How to Include Users in the Design and Development of Cyberinfrastructures?

Hashim Iqbal Chunpir[1,2,3](✉)

[1] Department of Computer Science, Universidade Federal de São Carlos, São Carlos, SP, Brazil
[2] Faculty of Informatics, University of Hamburg, Vogt-Kölln-str. 30, Hamburg, Germany
[3] German Climate Computing Centre (DKRZ), Bundesstraße 45a, Hamburg, Germany
hashim.chunpir@gmail.com

Abstract. Cyberinfrastructures have reached their production level as far as their capability to serve researchers and connect big data is concerned. However, users face difficulties while they perform complex operations via cyberinfrastructures by using their user interfaces (UIs), for big data analysis and research. Using these infrastructures users perform operations such as data access, data visualization to complete their research activities. Unfortunately, there are not enough studies and projects conducted so far that provides guidelines to developers to design and develop interfaces that meet user requirements in cyberinfrastructures. These infrastructures are also known as e-infrastructures, big data infrastructures, open data infrastructures, virtual research environments. In this work, guidelines are recommended so that user requirements can directly be incorporated into the design and development of cyberinfrastructure applications serving a particular target audience. In this paper, an example of a cyberinfrastructure is given, using which users can be involved in its design and the development. These techniques can then also be transferred to the designers and developers of other cyberinfrastructures to improve the user experience as well as usability of UIs and associated services. Furthermore, these techniques can be enhanced even further and generalized to meet the requirements of users of applications other than cyberinfrastructures.

Keywords: e-Research · e-Science · User experience (UX)
Human computer interaction (HCI) · Open data · Cyberinfrastructure (CI)
e-Infrastructures · Usability · Big data infrastructures · Citizen science

1 Introduction

The cyberinfrastructures (CIs) enable digitalization of data, access to data, communication and collaboration amongst the users i.e. researchers. They impact the social, material, technical, political relations of research and enable knowledge production of research communities. The acts of performing science through Information and Communication Technologies (ICTs) have been set under several labels such as big science, data-driven science, networked science, open science, digital humanities and

A. Marcus and W. Wang (Eds.): DUXU 2018, LNCS 10918, pp. 658–672, 2018.
https://doi.org/10.1007/978-3-319-91797-9_46

Science 2.0. Other terms used are: e-Science, e-Social Science, e-Research, e-infrastructure, e-Science infrastructures and CIs [1]. Yet other terms used for these infrastructures are: Science Data Infrastructures (SDI) [2], open data infrastructures [3], big data infrastructures, digital science, collaboratories, virtual science, Virtual Research Environments (VREs) and big data science [4]. E-Science domains include medicine, earth sciences, climate sciences, particle physics, bio-informatics, social sciences and other fields [5]. In this work e-infrastructures and cyberinfrastructures have been used interchangeably.

Since e-Science initiatives are funded by public money, the idea is to create, maintain and provide access to knowledge that is open to all scientific communities including general public. There are various communities that directly interact with the CIs, these are: Users of CIs, developers of e-infrastructures, service providers of CIs, funding agencies of data projects, research projects and other stakeholders [6]. The users interacting with the CIs are of different types and they have different needs, technical abilities and capabilities [6, 7]. These stakeholders not only use tangible and intangible resources e.g. hardware and software but they also develop tangible and intangible resources e.g. data, software, information and knowledge for other stakeholders.

Users need an interface to access resources offered by CIs, usually data and computation power. It is essential to know how these infrastructures are used in reality and who the users are. While there can be thousands of users of a CI, the users who use the infrastructure extensively can be very few [8]. The interface to access CI resources includes command line tools, web portals and Graphical User Interface (GUI) to access data assets; which are the main resources hosted. These interfaces provided by a CI are the key to perform operations facilitated by a CI such as: Creating data, collecting data, storing data, sharing data, publishing data, searching data sets, visualizing data and processing data. The UIs of a CI are designed and implemented normally by the developers.

The future work and research directions for the enterprises especially related to cloud and grid computing technologies can be guided by a greater research objective. Specifically, the objective of making software developers aware, particularly those who design UIs of applications (mobile applications as well as others), to design interfaces that meet users' requirements so that users are able to interact with applications without any trouble. It is suggested to focus the future research in this direction, related to CI based applications because users daily access many terabytes of data through these applications to do science and enable access to these data. In this context, the future research projects and collaborations are proposed to be based on five fundamental research questions:

I. To what extent the concepts of service orientation and meeting users' needs are incorporated into the business models of big data enterprises especially based on grid and cloud technologies?
II. How are the developers of the interactive UIs of applications (on mobile, web and desktop) capturing users' information about the user's interaction, and are addressing users' needs, especially the development of UI components of non CI applications whether mobile applications, web-based or desktop?

III. How are the developers of the interactive UIs of a CI capturing the user's information about the user's interaction and are addressing the user needs, especially the development of UI components of CIs?
IV. Are there any similarities or differences between the user requirements of general-purpose applications and the applications for data infrastructures, i.e. comparing and contrasting research question I and II?
V. What techniques or Human Computer Interaction (HCI) artefacts[1] (e.g. user scenarios, user stories or others) can be utilized in the software engineering process to capture users' requirements to provide better usability and UX to the users of non CI applications as well as applications for CIs?

Answering these research questions by initiating future projects will help users to enhance the user experience (UX) and the usability features in the projects, services and products offered by cloud based enterprises and its industrial partners. For instance, in the field of e-Science this research will help researchers to get the required data and information to perform e-Research using the UI interfaces of a particular CI, thus, encouraging user interaction with the CIs.

It is indeed needed to incorporate user's point of view in the process of development of software and interfaces, whether related to e-Science or not, in such a form that the developers can address users' UI requirements. Since there are different types of users of CIs, it is important to consider all groups of users [6]. Most of the stakeholders, e.g. in e-Science and ICT infrastructures include data scientists, data curators, computer scientists, domain experts, managers and most importantly interface designers as well as software engineers. All can provide input to enable better usability to the users. Usability is always associated with technology design with an aim of getting job done efficiently with satisfaction [9, p. 588]. Observing UX aspects allow developers incorporation of the fundamental design issues but these UX aspects are more about being amazed, having fun and experiencing emotional responses and creating interest [9, p. 588]. These attributes can incite more interest of scientists as well as other practitioners and can act as a catalyst to increase their participation in using scientific data through CIs. All users have different capabilities, the incorporation of user stories can make life easier of many scientists (users) to increase usability, data search and understanding of data sets.

Earlier, the UK e-Science Task Force has also explicitly indicated to consider broad set of user perspectives in studying a user, a support staff, an administrator and a CI developer [10]. This is because incorporating user perspectives in the form of user stories can maximise the use of e-Science technologies and applications to support new forms of scientific community and can be a source of boom of scientific discoveries. Providing better usability via new interfaces and representation of scientific concepts can directly increase the effectiveness of e-Science researchers [10].

In this work, I take a case study of a CI: Earth System Grid Federation (ESGF) to further explain the research directions and the research methodology as an example to

[1] Artefacts (also spelled as artifacts) are the objects modified by humans as opposed to objects that exist in nature. In HCI or interaction design terminology they mean a *tool* or even *activities* in a process, e.g. user stories, scenarios, use cases etc. [47]

guide future research in this field. This research methodology to involve users and designers has been illustrated in Sect. 4. ESGF is an international collaboration with a purpose to develop the software infrastructure needed to study climate change at a global scale. The background about CI and the knowledge gap is given in Sect. 2. The rationale and justification to do such type of research is provided in Sect. 3.

2 Background on ICT Infrastructures and the Knowledge Gap

In the last decades, the ways to conduct research and produce knowledge has been evolving and the traditional ways to perform science has been challenged. The acts of performing research nowadays have been closely related to the use of information and communication technology (ICT) infrastructures [1]. The CIs are pulling people together, allowing collaboration amongst people, facilitating joint initiatives across various cross disciplinary fields, institutions and geographies in the fields of science and humanities. These infrastructures offer new opportunities for sharing and connecting resources such as data, code, publications, computing power, laboratories, instruments and major equipment to generate meaningful information to enable discoveries in science. Infrastructures that have been deployed to access and share the scientific knowledge, data, computing resources and even human resources to facilitate the intra- and inter-disciplinary research are called e-Science infrastructures or e-infrastructures.

The underlying infrastructures that enable e-Science to take place are popularly known as “e-Science infrastructures” in Europe and “cyberinfrastructures” in the US [11]. In this paper we use the term CIs. Moreover, CIs have been widely deployed to share the knowledge by accessing the scientific data, computing resources with joint efforts of other human resources to facilitate the intra- and inter-disciplinary research called e-Science [12]. Other names connected to the concept of e-Science include e-Research, digital science, collaboratories, virtual science and big data science [4]. E-Science domains include medicine, earth sciences, climate sciences, particle physics, bio-informatics, social sciences, humanities and other fields.

There are number of challenges in the arena of open science and CIs. Dealing with the data generated in massive amounts is itself a big challenge [13]. Once the data are generated there are issues of storing this data in data archives [14]. Around 90% of the data that is produced during research projects are lost and not saved [2, 14]. The original raw data is not made accessible and only processed data is published to open data platforms [15, 16]. There are multiple platforms and not a single platform to fetch data. The scientists can easily lose the track of the fragmented sources of data behind the platforms [16]. At times data are temporarily not available on the platforms and some data are only partly available [16]. Some data is only available without relevant information [15]. The use of fragmented software applications, legacy systems complicate the publicizing and processing of data [15].

Besides, there is a resistance amongst scientists to make data public for number of reasons [14, 15]. Uncontrolled use of data, abuse of data and fear that the research results will be published by other scientists are some of the reasons behind this resistance [2, 14, 15]. There are also number of technical concerns related to data replication amongst data centre sites and distribution of data to data centres [14].

There have been some recent achievements in the last couple of years, these include: automation in data management, re-use, re-publish, global data availability, wide public access including existence of infrastructure components as well as improved security and access control providing trusted environments [2]. With these recent developments some of the challenges mentioned before are in an attempt to be resolved. However, it has been observed that the open data that is obtained from the scientific research institutions and organisations has no value in itself; it only becomes valuable when used [15, p. 266]. Yet, it is predominantly the case that the users' view is largely unfortunately neglected and the benefits of scientific data might not realise to full extent [15, p. 266]. Moreover, the use of open data might not be easy [15, p. 266]. In order to use open-data in Science the focus on service orientation cannot be ignored [17]. Moreover, a business model or a vision that provides an efficient user support process is inevitable accompanied with suitable UIs for researchers to enable better UX of CI to get the data and information that they need in a required format.

The barriers associated to the use of open science data via CIs are classified under the following categories as reported by Zuiderwijk et al. (2012) and Schulte et al. (2016) using literature review as well empirical evidence from interviews and discussion forums: (1) availability and access to data, (2) find ability of data, (3) usability, (4) understand ability, (5) quality, (6) linking and combining data, (7) comparability and compatibility, (8) metadata, (9) interaction with the data provider, and (10) opening and uploading [16, 18]. Most of these barriers are pertaining to the access and the use of scientific data via CI interfaces. A search on Google, Scopus, Web of Science with the keywords "e-Science", "e-Research", "user experience", "usability" and their combinations show that there are unfortunately fewer studies conducted that evaluates the current UIs provided by the CIs or analyses the techniques the developers use to incorporate user's perspectives while designing UIs of CIs. Fewer previous studies in this direction in which the author was involved were [19–24].

E-Science UI portals provide web-based or desktop tools to access and run scientific applications and data as well as searching data or linking these applications. The types of UIs used in CI can be categorized into four main categories: desktop utilities, smartphone clients, workflow wrappers and integrated research platforms [25]. There have been initiatives to provide better UI in all these categories, for example; e-Science initiatives in USA such as ESGF. In ESGF project, the new UI web platform powered by ESGF COG (www.earthsystemcog.org), was developed by a UI development team headed by scientists from University of Colorado and NASA.

Desktop utilities are GUIs; an alternate to command line interfaces that help users to run applications and launch commands on remote supercomputers. They provide features such as drag and drop. They hide complexity but at the same time they may limit access to certain features of applications [25]. According to the current ESGF management team, the new UI is user friendlier and provides more features for users than the previous one. These features include better data search, data download, navigation and data visualization facilities. Similarly, in FP7 ENGAGE CI project, advanced search functionality for data is provided via UI. Moreover, ENGAGE claims that data results that the user is looking for, is made easier [3, p. 263]. Navigation of the portal has been improved by providing clear buttons and breadcrumbs. Appropriate UI has been provided in ENGAGE based on state of the art technology using Spring Model-View-Controller (MVC) project [3, p. 263].

Handheld UIs for e-Science also exist e.g. in UK e-Science RealityGrid project, two handheld *smart phone clients* were developed which include features of real-time computational steering and interactive 3D visualisation [26]. Features offered were generic pan, zoom, rotate etc. *Workflow wrappers* are a collection of tools to automate common operations regarding a scientific workflow. Examples include KEPLER, or more recently in the bioinformatics domain, BIOCONDUCTOR or GENEPATTERN [25]. *Integrated research platforms* are tools that allow users to combine and/or explore several tools and perform data observations. Such tools are not mature and are currently hot research areas [25]. An example of such a platform is the Kidneyome portal. Other notable achievements in the UI of CIs include the development of friendly interfaces for life scientists, specifically for the community of human genetic researchers in the project Hope [27]. However, the user rating of the GUIs of any of these platforms has not been done. Moreover, there is no evidence that while offering UIs to the users the picture of different types of users is kept in mind by the developers during the development and design of UI of a CI. It is important to capture users' mental model of using the UI of a particular application. An interesting work in this area for the further reading and background in mental models is [28].

There is a substantial progress in development of user friendly UI for CIs. This notable advancement in the UI development has been achieved due to use of state of the art technologies that supports better UI development for users. Yet, it is important to note that the effectiveness of UI from the end-users or other stakeholder's point of view has not been studied.

In order to provide better usability to the users of CIs, the proposed research directions will contribute in producing Human Computer Interaction (HCI) artefacts that shall be used by developers of the software part of a CI to enable better usability and UX to the users. These artefacts will be used as a supporting tool for developers to discover and formalize the issues related to the interaction of the end-users with the CI. Furthermore, the purpose of these artefacts is to support developers of the software components of a CI, in meeting the user needs in order to provide suitable interaction possibilities for the users to interact with CIs to fulfil the needs of users. Better user experiences as a result can attract more and more scientists to perform e-Research; hence, discoveries can be made easier and faster due to user friendly interfaces provided by CIs. Similar HCI artefacts can also be used to capture user requirements of other general purpose ICT applications as well.

3 Justification and the Rationale

The data offered by a CI using cloud and grid technologies must not just be seen as a product [29] but as an ongoing movement to utilize this data as well as the infrastructure capacities by the users to come up with new scientific discoveries. The software applications that provide an interface to CI are the doors to get the required data and to transform this data into information, knowledge and create a shared understanding of a scientific observation or a concept thus contributing to wisdom. Key to these doors is providing better usability, UX and accessibility to the infrastructures. In essence, these e-Science applications can contribute to Data, Information, Knowledge, Wisdom (DIKW) hierarchy. This research agenda and pathway shown in this work will help all the stakeholders especially data users, data providers to search, fetch, download, visualize, re-use, annotate, re-analyse and perform computing operations on data. In this way, more value from the data can be derived that can benefit the data publishing as well as data users community by making their life easier.

Many studies claim the benefits that can be realized by the open data and open data infrastructures [13, 16]. However, to realize the benefits and creating a public value one needs to overcome the barriers, especially the barriers to use the CIs and their associated applications. Unfortunately, there is hardly a work already found in literature search that incorporates user perspectives in the form of artefacts for developing UIs for CIs. The proposed research direction is a first attempt to create value for users in ICT infrastructures.

Previously, the user support process of a well-established CI: Earth System Grid Federation (ESGF) was analyzed by collecting data from various sources such as interviews, survey questionnaire, field study from the stakeholders, thus enabling data triangulation [7, 22]. ESGF comprises of geographically distributed peer nodes that are independently administered, yet united by common federation protocols and application interfaces [30].

The data analysis from various sources provided the information about the view and the understanding of the previous user support process in ESGF. Based on this analysis, the problems were identified and the solutions were suggested in the form of a recommendation framework; Federated User Support Enhancement (FeUSE) framework was compiled [6, pp. 231–299]. Finally, based on the findings, the user support process in ESGF evolved. A significant outcome of that study was the identification of end-users and the separation of end-users from the other stakeholders [22, 31]. It was furthermore emphasized after studying the user support process of ESGF that the CI is made for the users i.e. researchers who want to find, access, download and process data and therefore the users are needed to be supported actively. Based on the user interaction with the CI in the form of e-mail queries, the categories of user problems, the workarounds and solutions were suggested. The detailed findings and the contribution can be found in [6, 7, 17, 22–24, 31–33].

This work in the form of a research agenda aims to further deepen this topic and initiate further projects and collaborations with new industrial partners. Furthermore, it intends to include the usability aspects especially in the form of introducing artefacts to the developers of the UIs of CI software applications and other domains to provide better usability and UX features. In order to study the UIs and the development of UIs of CIs, a strong partnership is recommended to be planned with the participating institutions of CIs. It is claimed that the UIs for the users in CIs like ESGF, ENGAGE and others are present and with the passage of time, there have been attempts to improve them [3, 34]. However, these interfaces have not been evaluated and secondly, there exists no aid mechanisms for developers in the form of HCI artefacts to incorporate user's point of view.

4 Research Methods and Research Plan

A case study method is selected to explore the concept of user's involvement, UX and the usability of ICT facilities in a CI. In future research, each relevant project with the UX and usability aspects that organisations may undertake can be seen as a single case, eventually each case can be compared later on. A case study focuses on understanding the dynamics present within single settings [35]. It can also be defined as an empirical inquiry that investigates a contemporary phenomenon within its real-life context [36]. ESGF[2], along with its all partner institutions and its associated data projects is chosen as a case study to explain this research agenda with the help of a study design i.e. research method. Furthermore, it is explained how the software UIs, that already exist, can be further designed and developed by the development team of ESGF. ESGF is one of the major peer-to-peer international federated CI effort in climate science research [37, 38]. It has elements similar to other CIs and that's why chosen as an example in this work.

The main research goal of this proposed future research agenda is to see how the user's requirements shall be integrated in the design process of UI and how the needs of users shall be met by the developers of the CI's software interface. To address this research goal, firstly, it is important to investigate what methods, techniques are currently being used by the developers to incorporate users' needs in the UIs of a CI and secondly, how the developers meet user requirements when they design and build tools for users. In order to accomplish this in this exposé; Design Science Research (DSR) approach combined with Cooperative Method Development (CMD) is chosen in the form of Software Cooperative Design Research (SoftCoDer) approach as proposed by Choma et al. (2015). CMD is an adaptation of Action Research (AR) [39] (see Fig. 1).

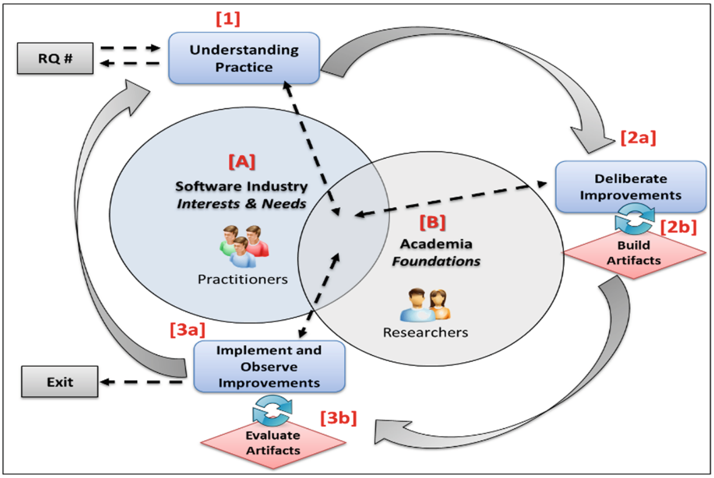

Fig. 1. The SoftCoDer approach [39].

2 www.esgf.llnl.gov.

DSR is based on building artefacts and evaluating them based on an observation and a prior knowledge about a particular business setting. According to Hevner, the major purpose of DSR is achieving knowledge to understand the problem domain by designing an artefact [40]. Therefore, in DSR process artefacts are built to suite the business needs in a particular context. DSR is commonly used in Information Systems (IS) as well as in Software Engineering [41]. DSR combined with CMD in the form of SoftCoDer approach is well suited for this proposed future agenda because of its relevance with the social settings, software engineering context, artefacts development and improvement and incorporation of the user's perspectives. Moreover, it helps to improve techniques and processes in problem-oriented social settings [39]. Another reason for the choice of SoftCoDer approach is because it is based on systematic body of evidence and is aimed at understanding and improving human performance [39]. The researchers need to look at the practitioners approach in the industrial projects about developing Software GUIs for the users that is why SoftCoDer approach is deemed suitable. The phases are suggested to the researchers as an example who would like to pursue research in this direction, keeping with the guidelines of the SoftCoDer approach.

5 Phases Suggested to the Researchers as an Example

The following phases are suggested to the researchers who would like to involve users in the design and development of a CI. Here ESGF CI is given as an example. These phases are illustrated in the form of guidelines to the researchers who would like to pursue research in embedding users in the design and development of CIs, keeping with the steps of the SoftCoDer approach:

5.1 Phase 1 - Understanding the Practice

In this proposed research, understanding the current process of UI development in CIs is needed to be observed by studying the ways, developers are currently meeting user requirements of the UI design and developing software facilities for users in CI. The researchers can then compare it with the UI development of general purpose applications i.e. non CI applications. This can be achieved by following steps:

Step 1: First a research problem must be defined based on the prior knowledge about this topic, literature and other sources. According to the literature review conducted so far, it is known that users are not able to exploit the potential of open data to the fullest [3]. There have been problems found in ICT facilities of CIs such as registration problems, data access and download problems [6]. These problems are related to the UI of applications of a CI directly or indirectly. Moreover, as according to Zuiderwijk et al. (2013): services for the use of open data towards end-users are often lacking. Keeping with the evidence present in the literature, the problem definition has already been done to some extent.

Step 2: In the second step, in order to achieve the research objectives, data should be collected from different sources in order to understand the current practices (also known as *"as-is"*) of the developers involved in designing and developing of UI, as indicated in the SoftCoDer approach. The following sub-steps are suggested to capture *as-is*:

(a) Reviewing the literature of the use of methods, software engineering techniques, HCI artefacts being used by developers and designers in industry or other sectors to incorporate user's requirements in UIs of information systems will be the foundation.

(b) A web-based survey questionnaire is proposed to be designed to get the current picture of UX of applications, keeping with the guidelines provided by Usability Metric for User eXperience (UMUX) [42–45] and System Usability Scale (SUS) [46]. After conducting the survey-questionnaire, the data shall be collected and then analysed. The target respondents of the survey-questionnaire shall be the stakeholders; mainly the users of a particular ICT infrastructure, for instance; ESGF and ESGF-related global projects. It is estimated that the total users of ESGF are 25,000 out of which only few hundreds are active. The responses from the users could provide the current picture of usability of ESGF.

(c) Semi-structured interviews shall be conducted with the stakeholders of organisations participating in ESGF e.g. about the current techniques the stakeholders apply to capture user requirements for UI of CI software applications. For instance, interviewing 20 stakeholders, especially developers out of the population of 55 active employees in ESGF is a good option. It is estimated that the people involved in the development of software of ESGF are not more than 20.

From these sub-steps of the first phase of SoftCoDer approach, the current UX and usability of UI of the applications offered by ESGF to select and download big data shall be gauged. Moreover, from these sub-steps the current methods to capture user requirements can be known. The problem statement shall be re-visited and re-defined. From the data gathered from these steps, proposed above; an overall-picture of the usability and user requirements in ESGF as well as associated projects shall be portrayed in the form of a report describing major features, strengths, weaknesses and the problems found in the aspects of usability and UX in e-Science. Moreover, the methods, techniques, artefacts suggested in literature and other software engineering projects to capture requirements of users must be studied and selected. Thus, in this way in the business model of e-Science organisation, user service orientation can be focussed on.

These methods, techniques, artefacts can be suggested to the developers in the next step called deliberate improvements, so that the developers and designers of UI could use them in the development process.

5.2 Phase 2a - Deliberate Improvements

Based on the observation of the current practices and theoretical groundings emanating from the first phase, the possible improvements in the process of development of UI for users in e-infrastructures shall be suggested. Furthermore, based on this analysis, recommendations shall be made to the developers to use the techniques, methods or artefacts that help to improve the usability features of UIs of CIs.

Moreover, it shall then be instructed to use these suggested HCI artefacts or techniques by the ESGF management that the developers agree upon to incorporate users' points of view in the CI. The HCI artefacts and techniques are intended to contribute to enhance the overall usability and UX of CIs and to make the UIs more usable and pleasurable.

5.3 Phase 2b - Building Artefacts

The suggested recommendations of the artefacts from the last phase shall be presented to the ESGF project stakeholder's committee and demonstrated to the developers. In fact, in this phase, the actual implementation of artefacts will be made in the UI development process and artefacts will be created in collaboration with the researchers as well as the development team of the collaborating partners e.g. in this case ESGF. These artefacts will then be used in the development of UI for users in CIs with the help of practitioners of ESGF.

5.4 Phase 3a - Implement and Observe Improvements

An online questionnaire shall be carefully designed to create a common agreement and consensus amongst the researchers as well as the practitioners of ESGF about the implementation of the recommended artefacts in the development process of UIs. Furthermore, these suggestions in the form of artefacts shall be put into practice under the guidance from the ESGF management.

In this phase, the researchers will observe the implementation of using artefacts in the development process of UIs of the desktop or web applications of a CI. From the group that used artefacts or other techniques, multiple sub-groups of developers, will be instructed to create a UI prototype for users using a single artefact. The other group (control group) will create a UI prototype for users without using any artefact.

Moreover, after the implementation, a discussion in the form of a workshop with the practitioners will be organised to see each other's UI prototype, rate it and suggest possible improvements. This step will be conducted with the help of ESGF management. The researchers shall guide ESGF management in this aspect to take appropriate decisions.

5.5 Phase 3b - Evaluate Artefacts

The UI prototype based on each artefact shall then be rated by the users via an online questionnaire. At the same time, the suggested artefacts must be rated by the UI developers from ESGF and outside ESGF. From the users and the practitioners involved in this evaluation and rating process the lessons can then be drawn.

5.6 Phase 4 - Communication and Dissemination of Results

The communication of the results of the project shall take place after every 6 months from the initiation of the project with the ESGF and its partner institutions. Finally, the lessons learned and the final results of the project and the artefacts will be communicated to the whole ESGF team in a meeting. Moreover these results and artefacts can then be communicated in the final report of the proposed project, technical reports, journal papers and conference proceedings.

A validation of results and evaluation of the outcome can be made with the help of focused groups again. Moreover, future directions can be given to other collaborating CI projects so that the results can then be generalized with other CI and non CI projects.

6 Expected Future Outcomes

The aim of this work is to provide recommendations in the form of artefacts to developers and the policy makers of big data enterprises especially from the government sector that fund CI projects, so that they are able to develop and design UI in CIs that are eventually usable by the users and offer better UX for future. The recommendations in the form of artefacts can be bundled in a proposed framework called *e-Science UI usability (eUIu) model* and then it can be compared with the usability guidelines for non CI applications. This model can provide significant input to the enterprise model and the business model of CIs and other enterprises. With this model both the developers of CI and developers of other infrastructures will benefit. Moreover, they will be able to enable better UX and usability for users in various domains of data science such as climate science, medical science, physics and others. Besides, a set of proposed web-based UI prototypes and visualization environments shall be created by the developers to support better UX in CIs. Furthermore, it is expected that a shared understanding amongst developers about the user's point of view will be developed, facilitated by the artefacts. As a result of better interfaces the open science data will be made easily accessible to the public and most predominantly to the researchers using CIs.

The eventual outcome of these future research directions is improvement in the usability and the UX of the UIs provided by a CI, e.g. ESGF. If the UX and usability is enhanced, there can be a significant boom in e-Research, scientific discoveries and more interest in the research community in using CIs. Consequently, this work is intended to start a usability movement combined with software development derived towards providing better UX and fulfilling users' needs. It can initiate some pioneer projects within and across the domains offered by CIs that can contribute to the age of easy to use e-Science in the form of reducing usability problems and promoting an open e-Research that in turn encourages collaboration amongst scientists, other practitioners and even public i.e. citizens at a later stage.

References

1. Karasti, H., Millerand, F., Hine, C.M.: Knowledge infrastructures: part I. Sci. Technol. Stud. **29**(1), 2–12 (2016)
2. Demchenko, Y., Zhao, Z., Grosso, P., Wibisono, A., De Laat, C.: 2012 IEEE 4th International Conference on Cloud Computing Technology and Science Addressing Big Data Challenges for Scientific Data Infrastructure, pp. 614–617 (2012)
3. Zuiderwijk, A., Janssen, M., Jaffery, K.: Towards an e-infrastructure to support the provision and use of open data. In: Conference for eDemocracy and Open Government, pp. 259–275 (2013)
4. Hey, T., Trefethen, A.E.: Cyberinfrastructure for e-science. Science **308**(5723), 817–821 (2005)
5. Chunpir, H., Moll, A.: Analysis of marine ecosystems: usability, visualization and community collaboration challenges. Procedia Manuf. **3**, 3262–3265 (2015)
6. Chunpir, H.I.: Enhancing User Support Process in Federated E-Science. University of Hamburg (2015)
7. Chunpir, H.I., Ludwig, T., Badewi, A.A.: Using soft systems methodology (SSM) in understanding current user-support scenario in the climate science domain of cyber-infrastructures. In: Marcus, A. (ed.) DUXU 2014. LNCS, vol. 8519, pp. 495–506. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07635-5_48
8. Freeman, P.A.: Is it possible to define cyberinfrastructure ? First Monday 6(12) (2007)
9. Preece, J.: Citizen science: new research challenges for human – computer interaction. Int. J. Hum. Comput. Interact. **32**(8), 585–612 (2016)
10. UK e-Science Usability Task Force. Usability Research Challenges in e-Science (2009). http://www.cs.nott.ac.uk/~psztar/UTF.pdf. Accessed 19 Sept 2016
11. Jirotka, M., Lee, C.P., Olson, G.M.: Supporting scientific collaboration: methods, tools and concepts. Comput. Support. Coop. Work **22**(4), 667–715 (2013)
12. Chunpir, H.I., Zaina, L.: Assisting users in open data infrastructures: a management perspective. In: 15th Internationa Conference WWW/Internet 2016, pp. 27–34 (2016)
13. Chen, J., et al.: Big data challenge: a data management perspective. Front. Comput. Sci. **7**(2), 157–164 (2013)
14. Rudolph, D., Thoring, A., Vogl, R.: Research data management: wishful thinking or reality? PIK – Prax. der Informationsverarbeitung und Kommun. **38**(3–4), 113–120 (2015)
15. Janssen, M., Charalabidis, Y., Zuiderwijk, A.: Benefits, adoption barriers and myths of open data and open government. Inf. Syst. Manag. **29**(4), 258–268 (2012)
16. Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., Alibaks, R.S.: Socio-technical impediments of open data. Electron. J. e-Gov. **10**(2), 156–172 (2012)
17. Chunpir, H.I., Ludwig, T., Williams, D.N.: Evolution of e-Research?: from infrastructure development to service orientation. In: DUXU 2015, Part III, LNCS, vol. 9188, pp. 25–35. Springer, Cham (2015)
18. Schulte, F., Chunpir, H.I., Voß, S.: Open data evolution in information systems research: considering cases of data-intensive transportation and grid systems. In: Proceedings of 5th International Conference, DUXU 2016, Held as Part of HCI International 2016, Part III, Toronto, Canada, 17–22 July 2016, pp. 193–201 (2016)
19. Chunpir, H.I., Curri, E., Zaina, L., Ludwig, T.: Improving user interfaces for a request tracking system: best practical RT. In: Yamamoto, S. (ed.) HIMI 2016. LNCS, vol. 9735, pp. 391–401. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40397-7_37

20. Chunpir, H.I.: Prioritizing tasks using user-support-worker's activity model (USWAM). In: Yamamoto, S. (ed.) HIMI 2016. LNCS, vol. 9735, pp. 379–390. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40397-7_36
21. Chunpir, H.I., Williams, D., Ludwig, T.: User experience (UX) of a big data infrastructure. In: Yamamoto, S. (ed.) HIMI 2017. LNCS, vol. 10274, pp. 467–474. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58524-6_37
22. Chunpir, H.I., Ludwig, T., Badewi, A.: A snap-shot of user support services in Earth System Grid Federation (ESGF): a use case of climate cyber-infrastructures. In: Proceedings of the 5th Applied Human Factors and Ergonomics (AHFE) Conference, July 2014
23. Chunpir, H.I., Rathmann, T., Ludwig, T.: The need for a tool to support users of e-science infrastructures in a virtual laboratory environment. Procedia Manuf. **3**, 3375–3382 (2015)
24. Chunpir, H.I., Badewi, A.A., Ludwig, T.: User support system in the complex environment. In: Marcus, A. (ed.) DUXU 2014. LNCS, vol. 8520, pp. 392–402. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07638-6_38
25. Harris, P.J., et al.: The virtual kidney: an eScience interface and Grid portal. Philos. Trans. R. Soc. London Ser. A Math. Phys. Eng. Sci. **367**, 2141–2159 (2009)
26. Mancini, E.P., et al.: Simulation in the cloud using handheld devices to cite this version: simulation in the cloud using handheld devices (2012)
27. Venselaar, H., Ah, T., Kuipers, R.K.P., Hekkelman, M.L., Vriend, G.: Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinform. **11**, 548 (2010)
28. Chunpir, H.I., Ludwig, T.: A software to capture mental models. In: Antona, M., Stephanidis, C. (eds.) UAHCI 2017. LNCS, vol. 10279, pp. 393–409. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58700-4_32
29. Zuiderwijk A., Janssen, M.: A comparison of open data policies and their implementation in two dutch ministries. In: ACM 13th Annual International Conference on Digital Government Research (Dg.O 2012) (2012)
30. Williams, D.N.: ESGF project overview (2015)
31. Chunpir, H.I., Ludwig, T.: Reviewing the governance structure of end-user support in e-Science infrastructures. In: Informatik 2014 Proceedings on Mastering Big Data Complexity. Lecture Notes in Informatics (LNI), vol. 232 (2014)
32. Chunpir, H.I., Ludwig, T., Curri, E.: Improving processes for user support in e-Science. In: IEEE 10th International Conference on e-Science (e-Science), vol. 2, pp. 87–90 (2014)
33. Chunpir, H.I.: Prioritizing Tasks Using User-Support-Worker's Activity Model (USWAM). In: Yamamoto, S. (ed.) HIMI 2016. LNCS, vol. 9735, pp. 379–390. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40397-7_36
34. Chunpir, H., Williams, D., Cinquini, L., Kindermann, S.: Third annual ESGF and ultrascale visualization climate data analysis tools face-to-face meeting report. Livermore, CA (2013)
35. Yin, R.: Case Study Research: Design and Methods, 5th edn. Sage Publishing, Thousand Oaks (2013)
36. Eisenhardt, K.: Building theories from case study research. Acad. Manag. Rev. **14**(4), 532–550 (1989)
37. Cinquini, L., et al.: The Earth System Grid Federation (ESGF): an open infrastructure for access to distributed geospatial data. In: 8th IEEE International Conference on E-Science, pp. 1–10 (2012)
38. Williams, D.N., Bell, G., Cinquini, L., Fox, P., Harney, J., Goldstone, R.: ESGF: federated and integrated climate data from multiple sources. In: Hiller, W., Budich, R., Redler, R. (eds.) Earth System Modelling, vol. 6, pp. 61–77. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-37244-5_7

39. Choma, J., Zaina, L., Silva Da Silva, T.: Towards an approach matching CMD and DSR to improve the Academia-Industry software development partnership. In: 29th Brazilian Symposium on Software Engineering, pp. 51–60 (2015)
40. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Q. **28**(1), 75–105 (2004)
41. Livari, J., Venable, J.: Action research and design science research seemingly similar but decisively dissimilar. In: Proceedings of ECIS 2009 European Conference on Information Systems (2009)
42. Finstad, K.: The system usability scale and non-native English speakers. English **1**, 185–188 (2006)
43. Finstad, K.: Interacting with computers the usability metric for user experience. Interact. Comput. **22**, 323–327 (2010)
44. Finstad, K.: Response interpolation and scale sensitivity: evidence against 5-point scales. J. Usability Stud. **5**, 104–110 (2010)
45. Finstad, K.: Response to commentaries on 'The usability metric for user experience'. Interact. Comput. **25**, 327–330 (2013)
46. Sauro, J.: Measuring Usability With The System Usability Scale (SUS) Measuring Usability (2011)
47. Papantoniou, B., et al.: The Glossary of Human Computer Interaction. Interaction Design Foundation. https://www.interaction-design.org/literature/book/the-glossary-of-human-computer-interaction/artifact. Accessed 22 Feb 2018