Virtual Physiological Human Network of Excellence

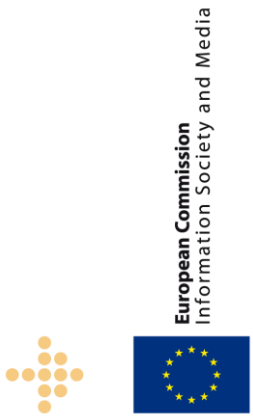
Grant Agreement: 223920

VPH ToolKit Guideline Document

Topic: Usability and Training

Version 1.0

2-Mar-11



This page is intentionally blank

Document Information

IST Project Num	FP7 – 2007 - ICT - 223920	**Acronym**	VPH NoE
Full title	Virtual Physiological Human Network of Excellence		
Project URL	http://www.vph-noe.eu		

Document	**Number**	G08	**Title**	Guidance (Usability)

Status	Version. 1.0	Final ☑
Dissemination Level	Public ☑ Consortium ☐	

Authors (Partner)	UCL	Haidar, Zasada, Nash
	UOXF	Cooper, McCormack
	CNRS	Friboulet, Cervenansky
	UPF	Omedas
	UOA	Lloyd
	UoS	Fenner

Responsible Author	Haidar		**Email**	Ali.Haidar@ucl.ac.uk
	Partner	UCL	**Phone**	+44 (0) 2076795300

Abstract (for dissemination)	This document provides guidance on the usability factors used in assessing the tools to be submitted for the VPH ToolKit.

Version Log

Issue Date	Version	Author	Change
22-July-10	0.1	WP3 – Haidar	First version
29-Jul-10	0.2	WP3 – Haidar	Second version with merged comments
05-Aug-10	0.3	WP3 – Haidar	Final draft before annual review
05-Mar-11	0.4	WP3 – Haidar	Final draft before EBI meeting
22-Mar-11	1.0	WP3	First public release

This page is intentionally blank

Table of Contents

This page is intentionally blank

EXECUTIVE SUMMARY

This guideline is intended to assist potential contributors to the VPH NoE Toolkit in assessing the usability of their tools.

There are several such guideline documents in this series, covering the full range of issues affecting content providers. They are being developed over a period of time and, once finalised, these guideline may be bound together into a single VPH NoE resource.

Introduction

This document is one of a series that together build to form a complete guide to the ideal content and presentation of materials offered for distribution via the Virtual Physiological Human Network of Excellence ToolKit Portal. The full set of Guideline Documents is summarised below.

Guidance Area	**Description**
Tool characterisation	The attributes important for inclusion in the documentation of Tools, including performance validation
Model characterisation	The attributes important for inclusion in the documentation of Models, including performance validation
Data Characterisation	The attributes important for inclusion in the documentation of Data
Ontological Annotation	The significance, benefits and methods of ontological annotation of ToolKit content
Interoperability	Key attributes concerning the additional specification of predominantly tools and models that will allow operation in a multistage workflow alongside other items of ToolKit content
Ethico-legal issues, provenance	The inherited responsibilities that are attached to any item of ToolKit content – perhaps particularly data – including legal, ethical and territorial restrictions
Licensing	The conditions that apply to the legitimate use of the content from a commercial and intellectual property standpoint
Usability & Training	The factors that are important for the easy use and ready acceptance of ToolKit content, taking into account the environment, the likely users and the need for interoperability. Additionally, the nature of training facilities of all types appropriate to particular content categories.

Usability and Training Topic

The VPH initiative is expected to provide a set of tools that allow scientists and clinicians to access a range of resources distributed across multiple administrative domains to perform patient specific modelling and simulations. These tools will be provided by VPH partners and from other scientific projects where these tools are well known and proven to be very successful.

Usability of these tools is paramount for their acceptance across VPH community. Many tools developed by research projects do not include usability evaluation as part of their standard software development process. A shining example of usability problems is encountered in the use of complex grid middleware and complex security mechanisms used to access High Performance Computing resources. Usability evaluation can provide evidence that there is problem and it has proven itself as a key ingredient in any organisation that wishes to attract users.

One key aspect of the VPH Toolkit is to provide potential end-users with guidelines to evaluate and assess the usability of the tools that will be exposed by the toolkit. The aim of this document is to provide criteria carefully tailored to help the VPH community study the usability factors that end-users are looking for in tools used for scientific research and in security mechanisms enabling secure access to resources used by these tools.

The document is structured as follows:

- ✓ The first section will give a definition of usability in the context of the VPH-NoE. This definition will allow specifying the scope of usability and gives an overview of current usability study techniques.
- ✓ The second part will describe in detail the usability criteria proposed for the VPH. Each element will be described in general and within the VPH context.
- ✓ The third and final section discusses a particular area in which usability is often lacking: the security mechanisms that will be used by the VPH ToolKit to enable secure access to patient's data. Having complex and unusable security controls results in many users engaging in practices which weaken the security of the environment.

Underlying Concepts

VPH tools usability is essential for the uptake of any VPH solution. ISO 9241-Part 11, a standard for usability, defines usability as follows:

"Usability is the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"

The study of VPH tools' usability deals with measuring how easily a user can employ a tool to perform prescribed tasks that successfully meet the goals intended by the designer (effectiveness and efficiency). In the VPH context, the user in question may be a clinician, a consultant, a scientist or a system administrator installing, maintaining or administering the tool. Usability is also used to describe the quality of user experience, which is supposed to be positive (satisfaction). To ensure the satisfaction of end-users, while maintaining efficiency and effectiveness, considerable care is required in the tool design process and could be better guaranteed with the use of formal usability evaluation procedures by tool developers. Implementing formal evaluation procedures can often save time and money, contrary to the common belief that usability testing increases overall production time and overheads costs on projects.

When a tool is submitted to the ToolKit, the first step is to look at how the tool was designed by looking at the tool's documentation and whether usability engineering principles were used in the design process. Most usability design principles look at the following properties which are relevant to VPH:

- ✓ Learnability: this involves looking at who are the users, what is the general background of the users? What do they know, and what can they learn? And how easy is it for users to accomplish basic tasks the first time they use the tool?
- ✓ Efficiency: Once users have learned the tool, how quickly can they perform tasks?
- ✓ Memorability: When users return to the tool after a period of not using it, how easily can they re-establish proficiency?
- ✓ Errors: How many errors do users make, how severe are these errors, and how easily can they recover from the errors?
- ✓ Satisfaction: How pleasant is it to use the tool?
- ✓ Can users easily accomplish their intended tasks? For example, can users accomplish intended tasks at their intended speed?

These factors are a subset of the usability factors, described in details in the ISO standard, provides detailed guidelines on the methodologies and techniques for ensuring the production of usable tools. The guidelines presented here are tailored for tools intended to be included in the VPH ToolKit.

Usability Evaluation Criteria

This section presents the criteria that will be used for evaluating tools submitted to the VPH ToolKit. The first question that arises when measuring the usability of a tool is: what aspects of the tool can be evaluated?

In the context of the VPH, it is important to set the goals for usability tests. These include success rates and the time it takes users to perform specific tasks, satisfaction and acceptance by users [1,2,3].

Tool Portability: The dependency of a tool on a specific platform is a major barrier for its adoption. VPH users come from various backgrounds: hospitals, academia and industry. Many hospitals adopt MS Windows platforms because they are closed environments and because of the ease of technical support provided by many companies to run such environments. In academia, researchers tend to use open source platforms such as the free Centos and Ubuntu Linux flavours. Hence a tool that is platform independent will be useful to all concerned. Using the Java programming language is one way to ensure that tools are interoperable. Another way to ensure tool portability is by exposing its interface as a Web Service. This will be described in the Interoperability guidelines document.

Installing the tool: The ease with which a tool can be installed and configured is a critical factor in the usability of the tool. Will a clinician or researcher be able to install a tool or does it need an expert user? The installation and configuration of grid middleware, such as Globus toolkit and UNICORE [10, 17], stand out as examples of problems encountered. Although there is documentation provided by the tools developer, often only expert users are capable of installing these tools [18].

Defining Benchmark tasks: Tools submitted to VPH ToolKit usually are designed for a specific purpose such as visualization, simulation, to run a job/application on grid resources or add security features such as authentication, authorization and auditing. Defining benchmark tasks in a tool is a common method for measuring:

1. Ease of use: this activity is a user-centric activity and investigates how can users easily accomplish their intended tasks in a timely manner.
2. The speed of performing tasks: can a user perform the tasks at an acceptable speed?
3. Completing a task: not being able to perform a task because of incomplete or erroneous documentation.

4. Errors: How many errors do users make, how severe are the errors, and how easily can they recover from the errors?
5. Satisfaction: How pleasant is the tool to use?

Participant Numbers: The number of participants to use when conducting usability evaluations depends on the method being used [4, 5, 6, 18]. Using too few may reduce the usability of a tool; using too many wastes valuable resources. When conducting VPH tool usability testing with end-users, it is important to test with different types of end-users: clinicians, researchers and experts. Also it is critical to have trained usability specialists as test facilitators and observers to ensure integrity of the test. The use of the power analysis technique [18] gives an approximation for the number of users required to take part in a study depending on the number of tasks provided by the tool.

Documentation: This is a vital part in the assessment of usability [7]. What documentation or other supporting materials are available to help the end-user? Can end-users find the answers to the queries they seek in this materials? The greater the number of exchanges of information with potential users, the better the developers' understanding of the users' requirements and the better documentation is provided. Documentation in the VPH context includes help documents and specifications of the tool functionalities that can be expressed in different ways such as use cases, UML diagrams or plain English. For more details about this item the reader is referred to the tool characterisation guidelines. A common practice with tools is to offer tutorials aimed at new users for performing typical tasks provided in the tool functionality; and this is considered as a good practice for addressing learnability.

Interoperability: This is a key point of interest for VPH software assessing a tool from the VPH Toolkit. The interoperability guidelines document addresses this issue in detail.

Learnability: For a tool to be acceptable, it is critical to provide end-users with sufficient training. The main issues in this area is how much training do users need and how the training materials are prepared? For example, by the developers of the tool or by a dedicated specialised entity such as the Association of Medical Education in Europe (AMEE www.amee.org). Organising workshops is one approach to train users on specific tools. For example, the EU grid infrastructure DEISA [11] organises regular workshops to train users to access High Performance Computers using light weight grid middleware – The Application Hosting Environment [18].

This area is dealt with in WP4 [http://www.vph-noe.eu/wp4]

"WP4 will address training and career development for both early and in-career VPH researchers. Activities will also pay special attention to the outcomes generated from other VPH-related projects and existing EC-initiatives (e.g. Marie Curie), in order to ensure complementarity with existing activities. The ultimate goal of this WP will be the design and implementation of actions directed at the development of VPH research education and careers, and will relate closely to WP5."

Evaluation of Security Mechanisms in the Toolkit

Another important area in the VPH tools usability guidelines is the usability of the security mechanisms used that enable access to patient's data and grid computing resources: having complex security mechanisms will deter end-users from adopting the tools [18,19]. It has been well documented [8] that most security mechanisms used in computational grids are not user-friendly, and only considering resource providers' requirements and ignoring end-users' requirements. The factors listed below are based on commonly encountered problems in computational grid projects and virtual organisations.

Number of Credentials: VPH tools will be used in virtual organizations (VO) that involve different institutions (hospitals, universities, national computation grid infrastructures) sharing different type of resources, such as visualization tools, supercomputers and patients' data. For a VPH user (clinician or researcher) to be identified at each institution of the VO, the user may need to maintain separate credentials for each organization in the VO. The number of separate identities that a user has to manage has a substantial impact on usability. The issue is compounded when the number of separate identities grows. As a result, users will not be able to easily remember the correct identity for each site, causing more confusion for users. In addition, having to be authenticated at several separate security domains can be time consuming, error prone and impractical from a usability perspective.

Acquiring Credentials (User enrolment): This step is jointly performed by the user and the organization providing the resources. It involves a user request for credentials and the organization issuing these credentials. For instance, when a new member of a staff joins a university, he/she fills a form for the IT help desk requesting credentials and an email address. The help desk checks with the university human resources department whether the staff is indeed an employee of the university prior to issuing the credentials. The process of acquiring credentials for accessing shared resources should be as simple as acquiring credentials for a local organisation. This is significant for potential users because a complicated and lengthy process will deter users from using VPH tools [8, 18]. Some VPH tools use digital certificates, issued by different certificate authorities (i.e. UK e-Science), to authenticate end-users, but issuing such certificates is time consuming.

Administration: The set of VO users using VPH tools is usually dynamic. New users are added and existing users are removed as required. Maintaining this dynamic membership set involves administrative overheads. In order to measure the administration aspect we consider four benchmark tasks:

1. Create new user account on the VO.

2. Change password operation for a VO user.
3. Remove an existing VO user.
4. Perform authentication to access VO resources.

Changes to Security Policies of Local Sites: In order to be a part of a VO, each organization may have to make some changes to their local security policies. The extent of these changes is a key factor in the evaluation of the security mechanism to be adopted by VPH toolkit.

Trust and Security Risk: Typically a VPH user will have access to a wide range of resources within a VO. The question to be addressed is: how does a user from an institution become a VO member? By joining a VO, new security risks are introduced to an organization. The source of these new risks comes from trusting new users, shared databases, or certificate authorities.

Scalability: The number of VO users can vary from a handful to thousands. How do factors such as usability, administration, policies and security threats scale with the growth of the number of VO users?

The tables below summarise the above criteria for assessing the usability of the tools and security mechanisms submitted to the Toolkit. They are intended as an illustrative example for comparing tools. Below is a brief description of the tables' entries:

Name: Application name. This field shall be enough clear to identify a tool.
Language: Programming language used in the development.
OS: List of supported operating systems (Windows, Linux, MacOs)
Command Line Support: Running the tool functionality from command line.
GUI Interface: The tool has a graphical user interface.
Help documentation: This includes tutorials ("getting started" guide) explaining to new users how run operations offered by the tool.
Specification documentation: This explains how the tools operations were designed i.e. what algorithm is used by a simulation etc. This is needed to ensure the correctness of the functionality offered by the tool. A common way used by software developers to model the specification is with diagrams, flows and UML notation.

Task Description: Describes the main tasks performed by the tool. Having a clear description will enable the comparison of the tools which perform the same tasks – how easy to run, time to task completion.

Installation Guide: This includes detail on how to install the tool on various operating systems.

We have recently completed a comprehensive usability study [18] that involved comparing several middleware products for accessing grid resources environments. These include the Application Hosting Environment (AHE) middleware, which comes with graphical and command line interfaces for accessing grid resources, a combination of AHE with Audited Credential Delegation security plugin that enables access grid resources with username-password pair, as well as UNICORE and Globus. There were forty participants drawn from different departments and faculties at UCL including Physics, Chemistry, Computer Science, the Medical School, the Business School, the Cancer Institute and the Law School. Each participant was asked to:

1. Run a simulation on a grid resource (NGS) using the different middleware.
2. Configure the security of their client tools, and use the credentials given to them (username/password, X.509 certificate).

A brief overview of the middleware characterisation is shown in the table below. For more detail about the study the reader is referred to [18].

	Grid Middleware				Security Mechanisms	
	Globus Toolkit GUI	**AHE Client GUI**	**UNICORE GUI**	**AHE GUI**	**AHE with Certificate**	**AHE with ACD**
Version	2.0	2.0	6.	2.0	2.0	2.5
Function	Middleware	Middleware	Middleware	Middleware		Security plugin
Speciality	Access to computational Grid resources	Access to computational Grid resources	Access to computational Grid resources	Access to computational Grid resources	Authentication	Authentication, authorisation and auditing of AHE functionalities
Language	Java	Perl	Java	Java	Java	Java
OS	Independent	Linux/Tomcat	Linux	Independent	Independent	Independent
Command line Support	Yes	No	Yes	Yes	Yes	No
Graphical Interface	No	Yes	No	No	Yes	Yes
Help Documentation	Yes	Yes	Yes	Yes	Yes	Yes
Specification Documentation	Yes	Yes	Yes	Yes	Yes	Yes
Tasks description	Yes	Yes	Yes	Yes	Yes	Yes
Installation Guide	Yes	Yes	Yes	Yes	Yes	Yes

	Middleware Tests				Security Tests	
	Globus Toolkit	AHE CLI	Unicore GUI	AHE GUI	AHE with Cert	AHE with ACD
Percentage of successful users	45.45	75.76	30.30	96.97	66.67	96.97
Percentage of users satisfied with tool	27.27	53.54	47.47	79.80	51.52	87.88
Percentage of users who found tool difficult to use	45.45	25.25	26.26	5.05	27.27	0.00

Figure 1. Summary of statistics collected during usability trials for each tool under comparison.

The above table shows very clear differences between the different tools tested. The table below, Figure 2, shows the mean time taken to run a task on the grid with each tool.

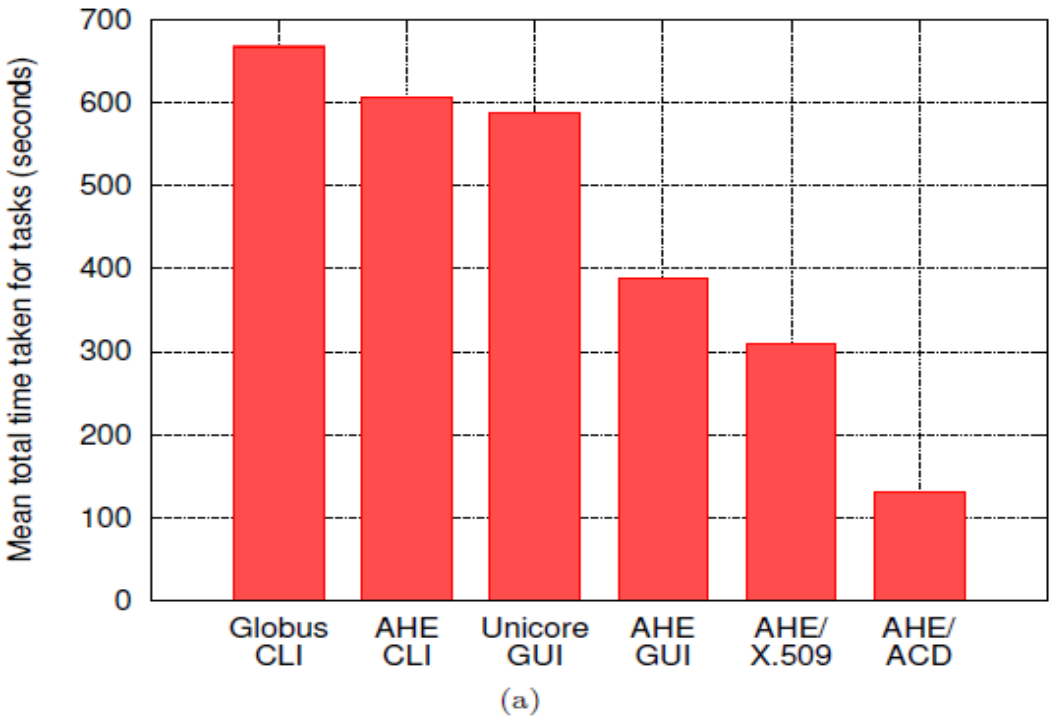


Figure 1. Mean time taken to complete a range of tasks with each tool.

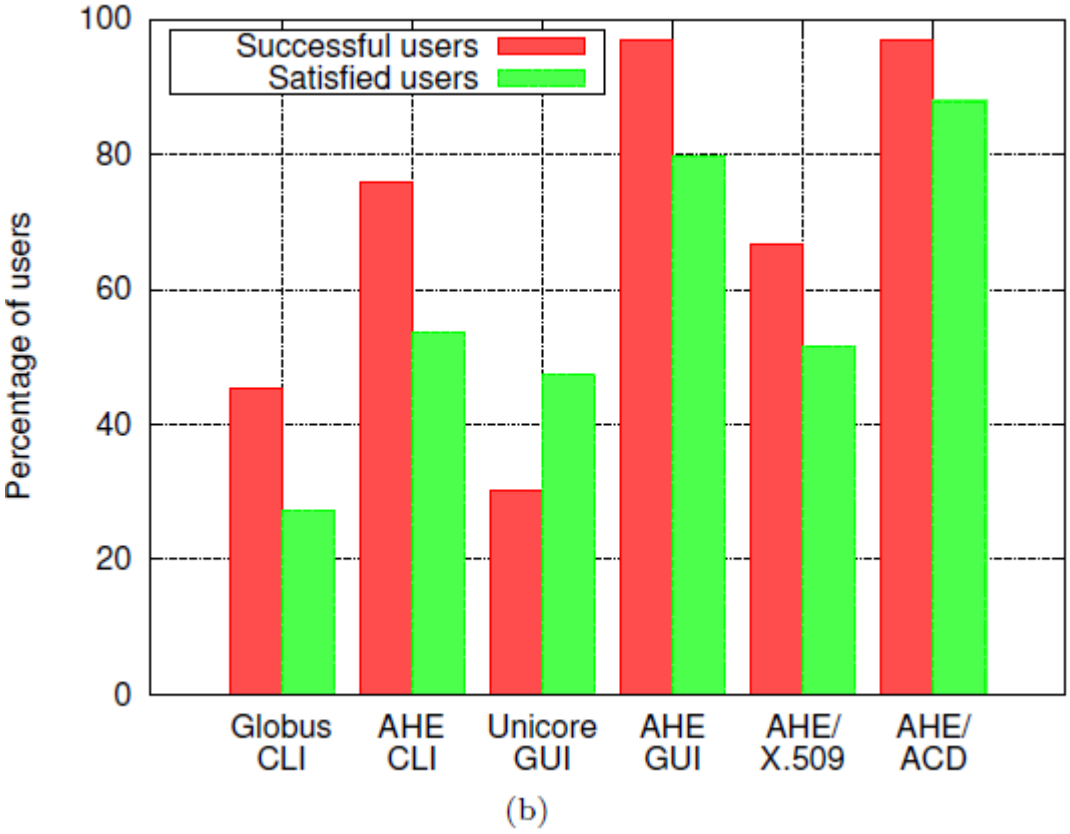


Figure 3. A comparison of the percentage of users who were satisfied with a tool and the percentage who could successfully use that tool.

In addition to the above factors for measuring usability of tools functionality the factors below, described in details in the previous section, are used for assessing the usability of security mechanisms used in the VPH toolkit:

Criteria	AHE with ACD (AHE 2.5)	Globus	UNICORE	AHE 2.0
Number of Credentials	1	1	1	1
Adding Users accounts	Yes First account takes 2 weeks. The second account is instant	Yes Takes 2 weeks	Yes 3 days	Between 3 days and 2 weeks.
Removing Users Accounts	Yes (instant)	24 hrs	24 hrs	N/A
Setting Permissions	Yes – fine grained	Group membership	Group membership	N/A
Acquiring Credentials	Less than 24 hrs	on average 3 days	On average 3 days	On average 3 days
Easy to Install/Configure	Yes	No	No	Yes
Changes to Security Policy	Minor	Minor	Minor	Minor
Scalability	Yes	Yes	Yes	Yes

Training

WP4 deals with arranging training policy for new tools emerging from VPH.

Further Information

Readers seeking more detail should be able to find sources listed and explained here

References

[1] Baca, B. & Cassidy, A. (1999). Intranet development and design that works. Human Factors and Ergonomics Society Annual Meeting Proceedings, 777-790.

[2] Bradley, R.F. & Johnk, L.D. (1995). Replacing a networking interface 'from hell.' Proceedings of CHI'95, 538-545.

[3] Sears, A., Jacko, J., & Borella, M. (1997). Internet delay effects: How users perceive quality, organization and ease of use information. Proceedings of CHI'97, 353-354.

[4] Cockton, G. & Woolrych, A. (2001). Understanding inspection methods: Lessons from an assessment of heuristic evaluation. In A. Blandford, J. Vanderdonckt & P.D. Gray (Eds.), People and Computers XV: Interaction without Frontiers (pp. 171-191). Heidelberg: Springer-Verlag.

[5] Bailey, G.D. (1993). Iterative methodology and designer training in human-computer interface design. Proceedings of InterCHI'93, 198-205.

[6] Brinck, T. & Hofer, E. (2002, April). Automatically evaluating the usability of web sites. Proceedings of CHI 2002, Extended Abstracts, 906-907.

[7] Osborn, S. & Elliott, G. (2002). Standard creativity: Creating flexible web development standards. Proceedings of the IEEE International Professional Communication Conference, 1-21.

[8] B. Beckles, V. Welch, J. Basney, "Mechanisms for increasing the usability of grid security", *Int. J. Human-Computer Studies*, Vol. 63, No. 1-2, pp. 74-101, July 2005.

[9] The TeraGrid project. [http://www.teragrid.org/]

[10] The Globus project. [http://www.globus.org].

[11] Distributed European Infrastructure for Supercomputing Applications. [URL: http://www.deisa.org].

[12] M. Arning, N. Forgó, and T.a Krügel, "Data protection in grid-based multicentric clinical trials: killjoy or confidence building measure?", Phil. Trans. R. Soc. A, July 13, 2009, 367:2729-2739.

[13] http://ww.owasp.org The Open Web Application Security.

[15] A. E. Abdallah and A. N. Haidar. "Usability Evaluation of Identity Management Schemes in Three Virtual Organisation Architectures". Int. J. of Information Assurance and Security, Special Issue on Information Assurance and Data Security, 4(6):560–570, July 2009.

[16] Cooper *et al.*, 2010. The Virtual Physiological Human ToolKit. Phil Trans R Soc A, 368:3925–3936, August 2010. http://www.vphnoe.eu/vph-repository/doc_download/167-the-virtual-physiological-human-toolkit

[17] http://www.unicore.eu/ The Uniform Interface to Computing Resources project.

[18] S. J. Zasada, A. N. Haidar, and P. V. Coveney. On the Usability of Grid Middleware and Security Mechanisms. 2011. Accepted for publication in the UK e-Science AHM 2010 theme issue of the Philosophical Transactions of the Royal Society A.

[19] A. N. Haidar, S. J. Zasada, P. V. Coveney, A. E. Abdallah, B. Beckles and M .A. S. Jones. "Audited Credential Delegation: A Usable Security Solution for the Virtual Physiological Human Toolkit". Accepted for publication in the VPH 2010 theme issue of Interface Focus of the Royal Society Interface Journal, March 2011.

