User Support System in the Complex Environment

Hashim Iqbal Chunpir[1,2], Amgad Ali Badewi[3], and Thomas Ludwig[1,2]

[1] German Climate Computing Center (DKRZ), Bundesstr. 45a, Hamburg, Germany
{chunpir,ludwig}@dkrz.de
[2] University of Hamburg, Department of Informatics, Hamburg, Germany
[3] Cranfield University, Manufacturing Department, Bedforshire, United Kingdom
a.badewi@cranfield.ac.uk

Abstract. e-Science infrastructures have changed the process of research. Researchers can now access distributed data around the globe with the help of e-infrastructures. This is particularly a very important development for the developing countries. User support services play an important role to provide researchers with the required information needs to accomplish their research goals with the help of e-infrastructures. However, the current user-support practices in e-infrastructures in the climate domain are being followed on intuitive basis, hence over-burdening infrastructure development staffs who partly act as human support agents. The main contribution of this paper is to present the environmental complexity with-in the contemporary user support practices of climate science e-infrastructure known as Earth System Grid Federation (ESGF). ESGF is a leading distributed peer-to-peer (P2P) data-grid system in Earth System Modelling (ESM) having around 25000 users distributed all over the world.

Keywords: e-Science, systems, research, user support, help desk, developing countries.

1 Introduction

User support has always been a key topic in Human Computer Interaction (HCI) and other fields as it refers to assistance provided to the users of technology and other products. The user-support process in e-Science infrastructures is an operational process that serves the end-users of e-Science infrastructure in achieving their goals i.e. using e-Science infrastructure (mainly) for their research but not necessarily limited to that. Users are mainly researchers and they accomplish various tasks of research within a specific time-frame via e-Science infrastructures. The user-requests are the inputs to this process and are processed by e-Science infrastructure staffs (also known as user support employees) with the help of tools and methods (whether automated or manual) to service the incoming user-requests, thus meeting the user support needs. However, support staffs in climate e-Science infrastructures, have also other tasks to be done, for instance; programming, strategic planning, node administration and others, apart from servicing end-user requests. This process of operating, maintaining and further expanding the infrastructure, including its data, is iterative in nature as the nature of support in e-Science infrastructure projects. The user support is

A. Marcus (Ed.): DUXU 2014, Part IV, LNCS 8520, pp. 392–402, 2014.

offered in the form of self-help via support websites, online tutorials, wikis or contacting an expert in the form of traditional help-desk [1], [2] and service-desk [3–5].

In this paper, the environmental complexity within the contemporary user support practices of a well-known climate science e-Science infrastructure known as Earth System Grid Federation (ESGF) is presented. The paper then presents the critique of the current user support process in ESGF and finally emphasizes on the need to streamline user-support in e-infrastructures. Moreover, the paper puts forward a recommendation to involve human resources as human support agents from research institutes in the developing countries as a part of remedy to help user support activities in ESGF. The rest of the paper is organized as follows: Section 2 describes the work related to e-Science and user support discipline. Before describing the organization of ESGF and its effects on the user support process in section 4; research methodology is explained in section 3. Finally, section 5 describes the critique of the current user support, followed by conclusion and discussion in section 6 and 7.

2 Related Work

The related work in this paper can be distributed into three sub-sections: e-science, user support, and user support in e-Science discipline.

2.1 e-Science

e-Science infrastructures have been widely deployed to access and share the knowledge, data, computing resources and even human resources to facilitate intra and inter-disciplinary research [6], [7]. There are different names associated to the concept of e-Science. The same concept is popularly known as “e-Science” in Europe and “cyber-infrastructures” in the US. Other names include e-Research, collaboratories, virtual science and Big Data Science [7].

In e-Science infrastructures, as more and more effort is being invested in improving the grid-based technologies like anatomy of data-grid [8], development of middleware, storage of data in grid [7] and socio-structural aspects of e-Science for instance “Virtual Organisations” (VOs), CWE (Collaborative Work Environments), VRE (Virtual Research Environments) [9]; the development in user support is being offered on intuitive basis with a focus on technology oriented methods that dominate the field [10–13]. The organization of user support is mainly based on the past experience, without studying and exploring the factors such as; nature of e-Science infrastructural domains, data application, scientific concerns, consideration of end-user and support staff requirements [14].

2.2 User Support (a.k.a. Help Desk)

User support in IT industry is known to be started in early 1980’s with the first “help desk” (HD) that had only a desk, a pen and a telephone used by human support agent [1], [15]. User support allows users to contact support staffs to address particular problem of a user [5]. Service desk is the concept that combined service management studies with the traditional customer support studies that used the term "help-desk" [12]. Until now there are different versions of business service frameworks such as IT Information Library (ITIL) that provides "best practice guidelines" for servicing end-users and customers especially in the commercial corporations and companies locally and globally [16], [17]. Some of these frameworks have been modified and adapted to academic setups such as universities [10], [17] or to governmental administrative bodies [12]. However, these frameworks have yet not been applied to the field of e-Science infrastructures and few studies address the issue of improving user support process of supporting users, keeping economic and human resource factors under control as well as fulfilling the expectations of all the stakeholders. Studying ESGF user support as a use case will contribute to the "service desk" or "customer services" concept in distributed, research oriented, non-commercial environments.

2.3 Defining User Support in e-Science Infrastructures

After examining the notion of user support and e-Science infrastructure individually, the user-support process in e-Science infrastructure can be defined as: The user-support process in e-Science is an operational process that serves the end-users of e-Science infrastructure in achieving their goals. The user-requests initiated by users are the inputs to this process and these user-requests are processed or transformed by user support staff with the help of tools and methods to provide solutions. Hence, meeting the user support needs. This process is iterative in nature as the nature of support in e-Science projects. The end-user support process is an example of a process. The environment is e-Science infrastructures. The mission of this process is servicing and satisfying end-users (of e-Science infrastructures) incoming queries. The constraints are user-support times, support staffs, financial resources, supporting technologies and interfaces. The resources are support staffs team and the support tools. The inputs are user requests / problems. The transformation is to understand the user-requests and provide a solution. The product is end-user support framework that provides solutions to the users.

Although lots of work has already been invested in the e-Science projects to form and operate a working user-support in e-Science infrastructures, yet there is a need to standardize and systematize user support. This study, through the first empirical investigation of end-user support process in e-Science infrastructure of climate domain, is aimed to fill this knowledge gap by providing a framework for understanding practices in e-Science in general and in climate science in particular. In this paper the elements of organizational structure of a climate e-Science infrastructure, ESGF, is presented.

3 Research Methodology

Case study research is used to depict the current user support practices in e-Science infrastructures [18]. Data collection methods are participatory observation, interviewing, archival analysis, and survey. Triangulating the result is useful to understand the current processes from different angles.

10 Interviews, from ESGF and C3Grid[1] e-Science infrastructures having different backgrounds and roles, are used to explore the current state and to understand the potential weakness in the current system. Participatory observation comes to achieve more understanding to the current processes.

Finally, questionnaire is used to describe the current operations of the support system. Online questionnaire was the method chosen among other methods for this survey since it is quicker, automated and supports complete anonymity of participants [18]. Questionnaire consists of 43 questions. Out of 36 responses received only 25 responses were useful. Despite small sample size of the respondents, it is a significant proportion of the whole targeted population in this case.

4 Organization and Governance of ESGF

In ESGF, software development and project management are done by different institutions that are project partners e.g. LLNL2 (leading partner), BADC3, DKRZ4, ANU5 and others. The model of an organizational structure of ESGF is shown in figure 1. In the figure, each symbol depicts an entity, having a specific function which is part of a climate e-Science infrastructure organization. The number of each entity may vary from time to time within a particular e-Science infrastructure and from infrastructure to infrastructure in e-Science, thus, creating a dynamic and complex environment. The institutions form the executive part of ESGF headed by principal investigators, technical team leads and the technical development team. An entity representing principal investigator is represented as a rectangle in the figure 1. Similarly, an entity representing a technical team lead including the technical team is represented as a diamond shaped notation in the figure 1. Each of the principal investigators may head different technical teams. There can be many principal investigators from different continents carrying diverse tasks. The executive part of ESGF is responsible for setting the strategic direction and overseeing technical activities of the project. The ESGF technical and maintenance teams are known as administrative bodies. Administrative bodies are a bit different to the concept of administrative domains6 in such a way that the former includes human resources. Administrative bodies are distributed in Asia, Australia, North America and Europe. Moreover, the forthcoming teams from South America and Africa will be joining soon. Each of these administrative bodies manages one or more nodes. In order to form a distributed

1 Climate Collaborative Community data and processing grid project.
2 Lawrence Livermore National Laboratory, USA.
3 British Atmospheric Data Center, UK.
4 German Climate Computing Center, Germany.
5 Australian National University, Australia.
6 A collection of hardware such as computers, databases, networks and other instruments under a common administration thus sharing common policies.

control, the nodes collaborate with other nodes to form a peer-to-peer (P2P) system not only technically but also socially, institutionally and administratively.

The sponsors are represented at the top of the figure 1. In ESGF, sponsors include DOE[7] (main funder), European Commission (EC) and others. For the ESGF operations there are tens of data centers around the globe that are working on data projects represented in the form of an oval in figure 1. The data projects vary in their nature and thematic area within the climate science domain. These data projects are known as data holdings. It is expected that in future the ESGF data archive system will serve the data holdings from the domains other than climate science as well [19]. CMIP5[8] is an example of one of the main projects of ESGF being served by ESGF data archive system represented in the form of a cylinder, in the figure 1. From CMIP5, data has been used to generate IPCC AR5[9] report on the basis of which political decisions are made. Looking at the history of ESGF, one can predict that in future the data holdings will keep on increasing. If users have specific queries about these data holdings then climate specialists are needed to be consulted. This implies that users need to have contact details of persons who can provide guidelines about scientific questions. Entertaining scientific queries is currently not an explicit part of the user support system in ESGF.

A node (represented in the figure 1, at the bottom) may have four different roles or flavors within itself as stated by Cinquini et al. [20]. These flavors are shown by the words "Data, Gateway, Security and Compute" attached to each node (see bottom part of the figure 1). The nodes including their respective flavors are managed by a particular team in an administrative body at a particular location in the world. "Data" stands for data holdings hosted by a node (also known as data node). A number attached the word "Data", for instance (or any of the four words mentioned before) represents the number of data nodes being hosted and maintained by a single administrative body. For example, if there is a label "3Data" it means that there are three data nodes which are managed by a particular administrative body. Every node has at least a data holding(s) part or flavor in it, by default. Additionally, a node may have a gateway "Gateway", a security set-up "Security" and a compute facility "Compute" as shown in the figure 1. A gateway is responsible for representing the data sets available in ESGF system to a user via user interface (UI) to interact with the system.

In ESGF terminology a security part of a node is also known as an identity provider "IdP". Security part of a node is responsible for registration, identification and authorization of a user to ensure that a user is a valid entity who is entitled to access data sets available in the ESGF system. A user accesses a data-set(s) hosted by the ESGF system via its UI "Gateway" from any ESGF node via single-sign on (SSO)[10]. The compute part of node is responsible for computing and visualization in ESGF via High Performance Computers (HPC) in a particular administrative body. The subsystems or system components in ESGF are distributed worldwide. The behaviour of

---

[7] Department Of Energy, USA.

[8] Coupled Model Intercomparison Project – phase 5.

[9] Intergovernmental Panel on Climate Change-Fifth Assessment Report.

[10] SSO facilitates a user to access multiple nodes without providing identification keys (e.g. a certificate) to each node separately.

each component effect each other and the user support system as well. For instance; the network topology of the ESGF P2P network directly influences its management and administration structure.

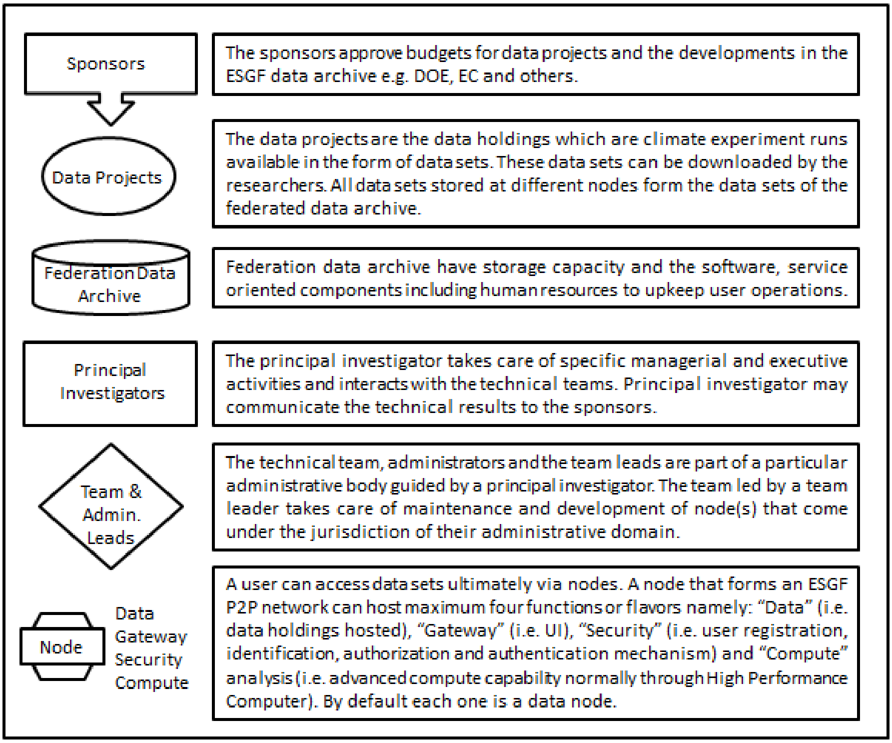

Fig. 1. e-Science infrastructure organization model in climate science domain and its elements

ESGF is stated as: "The Earth System Grid Federation (ESGF) is a multi-agency, international collaboration of people and institutions working together to build an open source software infrastructure for the management and analysis of Earth Science data on a global scale." [21] However, in this ESGF definition, user support services are not explicitly made part of the definition of ESGF. Keeping the anatomy of ESGF P2P data archive system in view, the user support system of ESGF has its sub-units. Though these sub-units are not formally designated as support units within administrative bodies, the support units are implicitly part of administrative bodies. Consequently, from the geographically distributed organization of ESGF P2P network, it is understandable that each and every administrative domain has its own practices of handling user-requests. This observation is also evident from the qualitative cum quantitative inquiry into the user support practices of ESGF undertaken by the authors. The diversity of practices in handling user queries by the user support staff, who themselves are developers of the ESGF system, form different support structures and models in each of the administrative domains, hence making a heterogeneous user support process. Subsequently, the user support system in ESGF does not comply with any set standards of processing user support requests.

From the figure 1, one can anticipate that the numbers of administrative bodies (principal investigation institutions plus developing teams of ESGF), data holdings, ESGF users, ESGF staffs participating in ESGF P2P system are subject to increase. Additionally, the role of an administrative body and its attached node(s) is subject to change, therefore this whole ESGF set-up is a complex, dynamic and evolving in nature. In ESGF system there is a continuous architecture re-design activities, software development, hardware changes, data publishing, data curation, data quality check and other activities. Attached to these core operational activities is the necessity of the user support activities that cannot be ignored. A dynamic and an ever-evolving infrastructure need a dynamic user support "service desk." Therefore, e-Science is likely to be confronted with demanding issues of long-term and continuity of service, particularly related to user support services which is quite similar to data curation and software development.

5 Critique

From the survey, the current user support model and organization of the current support process in climate e-Science infrastructure (ESGF) is not uniform. Therefore, it is vital to decide between centralized and de-centralized user support model. There are advantages and disadvantages associated to both of the models; however in user-support in general, centralized model is preferred because of single point of contact (SPOC) [22]. It is important to bring uniformity in the current user support process in climate e-Science infrastructure (ESGF). It is interesting to note that due to distributed nature of e-Science, the user support model is also distributed.

Very few (18%) support structures in administrative bodies rely on e-support or self-help. e-support is not much in practice in climate e-Science infrastructures. The reason for this is partly the distributed information on the websites as well as in the Wikis which is not completely known to users. On top of this, the information is not updated regularly. e-support should be promoted in climate e-Science projects. A practice should be followed that the users before writing to the help desk, try to get e-support by looking at the associated web sites and Wikis at first, to locate the relevant information.

The update of user-support information on the user support portal(s) should be made regular. A Service Level Agreement (SLA) can be suggested, where the concerned support staff updates the information regularly. The information can be checked by other support staff or a usability expert to make sure that these communication materials are accessible, easy to grasp and are understandable. Since there is no standardized classification into levels of escalation such as Second Level Support (SLS) and First Level Support (FLS); it is important to develop a norm. A support unit must stick to a particular global norm or a standard that may be introduced in the support structure of climate science e-Science infrastructures.

An effort could be made to reduce the time of personal response by introducing FLS employees covering various time-zones. For example, if a user has forwarded a request in Europe then it is vital to reply him/her from a support staff working within European time-zones. The greater the time of personal response, the lesser would be the efficiency of support process and the lesser would be the satisfaction of users. An flexible SLA or a quasi-SLA can be introduced in this case where up to 1 (working) day at maximum could be set as a standard i.e. the maximum time for personal reply. In case a concerned support staff is not available (e.g. on holidays etc.) then ersatz or substitute support staff should be available to service the user-requests in each support unit. This is very much appropriate in terms of showing empathy (building customer confidence in the support process) to the end-user because end-user might get frustrated if s/he does not get a personal reply.

Majority of ESGF employees who do user support (almost 80%) do not use automatic reply mechanism. It is important (in case of a request tracking system) to use automatic reply mechanism. The automatic reply should preferably look like a personalized message and not a machine response. The support process will improve if the reply time between support staff (e.g. between first and second level support staff) can be reduced. This can be achieved by an agreement or introducing a SLA for delegated response. A standard is needed to be defined and the support staffs who receive the delegated user-request might follow the SLA or an agreement. This will improve a delegated response time.

It is important to get the suggestions from users and encourage their participation for a number of reasons: Firstly, their feedback can improve the support process. Secondly, one can get an insight into their level of satisfaction. Thirdly, they may rate the solution suggested by the support staff. Finally, since the number of users is not so high, a personal interaction may create mutual trust, empathy and better relationship amongst end-users and the support staff.

Most of the support staffs (80%) do not collect user statistics and user request statistics. It is important to collect user-statistics to know:

- The users-base (coming from different continents).
- To be able to measure the overall user satisfaction.
- It is important to have statistics in order to measure total number of incoming requests versus resolved requests.
- It is important to try different suggestions to solve similar problem to find out a better solution etc. And to archive these solutions.

The first author has experienced from his participatory observation and meetings with other support staff that, at the moment in e-Science projects, in the support process both “Mailing Lists” (ML) and “Request Tracking tools” (RT) are used in parallel. According to ESGF support staff they need mix of mailing list’s features and request tracking tool’s features, mostly advantages of both. Therefore, in e-Science support process it is important to decide whether to use either RT or ML or a combination of both. An investigation should be done in this area to get a further opinion of the e-Science support staff and end-users to find out that which communication media is viable in e-Science projects of climate science.

It is important to note that there is no involvement of pure domain scientists in the user support process. It is significant to have a collaboration and exchange or sharing of knowledge between technical experts and pure scientific experts (also known as domain experts). Therefore, there is a need to incorporate scientific experts.

Since there are no dedicated user support employees, the employees of ESGF perform other activities parallel to supporting users. These activities affect support activity or vice versa, therefore it is important to have a dedicated support staff, preferably not as advanced as the current one. For instance, in ESGF top computer scientist are handling the simple and routine user queries. Simple and routine user queries may only be handled by FLS. People from developing countries may be employed to handle simple and routine queries. For the SLS, a norm can be created where a standard could be set in future where a support staff have to plan or reserve part of his/her time as a quota for user support activity. Some priority, recognition or value should be assigned to the user-support time in a working day where one can give as much or more credit to the support activity performed by the user-support staff as other activities that the support staff may perform in parallel. It is important to keep support activity productive and alive. It must be ensured by the support staff that all user queries are addressed properly.

6 Discussion

From the organization of ESGF e-Science infrastructure and the survey results one can say that the staffs who are involved in user-support process are over qualified to undertake first-level support activities which can be categorized as simple and routine. The employees of ESGF may rather be used as specialist for SLS or more levels of user support escalations. The (user-support) process owners who are active human support agents in different parts of the world may be nominated by the executive committee of ESGF from key institutes may recruit and educate the user support staff that may be engaged at the first line of user support as FLS preferably from climate computing institutes or universities in developing countries. The service provided by human support agents will not only be economical but also good for developing countries to not only promote the use of cyber-infrastructure of a particular in the local research arena of a developing country but also to introduce it in university curricula within developing countries. The training of the FLS staff from training countries may be done by the key process owners of the user staff by online sessions and development of instructive guidelines. As there is an upsurge in user queries with the passage of time due to inclusion of new data projects and increase of users of ESGF, it is important to start an initiative at this stage.

The technical and intellectual capacities in the developing countries are immense as already demonstrated by some of high-tech initiatives started there. For instance, formed more than a decade ago, National Database and Registration Authority (NADRA) is a high-tech initiative in Pakistan that proves advancement in IT sector and capability to develop high-tech solutions [23]. Moreover, for the people who reside in a particular region (urban or rural) in developing countries, an adaptive user interface (UI) can be designed that would fit the cultural background, language and interface expectation. Doing so may make it easier for them to get access to data-sets offered by a e-Science infrastructure as well as to get self-help or e-support. A UI can be created by the group of people who will be local human support agents after doing research. Furthermore, the time zones can be covered too, by allocating FLS at different time-zones. Outsourcing of parts of software of an e-Science infrastructure may be engineered in some developing countries economically and effectively.

7 Conclusion

In this study, it is evident from the observations that the user support within e-Science infrastructures in climate science is not being paid attention to and use support needs to be redefined within a complex and dynamic nature of e-Science infrastructures. In climate e-Science infrastructure there is no position of dedicated user support manager. The staffs of e-Science infrastructure are doing other activities apart from user support activities. The employees of e-Science infrastructure in climate domain, who are top computer scientists, are handling simple and routine user enquires. Handling of simple and routine user enquires might be economical to transfer to the institutes in developing countries care of by the staffs of e-Science infrastructure; once the e-Science support process in climate science and other domains is systematized. Currently, the authors are working on a conceptual model to systematize and standardize e-Science use support system in the climate domain. In this paper, a generalized organization structure of e-Science infrastructure, at least in climate domain but not limited to it, is presented. In the future, the authors will observe the effectiveness of transferring front line user support units to various institutes in developing countries.

Acknowledgement. We appreciate the sincere support of DKRZ and ESGF colleagues Dean Williams, Stephan Kindermann and others, including users who took part in our survey.

References

1. Leung, N., Lau, S.: Information technology help desk survey: To identify the classification of simple and routine enquiries. J. Comput. Inf. Syst. (2007)
2. Leung, N.K.Y.: University of Wollongong Thesis Collection Turning user into first level support in help desk: Development of web-based user self-help knowledge management system (2006)
3. Jäntti, M.: Improving IT Service Desk and Service Management Processes in Finnish Tax Administration: A Case Study on Service Engineering. pp. 218–232 (2012)
4. Jäntti, M.: Lessons Learnt from the Improvement of Customer Support Processes: A Case Study on Incident Management. In: Bomarius, F., Oivo, M., Jaring, P., Abrahamsson, P. (eds.) PROFES 2009. LNBIP, vol. 32, pp. 317–331. Springer, Heidelberg (2009)
5. Jäntti, M.: Examining Challenges in IT Service Desk System and Processes: A Case Study (c), 105–108 (2012)
6. Hey, A., Trefethen, A.: The data deluge: An e-science perspective. pp. 1–17 (January 2003)

7. Hey, T., Trefethen, A.E.: Cyberinfrastructure for e-Science. Science 308(5723), 817–821 (2005)
8. Buyya, R., Venugopal, S.: A Gentle Introduction to Grid Computing and Technologies (July 2005)
9. Jirotka, M., Lee, C.P., Olson, G.M.: Supporting Scientific Collaboration: Methods, Tools and Concepts. Comput. Support. Coop. Work (Ci) (January 2013)
10. Graham, J., Hart, B.: Knowledgebase Integration with a 24-hour Help Desk (2000)
11. Leung, N., Lau, S.: Relieving the overloaded help desk: A knowledge management approach. Int. Inf. Manag. ... 6(2), 87–98 (2006)
12. Jäntti, M., Kalliokoski, J.: Identifying Knowledge Management Challenges in a Service Desk: A Case Study. In: 2010 Second Int. Conf. Information, Process. Knowl. Manag., pp. 100–105 (Febraury 2010)
13. Roth-Berghofer, T.: Learning from HOMER, a case-based help desk support system. Adv. Learn. Softw. Organ., pp. 88–97 (2004)
14. Soehner, C., Steeves, C., Ward, J.: E-Science and Data Support Services (August 2010)
15. Kendall, H.: Prehistoric Help Desk!! Support World. Help Desk Institute, pp. 6–8 (October-November 2002)
16. Potgieter, B.C.: Evidence that use of the ITIL framework is effective (2002)
17. Arora, A.: IT Service Desk Process Improvement – A Narrative Style Case Study (2006)
18. Lazar, J., Feng, J., Hochheiser, H.: Research Methods in Human-Computer Interaction. Wiley, Indianapolis (2010)
19. Williams, D.N., Bell, G., Cinquini, L., Fox, P., Harney, J., Goldstone, R.: Earth System Grid Federation: Federated and Integrated Climate Data from Multiple Sources. vol. 6, pp. 61–77 (2013)
20. Cinquini, L., Crichton, D., Mattmann, C., Harney, J., Shipman, G., Wang, F., Ananthakrishnan, R., Miller, N., Denvil, S., Morgan, M., Pobre, Z., Bell, G.M., Drach, B., Williams, D., Kershaw, P., Pascoe, S., Gonzalez, E., Fiore, S., Schweitzer, R.: The Earth System Grid Federation: An open infrastructure for access to distributed geospatial data. In: 2012 IEEE 8th International Conference on E-Science, pp. 1–10 (2012)
21. Williams, D.N.: Earth System Grid Federation (ESGF): Future and Governance World Climate Research Programme (WCRP), Working Group on Coupled Modelling (WGCM)— Stakeholders and ESGF. pp. 1–17 (2012)
22. Middleton, I.: The Evolution of the IT Help Desk: From Crisis Centre to Business Manager in the Public and Private Sectors. MSC Thesis. The Robert Gordon University. Aberdeen, UK (1999)
23. Gelb, A., Decker, C.: Cash at Your Fingertips: Biometric Technology for Transfers in Developing Countries. Rev. Policy Res. 29(1), 91–117 (2012)