Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014, Kraków, Poland 19-23 July 2014
Edited by H.I. Chunpir, T. Ludwig and A. Badewi

A snap-shot of user support services in Earth System Grid Federation (ESGF): a use case of climate cyber-infrastructures

Hashim Iqbal Chunpir[12], Thomas Ludwig[12] and Amgad Badewi[3]

[1]German Climate Computing Center (DKRZ),
Bundesstr. 45a, 20146,
Hamburg, Germany

[2]Faculty of Informatics,
University of Hamburg,
Hamburg, Germany

[3]Department of Manufacturing,
Cranfield University,
Bedfordshire, England

ABSTRACT

Cyber-infrastructures have changed the process of research. Researchers can now access distributed data from all parts of the world with the help of cyber-infrastructures. User support services play an important role to facilitate researchers to accomplish their research goals with the help of cyber-infrastructures. However, the current user-support practices in cyber-infrastructures are being followed on intuitive basis (at least in climate e-infrastructures) thus over-burdening cyber-infrastructure employees. The main contribution of this paper is to present the snap-shot of the current user support practices in a cyber-infrastructure of a climate science 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 2700 users distributed worldwide. The questionnaire conducted with the climate cyber-infrastructure projects' employees presents the picture of the current user support situation by highlighting their profile, utilization of various communication media, user-request service time, attributes of incoming user problems and information requests. The respondents of the questionnaire were 26 support staffs of cyber-infrastructure projects, from different parts of the world. 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 cyber-infrastructures.

Keywords: Cyber-infrastructures, Climate Science Data-driven Research Projects, User Support, Help Desk, Service Desk, Human Systems Integration, Service Engineering, Globalization, Service as a Process

INTRODUCTION AND BACKGROUND

Services are everywhere in the world. They cover large variety of businesses types and fields including services in the scientific domains. Call center, user support, help desk are well established concepts in the service industry, where volumes of customer/client or user requests are processed daily, thus providing help to the requestors. The contemporary user support is offered, in the form of self-help via support websites, online tutorials, wikis; or contacting an expert via traditional help-desk (N. K. Y. Leung, 2006; N. Leung & Lau, 2007) and via service-desk (Jäntti, 2009, 2012a, 2012b) in commercial and non-commercial sectors. Cyber-infrastructures are deployed to access and share the data and computing resources mostly for the research purposes. They are also referred as e-Research, e-Science, e-Science infrastructures and collaboratories (Hey & Trefethen, 2005). Users of cyber-infrastructures use them to access data for their personal research goals. The user-support process in cyber-infrastructures helps researchers to mitigate all (technical but not limited to that) problems that might arise during the interaction of a user with the cyber-infrastructure. Users are mainly researchers and they need to accomplish various tasks of research within a specific time-frame via accessing and downloading required data stored at the various locations (nodes) of cyber-infrastructure, distributed globally. The user-requests are the inputs to the user support process and are processed by the staffs (also known as user support employees) of a cyber-infrastructure with the help of tools and methods (whether automated or manual) to service the incoming user-requests, thus meeting the user support needs. In cyber-infrastructure user support is provided; however, it needs more attention as well as investments (Soehner, Steeves, & Ward, 2010).

The user-support process in cyber-infrastructures is directly affected and is evolving with the evolution of cyber-infrastructures. It is vital to test and be aware of the user support procedures currently in practice, which are most of the times taken care by the infrastructure developers themselves. It is noteworthy to observe by the authors that the evolution of cyber-infrastructure technologies is partly dependent upon the communication taking place within the support staffs and users. The support staffs in their joint collaboration with each other as well as their interaction with the users discover new features that can be introduced in cyber-infrastructures while handling the user related issues. Therefore, it is important to analyze the support process and enhance its productivity as well as its efficiency in order to save time and supporting efforts of the cyber-infrastructure employees. This is possible if the current situation of the user support process in cyber-infrastructures is observed. In this study, information gathered empirically, by analyzing the methods, tools and techniques employed by cyber-infrastructure staffs (i.e. staffs who also are doing user support activities) via a survey questionnaire presents the part of the picture of the contemporary user support practices in Earth System Grid Federation (ESGF), at least from the employees' perspective. The outcomes of this study will enable future cyber-infrastructure projects to organize their user support process in a better way. This study is an important milestone in the field of cyber-infrastructure user services for the reason that it will be laying a foundation of a user support recommendation framework or good practices in cyber-infrastructures user support; that can be suggested after studying ESGF user support system. This user support recommendation framework can then be used to standardize other similar cyber-infrastructures with in the climate domain or other domains. Currently, the authors are working on it.

In this paper, the user support of a well-established climate cyber-infrastructure ESGF is taken as a case study. Moreover, the focus of this paper is on the outcomes of the survey-questionnaire, delineated in the form of a snap-shot of the current user support services in ESGF. It is important to note that, cyber-infrastructure staffs who support users have also other tasks to be done at the same time, for instance; programming, strategic planning, node administration and others, apart from servicing end-user requests. Therefore, it is pertinent to capture their thinking about the current user support practices in ESGF and other associated projects of climate e-Science. The rest of the paper is organized as follows: Section 2 describes the work related to cyber-infrastructure and user support concepts. Before depicting the snapshot of the current processes in ESGF 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 5 and 6, respectively.

RESEARCH METHODOLOGY

The snap shot of the current user support practices in cyber-infrastructures was captured in this study via survey-questionnaire and via participatory observation. The goal of the survey questionnaire was to investigate the current user-support process in use in main projects of cyber-infrastructures (related to climate science). There are different methods available to conduct survey for instance: by-mail (post), telephone etc. However, online questionnaire was the method chosen for this survey because it is quicker, automated and supports complete anonymity of participants (Lazar, Feng, & Hochheiser, 2010).

As this research focusses on improvement of the user support process in climate cyber-infrastructures, therefore conducting survey with the staffs of climate cyber-infrastructures is a vital step. The reason of conducting a survey with the staffs of cyber-infrastructures as the "target population" (Couper, 2000), is because these staffs though maintaining and further developing cyber-infrastructure are also partly are involved in supporting users. Therefore, these staffs can be partially called user support staffs. The questionnaire was divided into 5 different sections to get information about the following: Structure of support, time related factors in user-support process, communication factors in user-support process, user requests related factors in user-support process, users related factors in user-support process and collect data relating to sample description and introducing basic social biographic data with self-rating about their expertise.

LIMITATION OF THIS STUDY

The outcome of this survey questionnaire hinge on the responses of the cyber-infrastructure staffs that support users. There is a danger that these responses might be based on estimation of the staffs, who were the respondents in this survey-questionnaire, which might not reflect the actual picture of the current support process. There is little or no documentation available about the current support process in practice to verify the responses. However, keeping with the data triangulation scheme, the quantitative and qualitative data collected from this survey will later be not only tallied but also verified with different sources. These other sources include; interviews conducted with the support staff, users and other stakeholders, the mailing-list communication between the end-users and the support staffs. Finally, conducting focused group validation. Despite small sample size of the respondents, the data collected is still useful and the results are significant. Due to small sample size detailed statistical analysis is not shown in this paper; only results from descriptive statistical analysis are presented. Similarly, some of the basic biographic data is also not presented and discussed in this paper. During the time of writing of this paper, the results of survey-questionnaire are complete. However, results of interviews and archival analysis are yet to be compiled.

PORTRAYING THE SNAP-SHOT OF THE CURRENT USER SUPPORT PRACTICES

This empirical qualitative cum quantitative investigation revealed number of issues where attention should be paid to and the situations are needed to be improved in the current user-support process in the climate cyber-infrastructure projects. These issues include allocation of time, human resources, response time, time to solve the user-problems, characteristics of user requests, support tools, support structure and others.

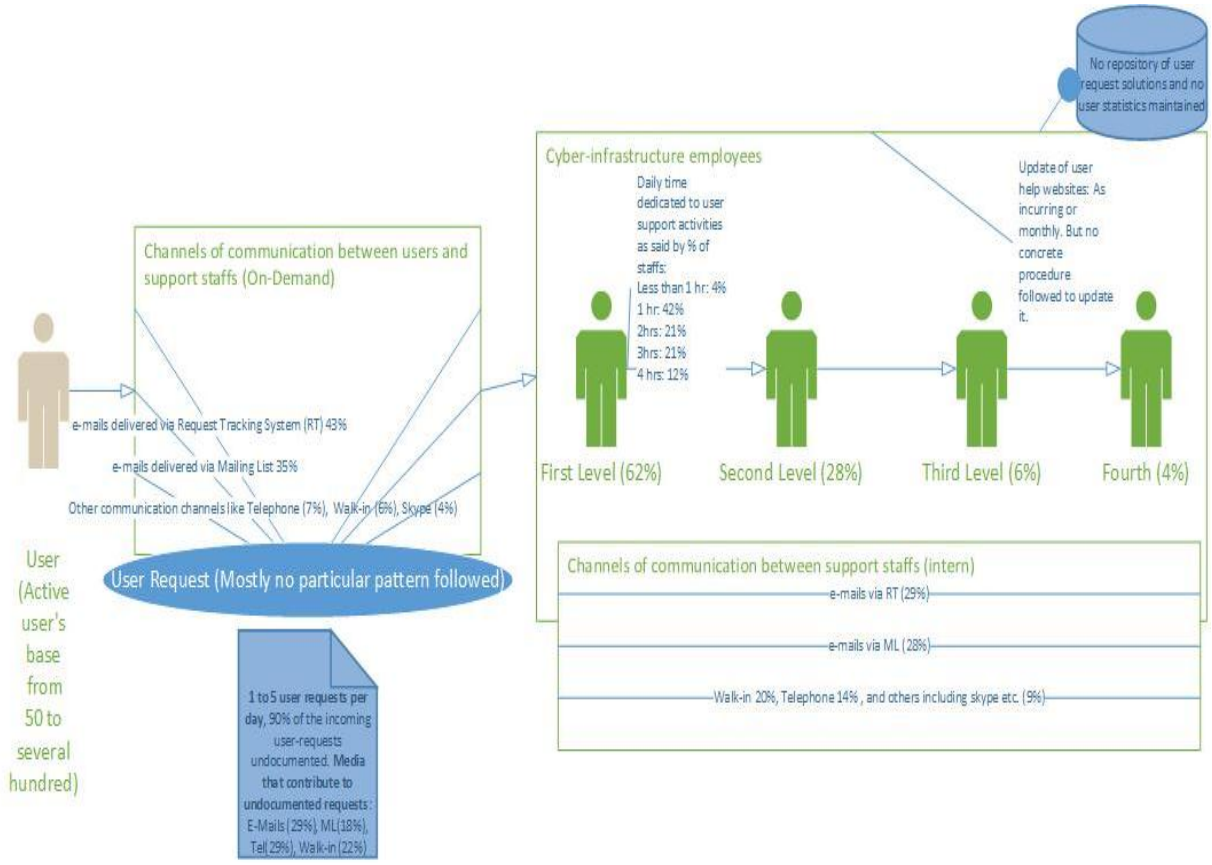

Figure 1 The current user support scenario in climate cyber-infrastructure projects.

In the survey questionnaire, mostly ESGF employees (17 out of 26) took part as respondents. Other respondents included the staff from C3Grid which is connected to the ESGF infrastructure as both of the project host common climate data projects. From the survey and participant observations one can see that male members are part of the cyber-infrastructure team as IT is dominated by male members, we need to encourage women to take this position. All support staffs have technical as well as scientific knowledge and all staffs have at least Bachelors as the minimum qualification and PhD as the maximum qualification. This indicates that the respondents are well-qualified and highly skilled. On top of it the employees are also sufficiently experienced with the support activities.

The current staffs that do user support in climate cyber-infrastructure projects like ESGF and others are dominated by the technical experts followed by some staffs who also have some climate domain knowledge. However, it is prominent from the survey questionnaire that the employees dedicate at least from 1 hour (as answered by 42% of the participants) up to maximum 4 hours (as answered by 12% of the participants) of their daily work time to user support activities. 21% of the participants dedicate 2 hours and 3 hours each from their daily work to user support activities (see figure 1).

It is an important observation that the participants are not only doing support activities, but also doing other activities, such as software development, infrastructure administration, research activities and other non-support activities. The following activities are popular amongst the support staff: Software development (as stated by 68% of the respondents), writing documents (68%), which are the major activities i.e. most popular activities amongst respondents which occupy their time. Other activities include system administration (60%), research work (44%), administrative activities (40%) and data related activities (40%). In ESGF, there are no dedicated user support employees i.e. an exclusive user support team hired for user support, thus performing user support activities only.

User support model is diverse and distributed. This observation is in accordance to the structure of an organization, which is distributed globally. Moreover, the user support type in climate cyber-infrastructures is mainly on-demand that is why time zones where users live are important. For example, for the users who are in Europe, the support people with in the European time-zone are relevant and the support people in USA are in a different time-zones. It is

important to take care of the time-zones while planning a timely and efficient user support process.

In user support process of climate cyber-infrastructures, there are no levels of a user query escalation formally defined. Therefore, the role of first level staff (FLS)[1] and second level staff (SLS)[2] is not hardcoded in support process of ESGF-like climate cyber-infrastructures. In the current climate cyber-infrastructure user support framework, taking the global picture, there is no definite classification in N-tiers i.e. FLS and SLS. In some support units there is a practice of two-tier user support system (41%) whilst in others there is one-tier (41%) support structure. The rest of the support units are not organized into any tiers Thus in ESGF user support system FLS and SLS both exist in parallel. However, sometimes SLS (Second Level Support) is getting questions directly. So there is no complete distinction between one or two tier support structures as in other support infrastructures in other IT domains.

More than half (62%) of the incoming user requests are treated by the staffs of climate cyber-infrastructure at first level i.e. without forwarding it to a second person. However, 28% of the incoming user requests are handled by two levels of staffs. This means that staffs who receives the request at first, refers the request to another colleague; for instance; for further treatment. This phenomenon is known as second level escalation. Similarly there are three and four level escalations with 6% and 4% of incoming queries each in the current user support process. From this result one can say that the incoming queries (62%) treated by single staff may contain *simple and routine*[3] queries, similar to the one observed by Leung and Lau (N. Leung & Lau, 2006). Only few users use a particular pattern of a request. If a user-request is treated by many support staffs i.e. many support levels, this shows the total number of human resources allocated to solve a user-request thus the treatment of user-requests might get expensive.

Most of the respondents (70%) claim that they need up to 3 hours to reply to the user-requests on the average. So, response time works fine. Almost 43% of the incoming user-requests (user-problems) can be solved within an hour. These incoming user-requests that are solvable within an hour may be classified as simple problems (with respect to solution time). Previous research by (Knapp & Woch, 2002; N. Leung & Lau, 2007) has confirmed that a majority of the incoming queries are simple and routine. Solution time of user request works fine in the current user support system.

Most of the support staffs (68%) believe that it might take a day or up to 2 days to get a reply after delegating it to a peer support staff. It is vital in the support process to reduce the reply time after request delegation is made to another staff to make support process faster and efficient. It is interesting to note that none of the respondents are of the opinion that the response from the second support staff can be within an hour. This observation further strengthens the distributed nature of the support process in climate cyber-infrastructures.

Most of the respondents believe that the user's base (total number of end users) range from 50 to several hundred. According to recorded responses, numbers of incoming requests from end-users per working day are from 2 to 5. This means in this case personal relationship can be made with the users. According to the survey results, it is evident that most of the end-users as according to the observation of support staff's (59%) perspective are satisfied with the support process. However, there is a room for further improvement in terms of satisfaction because no one is "very much" satisfied (as according to survey) with the support process, which is similar to the satisfaction-level of support staffs.

Most of the support staffs (80%) do not collect statistics of incoming user requests. Multiple delegation (MD)[4] is not

[1] First Level Support (FLS) is also known as Front Line Support meaning that the incoming user queries are handled by the front line staff which are generalists and non-specialists. The user queries that could not be handled by FLS are referred to Second Level Support (SLS) for further treatment.

[2] Second Level Support (SLS) is also known as Second Line Support meaning that the incoming user queries are handled by the second line staff which are specialists.

[3] Simple and routine queries are the queries that can be solved by FLS in very little time. In some cases, if adequate and relevant information is provided to the user then a user may solve the issue him/herself. Example of simple and routine queries include: Password queries, account queries, hardware and software enquiries.

[4] Multiple delegation happens when a user support staff sends the requests at the same time to various colleagues to treat the request further.

always made in user-support process in climate cyber-infrastructure projects. However, there is a practice of multiple delegations in the user support process only to a limited extent. The update of user-support information on the user support portal i.e. e-support is not regular. Most of the incoming user-requests are in English language.

From the results, one can deduce that in cyber-infrastructures, "e-mail" leading to RT and mailing lists (ML) is the major mode of communication between users and the support staff as opposed to telephone in some other organizations using IT help-desk support, also known as call centers (Major sources of user contact table 21 Lau and Leung 2007 Journal of CIS). According to respondents (support staffs), users prefer e-mails leading to RT (Request Tracking software) the most as it was ranked the highest. It is noteworthy to observe that this evidence is similar to the evidence from Q. 23 of the survey. The users just like support staffs prefer e-mails leading to Request Tracking software (RT) the most. Mailing lists (ML) are ranked second. Therefore, from these results one can deduce that in e-Science users do not prefer other media as a priority, only e-mails "leading to RT" is the popular mode of communication between users and the support staff. This result is in accordance to the distributed nature of cyber-infrastructures. From the results one can infer that even in the internal communication between cyber-infrastructure staffs, "E-mail leading to RT, followed by Mailing List" are also the popular media. It is important to note that some of the incoming user-requests remain unresolved, which is not a positive sign.

Most support staffs handle user-requests in a non-documented way i.e. after or during treatment of user requests the documentation about the nature of solution or the nature of problem is not made and stored in a repository. Most of the support staffs (90%) as according to responses collected from the survey, do not document the incoming user-requests. It is interesting to know the different forms of communication via which non-documented user-requests are received by the support staffs. From the survey results, it is evident that the unrecorded e-mails, telephone calls followed by walk-in and unrecorded mailing lists are the media which are used for undocumented user-requests. None of the support staff said that they need to write more than two times, which means that the incoming user requests are generally understandable.

In the current user support process there is no categorization of the incoming user requests done. From the survey, different possible sub-categories of the first proposed main category i.e. general information are suggested by respondents as follows:

- Authentication category; i.e. getting logged in
- Software usage category; model compilation, debugging, optimization, feature development, software installation
- Quality control; Queries about correctness of the data which have been downloaded
- Projects e.g IPCC AR5, CORDEX,, CFMIP, LUCID;
- Models: ECHAM6, MPIOM, JSBACH, HAMOCC, CCLM4
- Data publication category; how to publish data i.e. data publication?
- Post-processing of data
- Data citation
- Miscellaneous

CRITICISM AND SUGGESTIONS FOR IMPROVEMENT

In the current user support process in climate cyber-infrastructures, there is no assignment of user support roles. Consequently, top computer scientists end-up performing the duties of the front desk, in parallel to their main (core) duties. Cyber-infrastructure employees already have a daily plan of action (P) in their memory i.e. doing activities such as software development, data curation, writing documents etc. So, the planned tasks are known beforehand. The incoming events (E) or event tasks, which are user's or other employee's requests that can be triggered at any point in time, are actually interruptions to their daily work plan of action. Therefore, the incoming event tasks may affect the routine of employees, affecting the planned tasks. As a consequence, it requires cyber-infrastructure employees to allocate more time (T) to process user queries or alternatively dedicating less time to their core tasks. Employee's contribution to process user queries is not formally recognized by the cyber-infrastructure management in such a way that incentives are provided to the employees who process user requests.

There can be two types of event tasks: First, explicit event tasks, meaning that task performed on the request of a user or an employee i.e. responding to users' or other employee's e-mail received. The second type is an implicit event task which is performed by employees based on their own cognizance (awareness); for instance; updating the web page by inserting relevant information for users or inserting an alternate solution in the repository after its discovery. So all of the planned tasks and event tasks are needed to be prioritized and there is no priority rule existing in the current user support practices. The current priority is based on employee's own judgment, which is noteworthy to be investigated further by asking employees. After prioritization of tasks, tasks are then actually performed within a specific time (T) using certain means (M). T and M are generally also not recorded in the user support process of climate cyber-infrastructures. The current practice of T and M varies from location to location and time.

There is no central repository or database to record user requests and the solutions. Consequently, the solution to the redundant enquires or enquires of similar nature cannot be known or matched via a central information retrieval system. Though, help webpages are available at the ESGF and its associated web links, however, a central information retrieval system is currently not available to the users or to the employees where relevant information can be sought by browsing the categories of redundant queries and their solutions. It is in the interest of users as well as employees of cyber-infrastructures that a knowledgebase and information retrieval system may be designed, developed and made available online. At first it seems a lot of effort and it may require resources but in the long run it will benefit the users and employees both.

From the survey results, it is an important observation that not all questions are answered by the employees. Because of the multiple channels of communications and no ticket assignment mechanism in the user support process, some of the user requests might be over-seen (at some intervals). At times, the user support staffs that are scattered world-wide, may rely on other colleagues and may not find themselves responsible to answer or even to take all user queries. So keeping in view the current arrangement of user support process, no one feels answerable, which is not a positive aspect of the current user support process in climate cyber-infrastructures.

In a nutshell, there is a process of user support which is existing in the climate cyber-infrastructures but it is not formally defined in the cyber-infrastructures in such a way that it is measured in terms of quality, cost and standardization. Its resources are explicitly allocated or it is organized as per wishes of cyber-infrastructure users. So far, there was no survey conducted to capture and analyze users' wishes or to measure usability and accessibility of UI of ESGF and ESGF-like cyber-infrastructures and users' help portals. There is no process owner(s) of the user support process and in the ESGF governance scheme there is no explicit regulated process about user support process in ESGF.

Unfortunately, there is no categorization of the incoming user queries and no information is available about the categories. Therefore, all the incoming user requests are answered based on the experience and knowledge of cyber-infrastructure employees. There is no document, repository (database) or an information retrieval system available where the user request categories can be stored for future reference in order to avoid redundant effort of working similar user requests. Similarly, there is no user self-help culture where the users may look at the self-help retrieval system to get the initial information about the problems they are facing.

In the current user support process, users do not provide feedback or suggestions to improve the whole process. Similarly there is no mechanism of rating solutions. Users can be motivated to provide suggestions and rate the

solution by developing a facebook-like platform where a user may also comment on the problems of other users. The solution time, employee's responses time to user and between supports staffs, in case the request is delegated, is not uniformly defined and agreed upon by all stakeholders. Service level agreements (SLA) can be made and applied to regulate the current user support so that the employees know that they have to deliver some quality of service within a specified time. Plausible SLA can be formulated by forming a consensus between executive committee, the technical committee and other stakeholders of ESGF cyber-infrastructure.

CONCLUSIONS

Looking at the results of the survey one can say that in a nutshell, the user support process is working fine for the time being. However, this does not mean that it will work in the same way it is working now, in future forever. Therefore the current user support process in ESGF and ESGF-like cyber-infrastructure needs stepwise reforms and standardization. If appropriate measures are not taken at the right time (i.e. now) to standardize and study the requirements of end-users then the user support process can deteriorate in future. Because the data holdings offered by ESGF are on the increase day by day, besides ESGF is expanding its services to other domains such as chemistry and biology. So in future, as evident from history of ESGF; it is foreseen that there will be continuous changes in the architecture of ESGF and ESGF-like cyber-infrastructures as well as in their organization. These changes do affect the user support process on the whole. Since there are no dedicated user support staffs in ESGF and ESGF-like cyber-infrastructures, it is imperative to introduce dedicated user support staff at least at the first level support. The notion of user support staff or human user support agents is not formally defined in ESGF and its associated cyber-infrastructures in the same manner as the position of data manager or data steward was not part of e-Science in the past. However, with the passage of time the designation of data manager was formally recognized and nowadays in the field of e-Science data manager is a position that deals with data curation and data quality control activities (Karasti, Baker, & Halkola, 2006). In the same manner the importance and necessity of user support manager is being felt in climate cyber-infrastructures. One can assign a user support staff position to a student or a less qualified person who is able to mitigate the simple and routine queries in the user support process in cyber-infrastructures, so that the top computer scientists are not busy with doing trivial tasks. Moreover, it is suggested that the position of user support manager be introduced in the cyber-infrastructure projects and the user support manager can be a process owner at a particular user support unit at a particular location that can manage his/her user support activities falling in his/her jurisdiction.

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

Couper, M. (2000). Web surveys: a review of issues and approaches. *Public opinion quarterly*, *64*(4), 464–94. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11171027

Hey, T., & Trefethen, A. E. (2005). Cyberinfrastructure for e-Science. *Science (New York, N.Y.)*, *308*(5723), 817–21. doi:10.1126/science.1110410

Jäntti, M. (2009). Lessons Learnt from the Improvement of Customer Support Processes: A Case Study on Incident Management. In *10th International Conference, PROFES 2009, Oulu, Finland, June 15-17, 2009. Proceedings* (pp. 317–331). doi:10.1007/978-3-642-02152-7_24

Jäntti, M. (2012a). Improving IT Service Desk and Service Management Processes in Finnish Tax Administration : A Case Study on Service Engineering, 218–232.

Jäntti, M. (2012b). Examining Challenges in IT Service Desk System and Processes : A Case Study, (c), 105–108.

Karasti, H., Baker, K. S., & Halkola, E. (2006). Enriching the Notion of Data Curation in E-Science: Data Managing and Information Infrastructuring in the Long Term Ecological Research (LTER) Network. *Computer Supported Cooperative Work (CSCW)*, *15*(4), 321–358. doi:10.1007/s10606-006-9023-2

Knapp, M., & Woch, J. (2002). Towards a Natural Language Driven Automated Help Desk. In *Third International Conference on Computational Linguistics and Intelligent Text Processing*.

Lazar, J., Feng, J., & Hochheiser, H. (2010). *Research Methods in Human-Computer Interaction*. Indianapolis, IN: Wiley.

Leung, N. K. Y. (2006). *University of Wollongong Thesis Collection Turning user into first level support in help desk : development of web-based user self-help knowledge management system*.

Leung, N., & Lau, S. (2006). Relieving the overloaded help desk: a knowledge management approach. *... of International Information Management ...*, *6*(2), 87–98. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.98.1334&rep=rep1&type=pdf

Leung, N., & Lau, S. (2007). Information technology help desk survey: To identify the classification of simple and routine enquiries. *Journal of Computer Information Systems*. Retrieved from http://ro.uow.edu.au/era/859/

Soehner, C., Steeves, C., & Ward, J. (2010). E-Science and Data Support Services, (August).