Improving Processes for User Support in e-Science

Hashim Iqbal Chunpir
German Climate Computing Centre (DKRZ)
University of Hamburg
Hamburg, Germany
Chunpir@dkrz.de

Thomas Ludwig
German Climate Computing Centre (DKRZ)
University of Hamburg
Hamburg, Germany
Ludwig@dkrz.de

Endrit Curri
University of Hamburg
Hamburg, Germany
1curri@informatik.uni-hamburg.de

***Abstract*—Support processes play an important role to facilitate researchers (users) to accomplish their research activities with the help of cyber-infrastructure(s). However, the current user-support practices in cyber-infrastructure are being followed without an accountability and ownership of tasks thus over-burdening user support staffs who are performing user support on a free-will basis. This research project explores the current state of the user support process in Earth System Grid Federation (ESGF) from support staffs' perspective, taken as a case study in Earth System Modeling (ESM), a global cyber-infrastructure project. The study proposes an improved framework of best practices for servicing end-users of e-Science infrastructures that can be applied to enhance the user support process of ESGF and other cyber-infrastructures.**

Keywords—e-Science User Services; e-Research in Earth Sciences; Information Technology Services Management (ITSM); user support process; service desk; management of support activities; help desk.

I. INTRODUCTION AND RELATED WORK

Cyber-Infrastructures (CI) have been widely deployed to access and share the knowledge, data, computing resources and even human resources to facilitate the intra- and inter-disciplinary research called e-Science. There are different names associated to the concept of e-Science. The underlying infrastructure that enables e-Science is popularly known as "e-Science infrastructures" in Europe and "cyber-infrastructures" in the US [1]. Other names connected to the concept of e-Science include e-Research, digital science, collaboratories, virtual science and Big Data Science [2]. These names have been used in connotations other than e-Science as well. E-Science domains include Medicine, Earth Sciences, Climate Sciences, Particle Physics, Bio-Informatics, Social Sciences and other fields. An important example of a well-developed CI in the field of Climate Science and Earth Science Modeling (ESM) is Earth System Grid Federation (ESGF). It is one of the major peer-to-peer international federated CI efforts in climate change research [3]. ESGF is servicing more than 27,000 researchers from different continents who are accessing huge amount of climate data for climate-model inter-comparison purposes from widely distributed data-archives all over the world to study climate impact as well as core climate modelling [2,3]. Subsequently, ESGF is taken as a use-case in this study. Though ESGF is a unique CI use case in the climate domain having commonality features with most of the CI that make it substantially representative of an e-Science community. These basic traits include federated architecture, multiple data projects hosted, data search capability, distributed nature of data as well as compute nodes, visualization software and globally strewn users. This is the first study, clearly more studies are needed in other CI to compare and contrast findings of this study.

The *production grid and cloud technologies* are mature enough to achieve roughly reliable federated CI. However, the IT management and administration techniques including supporting users of the federated CI are currently deficient to ensure service quality in heterogeneous, large-scale and distributed resources. The user-support process in CI is an operational process that serves the end-users of e-Science infrastructures in achieving their goals i.e. using CI mainly for their research but not necessarily limited to that. This process can be compared with the service operations and service desk function described in ITIL-Service Operations [5]. A help-desk allows users to contact support staffs to address particular problems of a user [6]. Service desk is the concept that combines service management studies with the traditional customer support studies that uses the term *help-de*sk [7]. End-users are mainly researchers and they accomplish various tasks of research within a specific time-frame via CI. The user-requests are the inputs to this process and are transformed by the CI staffs involved in user support, also known as support employees with the help of tools and methods, whether automated or manual into solutions. This phenomenon is called servicing the user-requests. The CI staffs that service user-requests have also other tasks to be accomplished, for instance: programming, data curation apart from servicing end-user requests. Normally in CI there is less institutional expertise than is needed to service user requests. The user support process in CI projects is iterative and dynamic in nature same as the nature of CI itself.

Currently, the user support in ESGF and other projects from other domains including EGI, C3Grid is being offered in the form of self-help via support websites, online tutorials, wikis or contacting an expert in the form of traditional help-desk or service-desk. However, few studies have explored the user support process in CI via empirical investigation and included users and employees to get their perspective in the enhancement of the process. In addition, these CI projects did not carry out a detailed study and empirical investigation to

enhance the user support process or to measure user satisfaction. Thus, the end-user support is being offered based on the past experiences. This study provides a framework to streamline the user support process in CI by means of analyzing the methods, tools and techniques employed via CI governing bodies as well as human support agents and finally proposing improvements.

Until now, there are different versions of business service frameworks such as IT Infrastructure Library (ITIL) that provides best practice guidelines for servicing end-users and customers especially in the industry sector locally and globally [8]. Some of the commercial user support frameworks have been modified and adapted to academic setups such as universities [9] and to governmental administrative bodies [10]. User support in industry is different from CI in the sense that the user-queries often relate to technical and scientific problems with a globally distributed federated e-infrastructure where one or more components and support staffs may be under the jurisdiction of different institutions, present in different continents. Additionally, different legal codes, management paradigms, socio-cultural norms, sustainability, resources and dynamics are different amongst the partner organisations forming a CI [11]. This requires defining a model of the federated CI to deal with services it offers to users and the funding aspects [11]. Therefore, there are some efforts such as the gSLM[1] and FedSM[2] projects that were initiated to provide service management roadmaps for federated CI since it was identified that the structure and nature of CI is different from IT infrastructures in industry. However, these frameworks have not yet been tested empirically and applied to any particular field of e-Science. Moreover, there are few studies conducted that address the issue of improving the user support process to facilitate users as well as staffs in CI. Likewise, hardly any studies have been carried out to investigate economic and human resource factors involved in user support process of CI that keep this process feasible for all the stakeholders. It is clear that if the barriers of users in using the CI are not addressed, then the best utility of CI may not be achieved in spite of the technical excellence that has been achieved so far. Therefore, it is imperative to make the user support process smoother which would finally lead to identification of core problems in usability of CI and enable to mitigate them.

II. Research Background

In order to answer the main research question: "How to enhance the current end-user support situation in ESGF e-Science infrastructure?" it is important to investigate and have a clear understanding of *as-is* i.e. the current status of user support in ESGF CI. In addition, a clear picture of the user support process with a focus of interaction of the support staff with the users and vice versa via current user support tools, techniques and procedures are necessary to study. Equally, support staff's governance, organization and management of incoming user requests are imperative to consider. The main research problem can be further divided into the following research questions that provide a roadmap for addressing the main research question:

(i) What are the characteristics of CI that have an effect on user-support, particularly in the climate science domain and generally in other domains?

(ii) What is the current state of user-support in Climate e-Science? Finding out the strengths and weaknesses of the current user-support practices in the ESGF use case.

(iii) What are the common user-support procedures which are recommended in literature (e.g. in IT industry or other industries) that might be appropriate to improve the user support process in climate e-Science infrastructures?

(iv) What types of challenges are related to user-support operations (e.g. service desk, event management of e-infrastructure, problem management and incident management) in e-Science?

(v) How can user-support requests be classified in Climate Science domain of e-Science Infrastructures?

Finding an answer to these questions based on proper evidence will help to improve the state of end-user support in CI. Thus, on one hand this study is aimed to enable pain free operational advantage to the users of a CI to make advances in their field while using data for their research. On the other hand this research facilitates the management of CI to save resources while servicing users and providing satisfaction to the staffs that support the users and the users themselves as well. In essence this research enables the e-Research facilities to be more effective and usable.

III. Research Methodology

This research project uses the *mixed research approach*. The *mixed research approach* integrates multiple quantitative and qualitative methods to catch diverse perspectives on a subject matter under study [12]. However, the focus of this study is more on the qualitative nature of study than the quantitative investigation. A case study is "a research strategy which focuses on understanding the dynamics present with single settings" [8]. It can also be defined as "an empirical inquiry that investigates a contemporary phenomenon within its real-life context" [9]. The case study approach was selected as a research method because it suits well for studying service processes that are linked to a complex organizational context, distributed user and support teams in CI such as ESGF [10].

The research plan of this study comprises of six-phases: In the *first phase*, a research problem is defined based on some prior knowledge, participant observations, weekly meetings with focus groups, informal coffee table discussions as well as e-mails and relevant literature. In the *second phase* the following steps were conducted in parallel: (A) a survey questionnaire for the staffs of ESGF involved in supporting users was designed, piloted; the data was gathered and analyzed. 26 usable responses (N=26) were collected from the respondents of the survey questionnaire who filled it out. The

[1] http://www.gslm.eu/

[2] http://www.fedsm.eu/

respondents were the staffs of ESGF and its associated projects which support users. The total number of employees of ESGF and its associated projects involved in supporting users of ESGF range from 30 to 40. The level of their involvement varies as well. (B) 12 semi-structured narrative interviews were conducted with the stakeholders of ESGF including policy makers, users, senior managers and operational staffs to capture more insight into the user support process and to validate the data collected via the questionnaire. (C) A field study was conducted to observe the user-support activity especially interaction of support-staffs with the tools they use along with their interfaces to support users. However, the results have not yet been analyzed. (D) A mailing-list (ML) comprising of user-requests and answers from user-support staffs was also observed to find common patterns, classification of incoming user-requests, solutions, time-delays and other factors.

In *phase 3*, all data gathered from phase 2(A, B, C, D) was analyzed together and a picture of the current user-support process was portrayed in the form of a report describing the strengths and weaknesses of the current user-support process. Rich pictures were drawn in the form of diagrams to communicate the results of the analysis to the stakeholders. Recommendations were then made that can be implemented to enrich the current user-support process at the strategic level, the service design level and an operational level of ESGF CI. The recommendations are put into practice in *phase 4* by streamlining the whole user-support process in ESGF, where-ever possible within the resource limit. In *phase 5* an evaluation of the new user-support process is made and an analysis is performed. The improvements in the user-support process are monitored in the form of continual service improvement recommendations. In the last *phase 6*, the elements within the user-support process of ESGF are planned to be highlighted and should be generalized and hence put into practice in other CI support frameworks. Applying support practices and models to other fields are out of scope of this study.

IV. Status of Research

At this point of time, this research project is at *phase 2* and *phase 3* of the research plan. These phases contain 12 interviews and a survey questionnaire analysis. The survey questionnaire included 40 closed-ended questions and 3 open ended questions. The data-analysis of the results describes the current user support-process present in ESGF along with its strengths and weaknesses. It was found out that a substantial amount of time of ESGF employees was spent to support users from 1 to 2 hours daily as according to 63% of the respondents.

The ESGF employees are top scientists thus dedicating time to support users in trivial matters is not the best use of well-paid specialized human resources. Moreover, 62% of the incoming user queries can be solved within 5 minutes thus implying that most of these queries are of simple nature such as resetting password. Likewise, the culture of self-help amongst the users has not been followed partly because the relevant information on the web-page was either missing or was not completely comprehendible. Nonetheless, the staffs of ESGF support users on volunteer basis as it is not their core job and they are not paid extra to do it. It was also found out from the questionnaire that up to 15% of the incoming user queries are ignored. This result was also backed by the interviews conducted and is essentially an alarming situation because it has a repercussion on e-Science as a whole including the utility as well as usability of a CI. The main communication media used in the current user support process is text-based e-mail, received by the staffs via a mailing list or a request tracking system (RTS). There was a lack of consensus of using a particular RTS amongst the members of the partner organisations in ESGF. Therefore, during the study it was suggested to look for a platform where users can post queries and get answers in the same manner as the *stackoverflow*[3] forum and partly also to involve users. This suggestion was welcomed by the ESGF committee and *Askbot*[4] was chosen to replace the current mailing-list in the long run. It is important to note that some aspects of the outcomes of the survey-questionnaire have already been elaborated in detail in two papers published recently [15,16].

Improvements in the governance structure of ESGF has been suggested by defining roles and accountability after finding out the current structure of ESGF via interviews, participant observations and associated documents examined by the authors [16]. Similarly the organizational structure of ESGF has also been found out and generalized to fit other e-Science infrastructures as well. The parts of this structure have been identified in e-Science as sponsors, data-projects, the federated data archive i.e. CI, principal investigators and the node(s) [17]. The description and representation of these parts are given in Fig. 1.

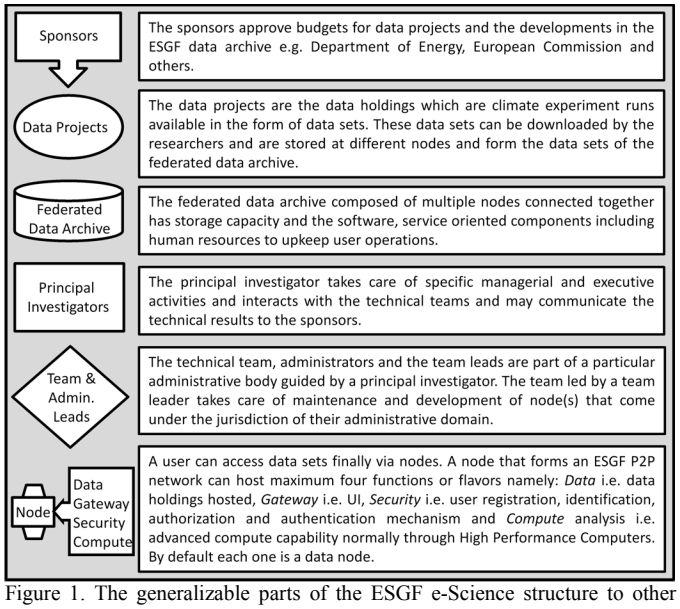

Figure 1. The generalizable parts of the ESGF e-Science structure to other ESGF.

[3] http://stackoverflow.com/
[4] https://askbot.com/

Furthermore, it was suggested to involve users in the user support process to get their opinion and continually improve the user support process. Additionally, a user support activity model has been suggested for the staffs of ESGF so that they can efficiently and effectively help the users of ESGF in such a way that their original tasks (other than user support) do not get affected. The application of social media has also been suggested to support users of e-Science infrastructures, particularly ESGF.

V. Conclusion and Future Work

This research assists in starting a new era of supporting the users of CI such as ESGF. In addition, it reduces the wastage of resources and redundancy of efforts by the staffs of e-Science infrastructures. Satisfaction of users will increase as the user support process gets smoother and painless. At the same time the shared understanding of human support agents to mitigate root causes of user problems will enhance. Studying the support process of the ESGF CI empirically is expected to reveal a good-practice user-support framework for e-Science infrastructures thus defining a generic support model that can be applied to CI of other domains too. This in turn can be a rich source for generating practical implications for system design, user modeling, and user instruction.

Once, the problems about the user support process in climate CIs have been identified, it might help in identifying similar problems in other fields. As an inter-disciplinary study, this research will contribute to IT organizations, Human Computer Interaction (HCI), Computer Supported Cooperative Work (CSCW) and e-Science. For IT organizations exploring the current support process in e-Science is useful for understanding the organizational structure of support staffs in e-science as well as their role in problem solving. For HCI, the exploration could provide a new system and User-Interface (UI), collaboration techniques involved in support staffs and end users communication in CI. In CSCW the exploration of the user support process of climate CIs at a micro-level will facilitate understanding the support process from a micro-perspective within the science—technology—sociality triangle, across different placements on the spectrum from development to theory and also across different orientations.

References

[1] M. Jirotka, C. P. Lee, and G. M. Olson, "Supporting Scientific Collaboration: Methods, Tools and Concepts," *Comput. Support. Coop. Work*, no. Ci, Jan. 2013.

[2] T. Hey and A. E. Trefethen, "Cyberinfrastructure for e-Science.," *Science*, vol. 308, no. 5723, pp. 817–21, May 2005.

[3] L. Cinquini and et al., "The Earth System Grid Federation: An open infrastructure for access to distributed geospatial data," in *2012 IEEE 8th E-Science Conf.*, 2012, pp. 1–10.

[4] Earth System Grid Federation, "ESGF," 2013. [Online]. Available: http://esgf.org/. [Accessed: 10-Jun-2013].

[5] Office of Government Commerce (OGC), *ITIL V3 2011 Service Operation*. 2011.

[6] M. Jäntti, A. Cater-Steel, A. Shrestha, and N. Campus, "Towards an Improved IT Service Desk System and Processes: A Case Study," pp. 203–215, 2012.

[7] M. Jäntti and J. Kalliokoski, "Identifying Knowledge Management Challenges in a Service Desk: A Case Study," *2010 Second Int. Conf. Information, Process. Knowl. Manag.*, pp. 100–105, Feb. 2010.

[8] B. C. Potgieter, J. H. Botha, and C. Lew, "Evidence that use of the ITIL framework is effective," in *18th Annual conference of the national advisory committee on computing qualifications,*2005, pp. 160–167.

[9] J. Graham and B. Hart, "Knowledge integration with a 24-hour help desk," in *Proceedings of the 28th annual ACM SIGUCCS conference on User services: Building the future*, 2000, pp. 92–95.

[10] M. Jäntti, "Improving IT Service Desk and Service Management Processes in Finnish Tax Administration: A Case Study on Service Engineering," pp. 218–232, 2012.

[11] O. Appleton, "A strategic roadmap for federated service management," 2014. [Online]. Available: http://www.fedsm.eu/sites/default/files/FedSM-D4.2-Qualification_Scheme -v1.1.pdf.

[12] R. Yin, "Case Study Research: Design and Methods," Sage Publishing, 1994.

[13] K. Eisenhardt, "Building theories from Case study Research," *Acad. Manag. Rev.*, vol. 14, no. 4, pp. 532–550, 1989.

[14] H. I. Chunpir, T. Ludwig, and A. A. Badewi, "Using Soft Systems Methodology (SSM) in understanding current user-support scenario in the Climate Science Domain of Cyber-infrastructures," in *HCII (10)*, 2014, pp. 495–506.

[15] Chunpir and Williams et al., "3rd Annual Earth System Grid Federation and Ultrascale Visualization Climate Data Analysis Tools Face-to-Face Meeting Report," 2013.

[16] H. I. Chunpir and T. Ludwig, "Reviewing the Governance Structure of End-User Support in e-Science Infrastructures," in *Informatik 2014, Beyond Data workshop proceedings*, 2014.

[17] H. I. Chunpir, A. Badewi, and T. Ludwig, "User Support System in the Complex Environment," in *HCI(11)*, 2014, pp. 392–402.