a natural interaction process among the involved stakeholders, high-level engagement conditions, a secure environment in which all the participants feel comfortable enough with the privacy issues, but with particular attention to the young ones, and a suitable dashboard for data analytics [26].

A Content Management System (CMS) provides a solid base to build online tools focused on managing information and supporting the interaction among the users. Nowadays, there are three CMSs that have been established themselves as the leaders: Joomla, Drupal and WordPress. The WYRED Platform is based on Drupal (release 7.x) because it is most technically advanced [27], with a powerful development framework, and the most recommended for community platform sites with multiple users [28] due to its own experience providing a space for working groups (https://groups.drupal.org). Although Drupal released version 8 on November 2015, not all modules that extend its core functionality are migrated to the new version. In particular, the base module to develop the WYRED communities (Organic Groups module) has a development version for Drupal 8 but not a stable version that can be used in a production site.

Regarding the functionality of the Platform, it is organized in multicultural and interdisciplinary communities where young people can develop research projects with the support of facilitators from different European institutions and associations. The communities have different tools such as forums to establish dialogues and coordinate research cycles, calendars to share dates and organize events or activities, surveys to develop or evaluate the projects. Also, it provides a tool to publish the results of the research projects. The users with a facilitator role inside a community are their administrators; they can manage user's roles and subscriptions, moderate the forum threads, create projects and invite new users to the community both registered and non-registered users. Moreover, there are two types of communities: public communities that are accessible for all registered users and anyone can subscribe; and private communities only visible to their members.

One of its main innovations is the strong commitment to user privacy; it is designed as a safe space in which children and young people can be free to express themselves as they wish. The Platform is only accessible to registered users. The registration process is restricted too; new users must receive an invitation with a unique link to have access to the registration form. The only contents that anonymous users can visit are the help section with videos and training actions and the terms of use section. Just like the communities, users can have three different roles at the Platform level, one for regular users and two for users with privileges—facilitator and administrator. A user with administrator access is the only one that can change the role of other users.

Finally, all registered users can answer a socio-demographic questionnaire about diversity and inclusion to get an overall perspective of the children and young people involved in the project. One of the most important things in the WYRED Platform is to maximize the users' privacy. For this reason, the user's private information collected during the registration process and with the questionnaire is saved in another software component of the WYRED ecosystem, the user management tool based on CAS.

A complete report with information about the last version of the WYRED Platform is available on [29].

4 Methodology

4.1 Participants

The study was carried out in the course of "Communication Techniques and Skills," a mandatory subject for undergraduate students, taught in the first semester of the first year of the Degree in Social Education of the University of Salamanca (Spain). There were 80 students enrolled in the subject in the 2017–18 school year, and the experimental group was formed by 77 students divided into 12 workgroups. Only 14.29% of the students are men, and 85.71% are women. These figures are due to the fact that in Spain, the percentage of female tertiary students that choose studies in the field of social sciences is 60% [30]. Regarding the country of birth, 68 students were from Spain, and 9 foreign (2 Dutch, 1 Argentinian, 1 Belgian, 1 Brazilian, 1 Chinese, 1 Italian, 1 Jamaican and 1 Portuguese); all of them speak Spanish. Table 2 shows the distribution of participants by the role in the WYRED Platform, gender and average age.

Table 2 Participants in the usability test ($n=77$)

Role	Female	Male	Average age
Facilitators	11	1	20.66
Students	55	10	20.02
Total	66	11	20.12

4.2 Instrumentation

The tool selected to measure the platform usability was the System Usability Scale (SUS). SUS is essentially a questionnaire that consists of ten simple items that the users of the system to evaluate will score.

This questionnaire is classified as a post-test questionnaire, meaning that it is applied after the users end a session with the system to be evaluated. Post-test or post-study questionnaires provide quantitative or qualitative data about the overall perceived usability of a system, contrasting with post-task tests, which are executed right after a task is completed, allowing the reaching of insights about particular or individual assignments. Post-test and post-task questionnaires can be combined to obtain a broader view of the perceived usability of the tested system, though for the WYRED Platform usability testing, a post-test questionnaire has been selected to make the usability measuring process simple and to obtain insights quickly to improve the platform.

In this case, as already mentioned, the perceived usability of the platform has been evaluated applying the SUS questionnaire. Although there were other options available to conduct a post-test questionnaire (the Questionnaire for User Interaction Satisfaction (QUIS) [31], the Software Usability Measurement Inventory (SUMI) [32] or the Post-Study System Usability Questionnaire (PSSUQ) [33]), SUS has become one of the most popular post-study standardized questionnaires [34]. This questionnaire is also one of the fastest to converge on the correct conclusion, meaning that SUS is a good option if the sample size is limited or it is suspected that it could be challenging to have a significant sample size of testers [35]. In addition to the last, the main reason of applying this kind of test is based on its own usability characteristics [36, 37]:

- It is an effective method: it provides a high-level measurement of subjective usability, being a valid and reliable tool [35];
- It is efficient: the SUS questionnaire is an efficient test, given the fact that it only consists of ten items. Filling out a long questionnaire can lead to frustration after prolonged sessions of testing, and therefore, the results could be skewed, providing inaccurate data. Because the WYRED Platform usability testing lasted for 3 weeks as will be described in the following section, a short questionnaire at the end of the testing is a good option to retrieve the opinion of the participants.

The SUS questionnaire can be applicable over a wide range of systems [38], including online platforms, making it suitable for the WYRED Platform. It is important to keep in mind, however, that the score obtained is the perceived usability of the users, so these scores are individual and every user could have a subjective usability score of the system tested based on its opinion.

The questionnaire items are positive and negative alternated statements (in order to avoid response biases) rated on a 1–5 Likert scale (from "strongly disagree" to "strongly agree," respectively) [39]. This Likert scale ranging from 1 to 5 (i.e., an odd number of steps) makes possible to indicate a "neutral" attitude when completing the usability questionnaire (the user can put his or her score at the middle of the scale to keep a neutral opinion) [34].

The validity and reliability provided by this test and the ease of its implementation made the System Usability Scale the selected usability measurement method for the WYRED Platform testing. To sum up, it is a quick test which can provide valuable insights (although not diagnostic) about the perceived usability of the platform.

For the WYRED Platform usability test, in addition to the SUS questionnaire predefined items, a set of demographic variables, as well as platform-specific variables were also collected by the instrument:

- Year of birth;
- Gender;
- Birthplace;
- Parents' birthplace;
- Family's main spoken language;
- Any eye diseases of the user that could affect the experience;
- Language used in the platform;
- Role of the user (student or facilitator).

Besides the aforementioned demographic variables, an open field was provided at the end of the survey to let the users express themselves and discuss any relevant experience they had during the testing of the platform. This field allows the usability test to count on qualitative feedback, in addition to the SUS score. (As the SUS score itself is not diagnostic, it only provides an overall usability rating). The narrow set of items that compose by design, the SUS questionnaire allows the recollection of these additional variables without creating a large form, which is beneficial because lengthy questionnaires can lead to frustration after a long session of testing. The main reason for the recollection of the previous variables is because of the aim of the WYRED Platform to provide technical support to social dialogues, so the gathering of socio-demographic variables could be valuable. Also, the analysis of the SUS scores based on the socio-demographic factors could help to reach more diagnostic insights about the usability of the WYRED Platform depending on the users' characteristics.

The instrument to collect this data was implemented using one of the components of the WYRED ecosystem, the tool to create questionnaires and surveys, a customized version of

LimeSurvey (https://www.limesurvey.org), an Open Source online statistical survey web application. The instrument was applied in Spanish because the context required it (the users participating in the usability test were Spanish speakers), making the translation of the items a vital task [40], although it is also available in English to be used for future usability studies in other partner countries.

4.3 Study design and data collection

The pilot experience was carried in the last weeks of face-to-face classes of the "Communication Techniques and Skills" course of the Degree in Social Education of the University of Salamanca (Spain) during the 2017–2018 school year. One of the main goals of the subject is to address conflict resolution.

During the pilot, the students were divided into 12 work groups to promote a series of social dialogues on topics that they considered of their interest and fit into the topics of the WYRED project.

The selection and configuration process of the social dialogues started with the presentation of the proposed topics in a collaborative board using the free version of Padlet (https://padlet.com), a multi-device application to make and share content with others. Then, the topics were explained in a face-to-face class with the purpose of guarantee the broadest range of possible topics. Each group selected a different topic and a group member as a facilitator.

The aim of the activity was argued about the selected topic using the WYRED Platform to prepare a report in which the following aspects would be reflected:

- Definition and description of the selected topic;
- To what extend the selected topic concern to young people;
- Proposals to solve or improve the state of the problem addressed;
- User experience and proposals to improve the WYRED Platform.

The topics selected by the students to work in the Platform were:

- Gender stereotypes and discrimination;
- Cyberbullying, humiliation and sexting;
- Construction and knowledge of ourselves through education and new technologies;
- Privacy on the Internet;
- Stress reasons among young people;
- Young people's access to Deep Web;
- The new influencers in social networks and the false myths;
- Personality in the context of social networks;

- Gender and the digital society;
- Sexting in relation to gender violence and cyberbullying;
- Machism in the social networks;
- Dangers of the Internet for young people.

After the initial phase in the classroom, the administrator of the WYRED Platform sent a registration invitation to the selected facilitators (12 students) with the data provided by the teacher. When each one of them finished the registration process, the administrator gave him/her a facilitator role in the Platform because the role cannot be assigned before the user exists in the Platform.

Each facilitator had two different tasks regarding their group mates. First, he/she had to create a community inside the Platform to carry out the activity. Second, the facilitator had to send a registration invitation to each of his/her group mates so that they had to register and start to work in the community.

The pilot experience lasted for 3 weeks: 1 week to organize the activity and 2 weeks to use the WYRED Platform to prepare the report.

Regarding the data collection, the teacher sent by email the link to answer the questionnaire 2 days before the pilot experience ended, and the questionnaire was available for 2 weeks. During that period, two reminders were sent by the teacher to request more answers.

Moreover, the interaction of the students was collected using Google Analytics (https://analytics.google.com), a web analytics service offered by Google to track website traffic. The WYRED Platform uses this tool principally to get information about the interactions of the different countries involved in the social dialogues and the employed devices to work with the Platform. This information is complemented with the information extracted from the WYRED Platform database to get the number of forum threads and comments.

4.4 Analysis

The instrument collected the data; however, to perform the analysis of the answers with external libraries, all the responses were downloaded in a structured format (CSV). The dataset obtained was cleaned to avoid useless data and, subsequently, to calculate the SUS score. Although the score calculation is relatively straightforward [39], the analysis of the responses has been made through the Python Pandas [41] library given its high performance and easy-to-use data structures. This library, among its other features, also allows filtering datasets by field, which can be useful to establish relations between the individual SUS score results and the other variables collected from each user.

Taking this into account, the individual SUS score was calculated for every participant. Once this task was completed, the average of every individual SUS score was

computed to obtain the final perceived usability of the WYRED Platform.

However, it was necessary to introduce another analysis taking into account that a technical error was found in the system during the usability testing. The average SUS score was also calculated for both the group of users exposed to the failure and the group of users that tested the platform after the error was fixed.

The analysis of the SUS also includes the aggregation of the scores based on the socio-demographic variables collected, in order to look for any differences in the perceived usability of the users given its own characteristics, as well as for a deeper analysis of the results, including the calculation of the confidence interval of the SUS scores obtained and the individual analysis of the SUS items responses collected.

The interpretation of the results is based on previous System Usability Scale studies and benchmarks [42, 43]. These studies aimed to establish average SUS scores depending on the type of system tested and to throw light about what an acceptable SUS score is, so they make possible the reach of insights about the results obtained and the perceived usability of the WYRED system.

All the source code developed for this analysis is available at https://github.com/AndVazquez/wyred-usability-analysis [44].

5 Results

5.1 Socio-demographic results

From the 70 participants that entered the questionnaire, 43 of them finalized it, with the following socio-demographic characteristics:

- 86.05% of students are female and 13.95% male;

- 41 students were born in Spain, 1 in China, 1 in Belgium and 1 in the Netherlands;
- 22 students were between 19 and 20 years (51.16%), 20 students were between 20 and 24 years (46.51%), 1 student was between 25 and 29 (2.33%). The average age is 21 years, approximately;
- 33 participants used the Platform in Spanish (76.75%), eight participants in English (18.60%) and two participants both in Spanish and English (4.65%);
- No users pointed out significant sight diseases that could affect their experience with the platform. However, 13 users indicated having some kind of sight issues:

    - Six users were afflicted by degenerative myopia
    - Two users were afflicted by astigmatism
    - Two users were afflicted both by myopia and astigmatism
    - Two users were afflicted both by astigmatism and hyperopia
    - One user was afflicted by hyperopia.

These socio-demographic results could help to describe and understand the sample of this usability test and analyze the results based on the users' characteristics.

5.2 WYRED Platform interaction during the testing

As was explained in the study design and data collection section, the users were told to execute a series of tasks to carry out the activity, including registering in the platform, creating communities and sending invitations (in case the user was a facilitator) and discussion of the topics selected within the WYRED Platform. This interaction of the students was recorded with Google Analytics, yielding a significant number of active sessions during the usability testing. As shown in Fig. 2, 88.24% of the sessions of the WYRED

Fig. 2 Active sessions and regions where these sessions belong while the WYRED Platform usability testing was taking place
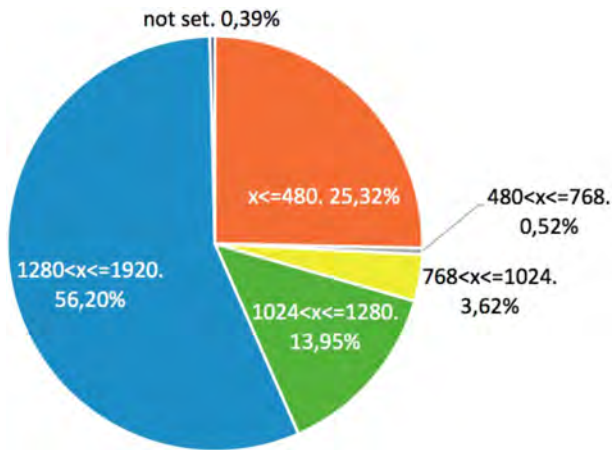
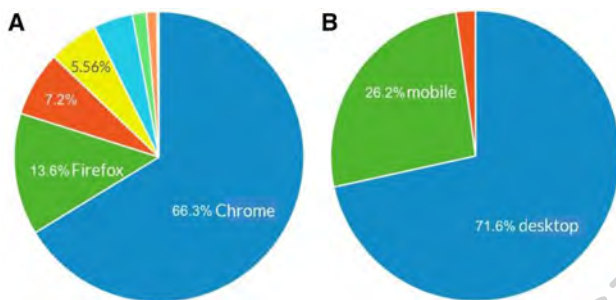**Fig. 3** Browser sizes used to access the WYRED Platform during the usability testing

Fig. 4 Browsers (**a**) and devices (**b**) used to access the WYRED Platform during the usability testing

Platform were originated in Castile and León (the Autonomous Region of Spain to which the University of Salamanca belongs), concluding that the majority of these sessions were caused because of the pilot experience. These data verify that the participants of the testing interacted significantly with the platform to be able to score its usability.

On the other hand, the analysis of the devices used to access the platform shows that 71.6% of the users used desktop devices to enter the WYRED Platform, contrasting to 26.2% of users that used mobile devices and only 2.2% that use tablets. Furthermore, the browser size covers a wide range of resolutions, from a 320×440 to 1920×970 pixels. The browser sizes used were cluster in six groups (Fig. 3): lower or equal than 480 pixels (25.32%); between 480 and 768 pixels (0.52%); between 768 and 1024 pixels (3.62%); between 1024 and 1280 pixels (13.95%); between 1280 and 1920 pixels (56.2%).

Regarding the browsers, most of the users used Chrome (66.3%), followed by Firefox with 13.6%, Edge with 7.2% and Safari with 5.56%. Figure 4 shows the distribution of the devices and browsers used. This information is also valuable

as the perceived usability could depend on the device used for the access [45], and even the System Usability Scale score could be afflicted by the context [46].

5.3 WYRED usability test (SUS) results

As already mentioned, 43 users completed the entire set of questions regarding the usability of the WYRED Platform. It is important to analyze only filled questionnaires; if some user abandoned the questionnaire without having completed it, its responses should not be taken into account as it would be incomplete. According to the literature, the System Usability Scale (SUS) is reliable with a minimum sample size of 12 participants [35]. Consequently, the 43 responses received could yield reasonably reliable results, and the sample size could even let us analyze the results by aggregating the users by their socio-demographic characteristics.

Although SUS was initially developed to provide a single score that indicates the (perceived) usability of a system [37], subsequent studies pointed out the two-dimensional nature of this scale [47], allowing the calculation of the system's learnability score (in addition to the usability score). Mainly, from the ten items that conform the SUS questionnaire, items 4 and 10 can be used to score the learnability of the system being tested, while the remaining items are used to obtain its perceived usability [47]. However, there are some caveats and a necessity to be cautious when using these two sub-scales of the SUS questionnaire, as there have been contradictory findings when analyzing the dimensions of this questionnaire [48].

Taking this into account, the SUS score was calculated following the scoring instructions [39] for every participant's responses. Additionally, the learnability score (from items 4 and 10) and usability score (from items 1, 2, 3, 5, 6, 7, 8 and 9) were also calculated and transformed to fit in a scale from 0 to 100 (as in the original SUS scoring method, in order to allow comparisons), so the results from the learnability items were multiplied by 12.5 and the results from the usability items were multiplied by 3.125 [34].

The calculations yielded the following results [49]:

- The average perceived usability of the WYRED Platform is 65.23, which can be considered as an acceptable SUS score, as it is close to the average SUS score (68.00) and falls around the 50th percentile (interpretation based on the studies done in [42, 43]);
- On the other hand, the perceived learnability seems to be slightly higher (66.28) than the usability (64.97), both being acceptable scores.

Figure 5 summarizes the outcomes of the SUS questionnaire, also including the individual scores for every participant (represented by overlapping circles) across the
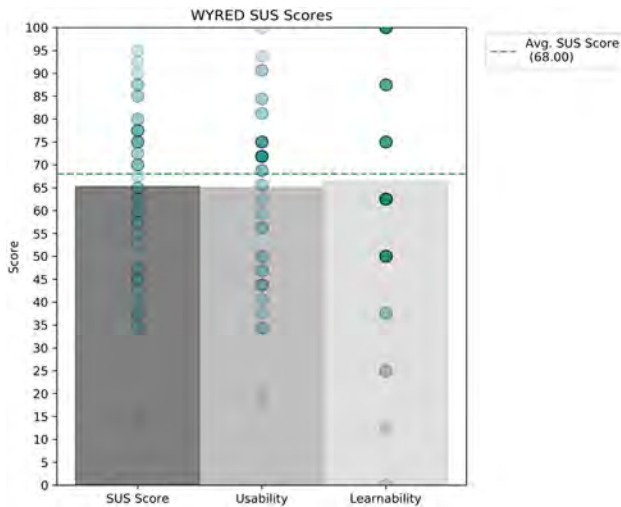
Fig. 5 Visual representation of the SUS questionnaire results regarding the WYRED Platform usability and learnability scores ($n = 43$)

three dimensions considered: SUS score, Usability and Learnability.

To reach more insights about this score, a more in-depth analysis has been carried out. The mean of all the System Usability Scale scores collected was computed for the target sample, but it is also relevant to know how precise these estimates are, as it is impossible to have access to the whole user population.

One of the most used methods to estimate the population parameters is to specify an interval of values (confidence intervals [50]) in which the population's SUS score mean would fall. To compute the margin of error to build the confidence interval for the WYRED Platform SUS score mean, it is necessary to compute a series of calculations [34]:

- Standard deviation: the standard deviation for the set of responses regarding the SUS score is 18.026. As shown in Fig. 5, some outliers can be observed regarding the given SUS score, affecting the mean and the standard deviation;
- Degrees of freedom: the sample size with one unit subtracted, in this case, as the sample size consist of 43 users, the degrees of freedom are 42;
- T-critical value: for a level of significance of 0.05, the *t*-critical value is 2.02.

The margin of error yields a result of 5.55, obtaining a confidence interval (with a 95% of confidence level) of the SUS score between 59.68 and 70.78, meaning that the correct SUS score for the population would be within this range.

However, as mentioned in the previous sections, there are some considerations regarding these results: during the evaluation of the WYRED Platform, a technical error was

found. This issue was resolved over the next 8 days, meaning that some participants were exposed to that error.

Considering this situation, the users were divided into two groups (users who tested the platform *before* the fix and users who tested the platform *after* the fix). The SUS score was calculated for each group, reporting new results.

Group 1 (users who tested the platform *before* the fix): the average SUS score for the 30 participants belonging to this group is 64.67, which is lower than the general one. Learnability and Usability scores are 65.27 and 64.48, respectively. For this SUS score, the confidence interval is the following:

- Standard deviation: 17.05;
- Degrees of freedom: 29;
- T-critical value (0.05 significance): 2.04;
- Margin of error: 6.37;
- Confidence interval (95% of confidence): 58.30 to 71.03.

Group 2 (users who tested the platform *after* the fix): in this case, the 13 users from this group gave an average SUS score of 66.54 (slightly higher than the SUS score obtained from the whole set of users). For these users, Learnability and Usability scores (68.27 and 66.11, respectively) fall closer to the average SUS score, based on the literature. In this case, the confidence interval calculation for the SUS score of this group yields the following results:

- Standard deviation: 20.78;
- Degrees of freedom: 12;
- T-critical value (0.05 significance): 2.18;
- Margin of error: 12.56;
- Confidence interval (95% of confidence): 53.98 to 79.10.

Another aggregation of the SUS score by the users' characteristics has been made to test if having the role of facilitator in the WYRED Platform affects the perceived usability. This analysis could yield insights about the difficulty of the tasks faced by users with this role; Jeff Sauro led further research concluding that users facing difficult tasks could score the perceived usability an 8% lower on average, compared with users that had to complete another kind of tasks [51].

The users were aggregated by their role (facilitator of the platform and general users), and the average SUS score for each group was computed with the following results:

- Users with facilitator role: the average SUS score for the 11 facilitators that answered the questionnaire was 61.82. The Learnability and Usability in this case yield scores of 62.50 and 61.65, respectively. In this case, the confidence interval is located between 50.01 and 73.62;
- Users with general user role: users that did not have facilitator tasks scored the platform yielding a SUS

	Sample size	Mean	SD	MOE	CI (95%)	Learnability	Usability
Overall	43	65.23	18.03	5.55	59.68–70.78	66.28	64.97
Before the fix	30	64.67	17.05	6.37	58.30–71.03	65.27	64.48
After the fix	13	66.54	20.78	12.56	53.98–79.10	68.27	66.11
Facilitator	11	61.82	17.57	11.80	50.01–73.62	62.50	61.65
Not facilitator	32	66.41	18.30	6.60	59.80–73.00	67.58	66.11
Age (over 21 years)	21	61.07	20.56	9.36	51.71–70.43	62.50	60.71
Age (under 21 years)	22	69.20	14.60	6.48	62.73–75.68	69.89	69.03
Sight problems	13	59.23	16.97	10.25	48.97–69.48	58.65	59.38
No sight problems	30	67.83	18.12	6.77	61.07–74.60	69.58	67.40

score of 66.41, with a confidence interval between 50.01 and 73.62. For this score, the items regarding Learnability and Usability scored 67.50 and 66.11, respectively.

There is a 7.42% decrease between the perceived usability scores of the users with facilitator role and general users.

Finally, users were also classified by age, creating two groups (users under and over the age of 21), and by their sight problems (users that pointed out any disease and users that do not have any disease). The results for these classifications and the summary of the previous results are given in Table 3.

5.4 System Usability Scale individual items results

The individual items of the System Usability Scale have also been analyzed. As already mentioned, this questionnaire is based on Likert scale items that users have to score. Although Brooke pointed out that "scores for individual items are not meaningful on their own" [39], it is possible to gain further information about the results by analyzing the particular statements [38].To analyze every individual item, the frequencies of the responses have been calculated, given the Likert nature of the scale [52]. The complete set of results can be consulted at [44].

The item that users scored higher was the item 9: "*I felt very confident using the system,*" the mean of the responses being 3.74. The majority of the users scored this item with 4 (which translates to "Agree" in the Likert scale). The frequencies of the scores for this item are shown in Fig. 6.

On the other hand, the worst score was obtained for item 1: "*I think that I would like to use this system frequently.*" In this case, the mean score is 2.94 given the fact that the majority of the users scored the item with 3 (which translates to a neutral opinion). The frequencies of the scores for this item are shown in Fig. 7.
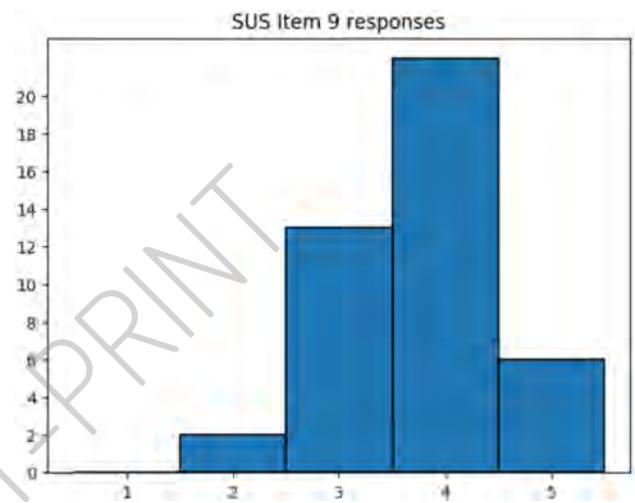

Fig. 6 Scores distribution for the ninth item of the System Usability Scale test ($n=43$)
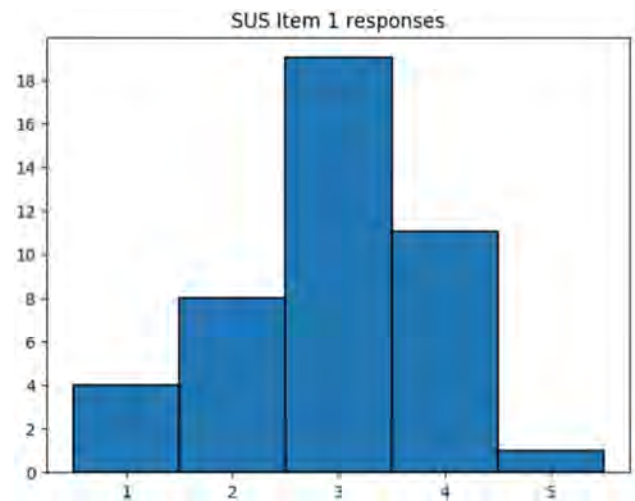

Fig. 7 Scores distribution for the first item of the System Usability Scale test ($n=43$)

6 Discussion

A total of 43 students completed the questionnaires regarding the usability of the WYRED Platform after interacting with it during a period. The sample size was enough for the purpose of this usability test to deliver reliable results, based on the literature [47]. This sample size also allows the classification of users by their socio-demographic characteristics, taking into account that the groups derived from the aggregation should also have around 12 individuals to perform a valid analysis of the scores. However, the characteristics of the sample did not allow the analysis of the usability by interesting demographic variables as gender or country of birth, due to the distribution of this values: among the 43 users of the sample, only six individuals were male, and three individuals were not from Spain leaving tiny groups (with less than 12 participants) to analyze if the classification by those categorical variables was done.

The overall average usability of the WYRED Platform for the 43 students is 65.23, with a 95% confidence interval between 59.68 and 70.78. As commented in the previous sections, a score of 65.23 is an acceptable score, although it is below the average of perceived usability in web systems (67–68) [34, 38]. Specifically, this score is in the "marginal high" in terms of the acceptability ranges, and falls in the "OK" and "Good" interval as defined in the SUS adjective ratings [42] and it is associated with the "C" grade regarding the Sauro–Lewis Grade [34]. The same interpretation applies to the individual Learnability and Usability scores.

While the results can be considered as acceptable, they also reveal that the WYRED Platform usability has a margin to be improved. The non-diagnostic nature of the SUS usability test makes the obtained score just an overall, quantitative rating for the system's perceived usability that can be compared with other systems' scores thanks to the existing benchmarks. However, the comments collected along with the survey offer qualitative information about the opinions of the participants regarding their experience with the WYRED Platform and some hints about what features to be improved and even about what factors lowered the given scores.

One of the issues pointed out in the commentaries were about the language of the Platform, for which translation is in progress. Also, some participants also mentioned in the commentaries' section the technical error found during the testing (an issue that affected the login tool). This issue was mitigated, and the users that completed the questionnaire after the fix rated the usability of the system better (with an average score of 66.54). Although the sample size of this set of users is lower than the group that tested the Platform before the fix (13 users against 30 users, respectively), the results obtained from this segment of users is equally valid, because the sample size is enough to apply the SUS test [47].

Other complaints found in the comments refer to inherent contents of the platform, like the required initial questionnaire before the registration is done. In the end, the comments collected also reflected that although some users rated the Platform low, they felt that the WYRED Platform is a practical and valuable tool.

Given the fact that a set of categorical variables were collected along with the questionnaire, a more in-depth analysis was made by dividing the participants into groups based on their characteristics to check if there were any differences in their average SUS scores. First, the role of the user was analyzed. The role of facilitator in the WYRED Platform comes with additional tasks that could be perceived as difficult. The results showed that facilitator users scored the Platform lower (with a score of 61.82) than general users (with a score of 66.41). Specifically, facilitator users scored the Platform a 7.42% lower, which could mean that the particular tasks of users with the facilitator role are seen more difficult than the ones that general users face [51].

Users were also classified by age to check if this variable affected the score, obtaining a SUS score of 61.07 for users over the age of 21 and a score of 69.20 for those users under the age of 21. This result may indicate that younger users perceived the platform more usable than other older users, though further research should be done with more significant samples to find a correlation between age ranges and the usability score given to the WYRED Platform. However, this difference between the scores may be due to the experience of younger users with this kind of systems, as the usage of social networks is more intense in younger populations. That is why further usability tests could be improved by adding an item to check the experience of the user on similar systems or platforms, as the experience could influence the SUS score (inexperienced users can lead to a decrease in the usability score [53]).

Another classification was made regarding the vision impairments pointed out by the users. Vision impairments are an essential factor to take into account as they can disturb the experience of the user with the system. The results yielded that users without any vision impairments scored the Platform better (67.83) than those users with sight diseases (59.23). Nevertheless, the sample sizes, in this case, were uneven (30 users without sight issues and 13 users with any kind of sight diseases), which made the analysis more difficult. Also, the vision impairments pointed out by the participants (myopia, hyperopia or astigmatism) are not severe impairments that could affect the user experience significantly. It would be great to count on the opinion of people with these kinds of severe sight diseases (blindness or color blindness) in further usability tests to check if the Platform is usable for them too or to check what kind of problems they face when using the Platform.

Finally, the analysis of every particular SUS item, although originally not supposed to be analyzed individually, gave information about specific usability factors. For example, the ninth average score was 3.74, which is above the average score for this item according to other benchmarks [38]. This score indicates that users feel confident using the WYRED Platform, being a valuable quality for a system in which the target group is young people.

On the other hand, other items were scored individually below the average, which is the case for the first item of the questionnaire, rated with an average score of 2.94. The mean for this item is 3.68, meaning that it could be necessary to investigate why users would not like to use the WYRED Platform frequently and act accordingly to improve these users' opinion.

To sum up, the application of the SUS test provided insights about the usability of the WYRED Platform and placed the system in the acceptable range of usability, with room for improvement based on the questionnaire feedback and the individual analysis of every SUS items' responses.

7 Conclusions

The WYRED Platform is a software component of the WYRED technological ecosystem that is focused on supporting the social dialogues about the digital society between children and young people. A first version of the Platform is already developed, but it is important to ensure the acceptance by the final users and to know their opinions to evolve the Platform and, therefore the ecosystem.

The System Usability Score has been applied to the WYRED Platform to obtain insights about its usability. This test is not diagnostic; it gives an overview of the usability of a system. The WYRED Platform obtained an average score of 65.23, which is a score below the average (68) but considered a decent result.

However, although this score does not allow the identification of usability issues, adding an open field in the questionnaire for comments offered hints about the improvements to be made to increase usability. Moreover, it provided valuable feedback about the general thoughts regarding the Platform, the majority of them being positive despite the score given.

There are some important changes which need to be made in the WYRED Platform to improve the SUS score. Future work will be focusing on solving some technical problems detected by the participants and the experts during the pilot experience, which could influence in the study results. This usability testing will be a reference to apply further tests to the platform, allowing the comparison of the results over time and showing the evolution of the platform regarding usability.

References

1. García-Peñalvo, F.J., Rodríguez-Conde, M.J., Seoane-Pardo, A.M., Conde-González, M.Á., Zangrando, V., García-Holgado, A.: GRIAL (GRupo de investigación en InterAcción y eLearning), USAL. IE Comun. Rev. Iberoam. Inform. Educ. **15**, 85–94 (2012)

2. GRIAL Group: GRIAL Research Group Scientific Production Report (2011–2017). Version 2.0. GRIAL Research Group, University of Salamanca, Salamanca (2018). https://doi.org/10.5281/zenodo.1217097

3. Rodríguez Conde, M.J., García-Holgado, A., Zangrando, A., García-Peñalvo, F.J.: Delphi study to identify the young people priorities about digital society. In: García-Peñalvo F.J. (eds.) Proceedings of the 6th International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM 2018) (Salamanca, Spain, 24–26 Oct 2018). ACM International Conference Proceeding Series (ICPS). ACM, New York (2018). https://doi.org/10.1145/3284179.3284183

4. García-Peñalvo, F.J., Kearney, N.A.: Networked youth research for empowerment in digital society: the WYRED project. In: García-Peñalvo F.J. (eds.) Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'16) (Salamanca, Spain, 2–4 Nov 2016). ACM International Conference Proceeding Series (ICPS), pp. 3–9. ACM, New York, NY, USA (2016). https://doi.org/10.1145/3012430.3012489

5. García-Peñalvo, F.J.: The WYRED project: a technological platform for a generative research and dialogue about youth perspectives and interests in digital society. J. Inf. Technol. Res. **9**(4), vi–x (2016)

6. García-Peñalvo, F.J., García-Holgado, A.: WYRED, a platform to give young people the voice on the influence of technology in today's society. A citizen science approach. In: Villalba-Condori K.O., García-Peñalvo F.J., Lavonen J., Zapata-Ros M. (eds.) Proceedings of the II Congreso Internacional de Tendencias e Innovación Educativa—CITIE 2018 (Arequipa, Perú, 26–30 Nov 2018), pp. 128–141. CEUR-WS.org, Aachen (2019)

7. Fidalgo-Blanco, A., Sein-Echaluce, M.L., García-Peñalvo, F.J.: Epistemological and ontological spirals: from individual experience in educational innovation to the organisational knowledge in the university sector. Program Electron. Libr. Inf. Syst. **49**(3), 266–288 (2015). https://doi.org/10.1108/prog-06-2014-0033

8. Fidalgo-Blanco, A., Sein-Echaluce, M.L., García-Peñalvo, F.J.: Knowledge spirals in higher education teaching innovation. Int. J. Knowl. Manag. **10**(4), 16–37 (2014). https://doi.org/10.4018/ijkm.2014100102

9. García-Holgado, A., García-Peñalvo, F.J.: Architectural pattern to improve the definition and implementation of eLearning

ecosystems. Sci. Comput. Program. **129**, 20–34 (2016). https://doi.org/10.1016/j.scico.2016.03.010

10. Alspaugh, T.A., Asuncion, H.U., Scacchi, W.: The role of software licenses in open architecture ecosystems. In: Proceedings of the First International Workshop on Software Ecosystems 2009, pp. 4–18. CEUR-WS (2009)

11. Chang, E., West, M.: Digital ecosystems a next generation of the collaborative environment. In: Kotsis G., Taniar D., Pardede E., Ibrahim I.K. (eds.) Proceedings of iiWAS'2006—The Eighth International Conference on Information Integration and Web-based Applications Services, 4–6 Dec 2006, Yogyakarta, Indonesia, pp. 3–24. Austrian Computer Society (2006)

12. Chen, W., Chang, E.: Exploring a digital ecosystem conceptual model and its simulation prototype. In: Proceedings of IEEE International Symposium on Industrial Electronics, 2007 (ISIE 2007), pp. 2933–2938. IEEE (2007). https://doi.org/10.1109/isie.2007.4375080

13. Laanpere, M.: Digital learning ecosystems: rethinking virtual learning environments in the age of social media. Paper Presented at the IFIP-OST'12: Open and Social Technologies for Networked Learning, Taillinn, Estonia (2012)

14. Pata, K.: Meta-design framework for open learning ecosystems. Paper Presented at the Mash-UP Personal Learning Environments (MUP/PLE 2011), London, UK (2011)

15. Mens, T., Claes, M., Grosjean, P., Serebrenik, A.: Studying evolving software ecosystems based on ecological models. In: Mens T., Serebrenik A., Cleve A. (eds.) Evolving Software Systems, pp. 297–326. Springer, Berlin (2014). https://doi.org/10.1007/978-3-642-45398-4_10

16. Lungu, M.F.: Towards reverse engineering software ecosystems. In: 2008 IEEE International Conference on Software Maintenance, ICSM 2008 (Beijing, China, 28 September–4 Oct 2008), pp. 428–431. IEEE (2008). https://doi.org/10.1109/icsm.2008.4658096

17. Lungu, M.F.: Reverse Engineering Software Ecosystems. University of Lugano, Lugano (2009)

18. Nachira, F.: Towards a Network of Digital Business Ecosystems Fostering the Local Development. European Commission, Brussels (2002)

19. García-Holgado, A., García-Peñalvo, F.J.: The evolution of the technological ecosystems: an architectural proposal to enhancing learning processes. In: Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality (TEEM'13) (Salamanca, Spain, 14–15 Nov 2013), pp 565–571. ACM, New York (2013). https://doi.org/10.1145/2536536.2536623

20. García-Holgado, A., García-Peñalvo, F.J.: Architectural pattern for the definition of eLearning ecosystems based on Open Source developments. In: Sierra-Rodríguez J.L., Dodero-Beardo J.M., Burgos D. (eds.) Proceedings of 2014 International Symposium on Computers in Education (SIIE) (Logroño, La Rioja, Spain, 12–14 Nov 2014). Institute of Electrical and Electronics Engineers, pp. 93–98. IEEE Catalog Number CFP1486T-ART (2014). https://doi.org/10.1109/siie.2014.7017711

21. García-Peñalvo, F.J., García-Holgado, A. (eds.): Open Source Solutions for Knowledge Management and Technological Ecosystems. Advances in Knowledge Acquisition, Transfer, and Management (AKATM) Book Series. IGI Global, Hershey (2017)

22. García-Holgado, A., García-Peñalvo, F.J.: Preliminary validation of the metamodel for developing learning ecosystems. In: Dodero J.M., Ibarra Sáiz M.S., Ruiz Rube I. (eds.) Proceedings of the 5th International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM 2017) (Cádiz, Spain, 18–20 Oct 2017). ACM International Conference Proceeding Series (ICPS). ACM, New York (2017). https://doi.org/10.1145/3144826.3145439

23. García-Holgado, A., García-Peñalvo, F.J.: Human interaction in learning ecosystems based in Open Source solutions. In: Learning and Collaboration Technologies. Design, Development and Technological Innovation. 5th International Conference, LCT 2018. Held as Part of HCI International 2018, Las Vegas, NV, USA, 15–20 July 2018, vol. 10924. Springer International Publishing (2018). https://doi.org/10.1007/978-3-319-91743-6_17

24. Rubio Royo, E., Cranfield McKay, S., Nelson-Santana, J.C., Delgado Rodríguez, R.N., Ocon-Carreras, A.A.: Web knowledge turbine as a proposal for personal and professional self-organisation in complex times: application to higher education. J. Inf. Technol. Res. **11**(1), 70–90 (2018). https://doi.org/10.4018/JITR.2018010105

25. García-Holgado, A., García-Peñalvo, F.J.: A metamodel proposal for developing learning ecosystems. In: Zaphiris P., Ioannou A. (eds.) Learning and Collaboration Technologies. Novel Learning Ecosystems. 4th International Conference, LCT 2017. Held as Part of HCI International 2017, Vancouver, BC, Canada, 9–14 July 2017. Proceedings, Part I, vol. 10295, pp. 100–109. Springer International Publishing (2017). https://doi.org/10.1007/978-3-319-58509-3_10

26. García-Peñalvo, F.J., Durán-Escudero, J.: Interaction design principles in WYRED platform. In: Zaphiris P., Ioannou A. (eds.) Learning and Collaboration Technologies. Technology in Education. 4th International Conference, LCT 2017. Held as Part of HCI International 2017, Vancouver, BC, Canada, 9–14 July 2017. Proceedings, Part II, pp. 371–381. Springer International Publishing (2017). https://doi.org/10.1007/978-3-319-58515-4_29

27. Mening, R.: WordPress vs Joomla vs Drupal + CMS "comparison chart" (2014). http://websitesetup.org/cms-comparison-wordpress-vs-joomla-drupal/. Accessed 17 July 2019

28. Rackspace: CMS Comparison: Drupal, Joomla and Wordpress (2013). https://support.rackspace.com/how-to/cms-comparison-drupal-joomla-and-wordpress/. Accessed 17 July 2019

29. WYRED Consortium (2018) Platform v2. WP3_D3.3. V1.2. https://doi.org/10.5281/zenodo.1248813

30. Ministerio de Educación Cultura y Deporte: Datos y cifras del sistema universitario español. Curso 2015/2016. Ministerio de Educación, Cultura y Deporte, Madrid, Spain (2016)

31. Chin, J.P., Diehl, V.A., Norman, K.L.: Development of an instrument measuring user satisfaction of the human–computer interface. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 213–218. ACM (1988)

32. Kirakowski, J., Corbett, M.: SUMI: the software usability measurement inventory. Br. J. Edu. Technol. **24**(3), 210–212 (1993)

33. Lewis, J.R.: IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. Int. J. Hum. Comput. Interact. **7**(1), 57–78 (1995)

34. Sauro, J., Lewis, J.R.: Quantifying the User Experience: Practical Statistics for User Research. Morgan Kaufmann, Burlington (2016)

35. Tullis, T.S., Stetson, J.N.: A comparison of questionnaires for assessing website usability. In: Usability Professional Association Conference, pp. 1–12 (2004)

36. ISO 9241-11: Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs), vol. 45, p. 9. The International Organization for Standardization, Geneva (1998)

37. Brooke, J.: SUS: a retrospective. J. Usabil. Stud. **8**(2), 29–40 (2013)

38. Bangor, A., Kortum, P.T., Miller, J.T.: An empirical evaluation of the system usability scale. Int. J. Hum. Comput. Interact. **24**(6), 574–594 (2008)

39. Brooke, J.: SUS-A quick and dirty usability scale. Usabil. Eval. Ind. **189**(194), 4–7 (1996)

40. Finstad, K.: The system usability scale and non-native english speakers. J. Usabil. Stud. **1**(4), 185–188 (2006)

41. McKinney, W.: Pandas: a foundational Python library for data analysis and statistics. Paper presented at the PyHPC 2011, Workshop on Python for High Performance and Scientific Computing. Seattle, WA, USA (2011)

42. Bangor, A., Kortum, P., Miller, J.: Determining what individual SUS scores mean: adding an adjective rating scale. J. Usabil. Stud. **4**(3), 114–123 (2009)

43. Sauro, J.: A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices. Measuring Usability LLC, Denver (2011)

44. Vázquez-Ingelmo, A.: Code repository that supports the analysis of the usability test (through the System Usability Scale) applied to the WYRED Platform (2018). https://github.com/AndVazquez/wyred-usability-analysis. Accessed 17 July 2019

45. Nielsen, J., Loranger, H.: Prioritizing Web Usability. Pearson Education, London (2006)

46. Kortum, P., Sorber, M.: Measuring the usability of mobile applications for phones and tablets. Int. J. Hum. Comput. Interact. **31**(8), 518–529 (2015)

47. Lewis, J.R., Sauro, J.: The Factor Structure of the System Usability Scale. Human Centered Design, pp. 94–103. Springer, Berlin (2009)

48. Borsci, S., Federici, S., Bacci, S., Gnaldi, M., Bartolucci, F.: Assessing user satisfaction in the era of user experience: comparison of the SUS, UMUX, and UMUX-LITE as a function of product experience. Int. J. Hum. Comput. Interact. **31**(8), 484–495 (2015)

49. Vázquez-Ingelmo, A.: Code repository that supports the analysis in the paper "usability test of WYRED platform" (2018). http://doi.org/10.5281/zenodo.1164124

50. Yamane, T: Statistics: An Introductory Analysis. Harper & Row, NY (1967)

51. Sauro, J.: SUSTisfied? Little-known System Usability Scale facts. User Exp. Mag. **10**(3) (2011). http://uxpamagazine.org/sustified/?lang=es

52. Boone, H.N., Boone, D.A.: Analyzing likert data. J. Ext. **50**(2), 1–5 (2012)

53. McLellan, S., Muddimer, A., Peres, S.C.: The effect of experience on System Usability Scale ratings. J. Usabil. Stud. **7**(2), 56–67 (2012)