IST-2001-33127

SciX

Open, self organising repository
for scientific information exchange

D11: Overall assessment
and evaluation report

Authors: Almudena Fernandez Valero,
Grahame Cooper, Brian Clifton, Turrid
Heglund

Access: private
Status: relesed

Version: 1.0

Date: March 30, 2004

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 2

EXECUTIVE SUMMARY:

This report describes the assessment and evaluation of the SciX project. It consists of five sections dealing with the different aspects of the work:

- Assessment and evaluation of the conceptual work which was performed by external experts such as Jean-Claude Guedon, John Cox and Dr. Heleen Gierveld.

- Usability of the pilots was done by professions Web site auditors at Indra.

- Comparison of the SciX pilots against criteria set by the Open Software Institute.

- Fulfilment of the objectives in the Technical Annex.

- A brief SWOT analysis of the pilots.

The evaluation had a positive impact on the work with SciX and is being taken into account while preparing new versions of the software. It will have an impact on the project in the last two months as well.

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 3

RELEASE HISTORY

date   changes

3.11.2002   Version 0.1 … proposed methodology only

30.3.2004   Version 1.0 … including all main chapters, subject to an update with regard to comparison of SciX with Open Software Institute

TABLE OF ABBREVIATIONS

SWOT ... strengths, weaknesses, opportunities, threats

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 4

TABLE OF CONTENTS:

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 5

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 6

1. VALIDATION OF CONCEPTIAL WORK

Both the as-is model as well as the to-be model has been shown for comments to several leading experts on the processes described, among others prof. Jean-Claude Guedon from the University of Montreal, Leo Waaijers from the Dare project in the Netherlands and Ingegerd Rabow from the Lund University library. Most comments have been informal presented on the workshop arranged in Helsinki in January 2003 and on conferences and visits and not in writing. They have all been valuable in developing the to-be model.

The following comments were made in an e-mail by Dr. Heleen Gierveld from the University of Technology at Delft.

- The model is quite readable, especially the almost "XML-like" listing of the activities, that provided a good orientation / map. And the inclusion of Control and Mechanism quite important for relating the processes to socio-economic issues. It will be interesting if these can be included in the model.

- I like the embedding of the actual research efforts in the information transfer activities. Disseminating is indeed only part of the costs and a means to gain the overall insight and new scientific knowledge. But I tend to support the –conservative- reasoning that administration will probably not be too bothered about these costs, as the costs of libraries are only peanuts compared to university research budgets. I like the argument you bring forward that all those research efforts are poorly distributed, and that hence the effectiveness of the research done is at stake. Yes, the changes would probably come from such concerns; not so much from library budget concerns.

- Then the model focuses on A2 and A3, as these are the processes that the information sciences are concerned with. I noticed, however, that you do include both business management processes (eg A2231) and information processes. Of course both have an impact on the cost model, but would they not make it more complex to model the costs at a generic level? Anyway, if management processes are included, how do you feel about including R&D activities and market research – to me these seem needed in order to sustain a product or an organization, especially in a digital environment where so much is going on.

- Then, the dissemination of knowledge is the end result of A2 and A3. This (the end result) probably compares to what I call functions. The four functions that I have adopted (Registration, Awareness, Certification, Archiving) all contribute to the dissemination of knowledge. I like to think in functions, but I admit that I have taken these four almost as an axiom, so I would be happy to reconsider them. However, I feel that none of these four can be omitted, if the are led back to the user information needs, which should be the basis of all processes. Regardless changing information processes, the same functions should be fulfilled. I like it that the model includes information seeking activities, both the active (search) and the passive mode (alerted).

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 7

In the following section the comments are from John Cox, the John Cox Associates who has validated the SPLC model. John Cox whole evaluation of the model can be found in the to-be report in deliverable 2.

First, a general observation must be made. The SPLC model focuses on the process of authoring, publishing, retrieving and utilising the journal article itself. Except in the case of indexing and archiving (referring to Diagram A23) it is a very detailed and faithful representation of the processes involved. However, it gives insufficient weight to the publishing organisation's infrastructure that supports the article (see Diagram A2231) and, if taken at face value and out of the context of that infrastructure, may mislead and oversimplify the organisation and management functions that are necessary for the processes to work.

- Diagram A2232 is misleading; taken together, the diagram and the accompanying text imply that the need for marketing – to generate demand for subscriptions – is reduced for open access journals. However, marketing includes the broader range of activities. There is little evidence of active marketing of open access journals by the university departments or societies that are responsible for most of the titles listed in the Lund Directory, with the exception of BioMed Central, which itself is a for-profit publisher managed by experienced publishing professionals.

- The hierarchy and Diagram A231, and the accompanying text, state that the need for commercial indexing services may reduce and that the use of general or specialist search engines will increase. This remains to be proven. For most academic researchers, it remains true that ISI's Citation Indexes remain the indexes of choice; subject-based indexes are widely used to identify and locate relevant articles. Open access journals are only just beginning to be indexed and catalogued.

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 8

2. USABILITY STUDY ON THE PILOTS

2.1 INTRODUCTION

This document contains the findings of the usability testing activities conducted on the design of two web sites framed under the SCIX initiative: ITC Digital Library (http://itc.scix.net) and the Electronic Journal of Information Technology in Construction (http://www.itcon.org).

These activities took place between November, 2003 and January 2004, in Madrid and Ljubljana respectively. The document details what methodology has been followed, what were the characteristics of the user sample, the test activities comprised, the results of these exercises and a set of recommendations to improve the overall user experience and usability of the websites tested.

2.2 Methodology

Methodologies based on user-centred design principles establish usability testing as a necessary milestone in the design process. The methodology applied to these usability testing activities have been determined in accordance with ISO 13407I (Human-centred design processes for interactive systems) that establishes the need for user based evaluations on late stages of a project.

Although some of the activities and metrics of the test are quantitative, its nature is essentially qualitative, due to the scope of the users' sample and the type of findings expected.

The test was composed of three types of activities. The first set was related to users' immediate recognition of the different elements of the interface. The goal was to detect first layer usability issues that have to do with the design of basic interactive elements such as links, icons, labels, etc.

The second group of activities was intended to measure users' performance and efficiency. In order to measure this parameter, users were required to perform several tasks that would be likely to be done by target users on a daily basis. Both the rate of success in achieving the tasks and the problems encountered during the process provided valuable information on the usability of the whole sequence and flow.

Finally, the third set of questions was intended to register the personal and subjective perception of the user, after he or she was exposed to the user interface. This activity, which was conducted trough a five scale point Likert Questionnaire, brought information on issues that are not directly related to objective usability, but with the subjective perception of it. Subjective usability is also a relevant issue: it reflects the level of motivation and expectations a user has on a given system, and it also influences directly on the overall performance.

The usability test was conducted on two different locations: Indra's facilities (10-14 Nov. 2003) and the Department of Civil and Geodetic Engineering at the University of Ljubljana (12 Jan. 2004).

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 9

2.2.1 Users

Both websites are targeted to researchers and scholars all over the world, interested in the subject of informatics in construction (and related issues), specifically those who are publishing research on the subject or may want to have access to it. Cultural issues are not likely to be relevant to this test, since the research community is used to retrieve information from international sources.

Nevertheless, the test has been conducted in more than one country to avoid possible confusion between issues derived from cultural particularities and those derived from usability problems.

One element taken into account has been the degree of familiarity with the interface that each user may have. The overall performance would be different between those who know the website and have used before and those who's first contact is happening during the test. Therefore, one variable in selecting the sample has been familiarity with the website.

	User 1	User 2	User 3	User 4	User 5	User 6
Name	Cristina Parrilla	Carlos Estevez	Isabel Ponce de León	Matej Smid	Etiel Petrinja	Tomaz Parlar
Age	24 years old	24 years old	26 years old	28 years old	25 years old	25 years old
Previous exposure to research in construction informatics	Low (studying Civil Engineering)	Medium (working as a civil engineer)	Low (Draughtswoman)	Low (SCIX administrative tasks)	Medium-High (Civil Engineering researcher)	Medium-High (Civil Engineering researcher)
Experience using computers	Low (5 years – only for student work)	High (14+ years)	High (6+ years – personal and professional)	High (15+ personal + profesional)	Very High (10+ programer level)	High (10+ personal + profesional)
Experience with other interactive equipment	Low	High	High	Medium	Medium	Medium

Some other issues such as technological literacy, age and field of study have been also recorded for guidance when analysing results, but was not relevant for the selection of the sample.

2.2.1.1 Users (Madrid)

Cristina Parrilla (http://itc.scix.net)

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 10

Cristina parrilla is a 24 year old girl studying Public Works, a technical degree in the field of public engineering.

She has been using computers since she started college, 5 years ago. Her experience with computers is broad but not deep: she usually browses the internet when she needs information for her assignments or for personal matters such as web based email, news, etc. She usually visits Spanish websites, since her level of English not advanced.

Cristina has never visited ITC, and she is not yet familiar with the concept of online libraries or technical journals.

Carlos Estevez (http://itc.scix.net)

Carlos is a 24 year old production manager in public works. After his graduation on civil engineering, he began working for a company where he is assigned to a public work that he has to manage.

He has a very broad and extensive experience with computers. He began using them 14 years ago, and he considers himself part of the so called Spectrum Generation: those guys who grew up with a Sinclair Spectrum 48k as their main toy.

Carlos has never visited before ITC, although he is familiar with the concept of research on the topic.

Isabel Ponce de Leon (http://itcon.org)

Isabel is a 26 year old draughtswoman who works for an Architecture firm. She studied a professional module specialised on Buildings and public works. She hopes that, after a couple years, she will be promoted and will work as a "projectist", being responsible for choosing materials in constructions.

Isabel is a technology savvy person. She started using computers 6 years ago, and she has been using them intensively both for work and at home. She is very familiar with databases and with searches on the internet, very important for her job.

She has not been exposed to ITC before, but she will start using such repositories and journals when she gets her promotion.

2.2.1.2 Users (Ljubljana)

Mateja Smid (http://itcon.org)

Mateja is a 28 year old person who has a great level of involvement with the SCIX project. Even thou she studied sociology, she works at the University of Ljubljana taking care of all the administrative issues involved in the SCIX project.

Besides being very familiar to the SCIX project, Mateja is also a technological savvy person. Her first computer was a Sinclair Spectrum, fifteen years ago. Thus she doesn't have any problem in acquiring and getting used to new interactive models or languages.

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 11

Etiel Petrinja (http://itc.scix.net)

Etiel is 25 year old. He lives in Ljubljana (Slovenia) where he studied Civil Engineering. Right now, he is completing his graduate studies to get his Ph.D. At the same time, he is a researcher within the Chair of Construction Informatics in the University of Ljubljana.

Etiel is somewhat familiarized with journals and research on construction informatics. He used itc.scix.net several times before the current usability test. His knowledge of computers is broad and deep, to the extent of knowing some programming languages (C++, Turbopascal, ...)

Tomaz Pazlar (http://itc.scix.net)

Tomaz is 25 year old. He lives in a town a bit far from Ljubljana and drives every day to go to work. He works as a "young researcher" in the Chair of Construction Informatics at the same time he is working on his doctoral thesis. His field of study is Civil Engineering, specializing in Structural Engineering.

Tomaz is often in contact with online resources for research on his field. He knew SCIX before but had almost no exposure to it. Besides, his level of computer literacy is rather high: he started using computers 10 years ago and use them on a daily basis.

2.2.2 Protocol

A usability test has to follow a consistent and fixed protocol, followed identically with every user. Altering this protocol may cause distorted findings that would difficult the obtainment of valid information.

These are the activities/milestones followed in this Usability Test:

1. Presentation: the test facilitator introduces himself and explains the purpose of the test.

2. Socio-demographic data questionnaire: the facilitator fills the form in a semi-structured interview with the user.

3. Introduction and explanation: The facilitator explains what exercises will be done in the test. He or she introduces the user to the "think aloud protocol", making sure that the user understands its importance. Assurance that the user understands that the test is about evaluating the interface and not himself / herself.

4. Recognition exercises: (see results section for detail on the exercises conducted)

5. Performance exercises: (see results section for detail on the exercises conducted).

6. Subjective perception questionnaire: (see results section for detail on the exercises conducted).

7. Thank the user and farewell: the facilitator expresses the organizations' gratitude and explains the usefulness of the activities. A small gift certificate is given to the user.

SCIX
SCIENTIFIC INFORMATION EXCHANGE

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 12

2.3 Results

2.3.1 ITC.SCIX.NET

Recognition: itc.scix.net

Homepage *Describe this homepage*	Purpose of the site is difficult to get at a glance for first time users (more than 1 minute for U1 and U2) Low visual association between images and the site's topic among first time users Left navigation menu barely perceived on initial exploration
available functionality *What is this site for?* *What can you do here?*	Dificulties in getting "what you get from it". Some answers were: links to resources, news on the sector. Unclear wether content would be free, under registration or paid. Low level of expectation on how the real content will be offered (links, abrstacts, complete pdf,...?)
bottom icons *What do you think this picture means?* *What do you expect to happen when you click on them?*	Registration icon misunderstood or not understood (neither the meaning nor the state)
my favourites *What do you think this link means?* *What do you expect to happen when you click on it?*	No usability problems found.

Recognition issues are related to initial perception of the site by users. These issues determine the first behaviors of the users in the site. Wether users will continue using the site or will abandon is sometimes determined by the some of these aspects.

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 13

SCIX
SCIENTIFIC INFORMATION EXCHANGE

Performance: itc.scix.net

Search *Find a specific paper* *Find papers on a specific topic*	No relevant problems in main (simple) search. Users get a bit overwhelmed when trying advanced search (too complex at a glance) Downloadable result of advanced search perceived as an error by most users. Controls for browsing along the results are barely noticed
Contribution *Upload a paper (how to...)*	Contribution procedure difficult to understand for first time users / contributors.
Information *Find out what the review and acceptance procedures are*	Users take some time to scan text to find what the review procedure is.
Registration *Sign in*	No relevant problems in signing in. Benefits / differences between being registered and not being so are not perceived at a glance.

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 14

Subjective perception: itc.scix.net

	completely agree	somewhat agree	Neither agree nor disagree	somewhat disaagree	completely disagree
Ease of use *I find the website easy to use*		xxxx			
Message clarity *I find the messages and texts of the site clear and easy to understand*	x	xxx			
Ubicability *I think it's easy to know where you are inside the website at any moment*	xx	xx			
Functionality *I find easy to know what I can do and what functions are available at any point of the website*	x	xx	x		
Design *Overall, I find the design of the website nice and pleasant*		xx		xx	
Usefulness *I find this website useful*	xx	x		x	
Comparative value *This website is better than other similar sites that I know*	x		xxx		

Responses to subjective perception questions should always be taken with "a grain of salt". They should be accepted as orientative but never as objective data.

By nature, all responses tend to be biased and more positive than the real opinions users may have. Therefore, a good way to interpret them is to lower all the responses one degree.

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 15

2.3.2 ITCON.ORG.NET

Recognition: itcon.org

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 16

Homepage *Describe this homepage*	Low visual association between images and the site's topic among first time users
available functionality *What is this site for?* *What can you do here?*	Search functionality barely perceived for first time users Not clear wether the site works under paid subscription (at a glance)
bottom icons *What do you think this picture means?* *What do you expect to happen when you click on it?*	Registration icon misunderstood or not understood (neither the meaning nor the state) Browse buttons are recognized but with some confusion on the meaning of each one.
Top links "ITcon" and "papers" *What do you think these links mean?* *What do you expect to happen when you click on them?*	Wrong expectations on the behavior of these links: not recognized as links at a glance, and not expected to be two separate links. Second link ("papers" in the detail) leads to confusion (is it a link or a status indicator?)

Recognition issues are related to initial perception of the site by users. These issues determine the first behaviors of the users in the site. Wether users will continue using the site or will abandon is sometimes determined by the some of these aspects.

Performance: itcon.org

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 17

Search *Find a specific paper* *Find papers on a specific topic*	Help text in simple search tends to confuse users on the usage instructions instead of clarifying them. no other relevant usability problems.
Contribution *Upload a paper (how to...)*	Contribution procedure difficult to understand for first time users / contributors.
Information *Find out what the review and acceptance procedures are*	Acceptance procedures difficult to find (users needed to dig dor more than 1 minute in all cases)
Registration *Sign in*	No relevant problems in signing in. Benefits / differences between being registered and not being so are not perceived at a glance.

Subjective perception: itcon.org

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 18

	completely agree	somewhat agree	Neither agree nor disagree	somewhat disaagree	completely disagree
Ease of use *I find the website easy to use*	x	x			
Message clarity *I find the messages and texts of the site clear and easy to understand*		x	x		
Ubicability *I think it's easy to know where you are inside the website at any moment*	x	x			
Functionality *I find easy to know what I can do and what functions are available at any point of the website*	x	x			
Design *Overall, I find the design of the website nice and pleasant*	xx				
Usefulness *I find this website useful*	x		x		
Comparative value *This website is better than other similar sites that I know*		x	x		

Responses to subjective perception questions should always be taken with "a grain of salt". They should be accepted as orientative but never as objective data.

By nature, all responses tend to be biased and more positive than the real opinions users may have. Therefore, a good way to interpret them is to lower all the responses one degree.

2.4 Recommendations

2.4.1 Short Term (Necessary Corrections)

1. Group around three main blocks: "search", "contribute" and "about".

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 19

The mental model of the users is conformed aroun the main three things they can do: search for papers, upload their papers and eventually know more about the site and who is behind it.

Grouping information and functionality around tese three categories would reinforce the purpose of the site to users' eyes and would reduce confusion on available funcionality, as well as procedures and other information related to contributions.

2. Include a short descrition of the page in the homepage.

Incorporating a short and clear explanation, in one paragraph of what the site is about, and remarking it visually (different color for text and background) would help first time visitors to undestand the porpuose and functionality of the page.

Besides, such a measure would improve the efectiveness of the results of people searching for the websites on Google and other search engines.

3. Give priority to searches, starting from the homepage.

Search for papers is the main task to be performed on both websites. Therefore, it cannot have a secondary placement. Placing the search engine in the homepage we would make users' tasks easier and would reinforce the message of the site.

4. Use visual resources that relate to the topic of the sites.

A moderate and savy usage of pictures and pictograms related to construction and informatics, both sites would transmit more efficient and internationally the topic they cover.

5. Use standard icons for main functionality

Using standad icons for main functions would allow a faster and more effective recognition of available functionality.

The main functionality of the site (search, upload, download, login, logout, …) is fairly standard and there are many collections of icons to choose from that are commonly recognized.

6. Explain the benefits and differences between being registered and being anonymous.

A clear and visible policy on why register, along with a table of differences between being registered and not would help understand the value of this function, and would reduce confusion. If benefits for users are still unclear, it is recommended to allow free access to papers.

2.4.2 Long Term Improvements

1. Create a wizard for the submission procedure

Designing a step-by-step procedure that explained each information at the right moment would make the submission procedure easier for first time submitters.

2. Create an info page for first time users

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 20

First time users of digital journals don't have a strong mental model of how these services work and how are they structured.

An explanatory page devoted to this type of users would strenghten their mental model and would improve overall usability of the site.

3. Provide summaries of the page in different languages

Providing a single page sumary of the page in all EU languages would allow many users to better understand the purpose and function of the site at a glance.

4. Simplify the search interfaces to adjust to internet standards

Users' mental model for searching for documentation is closer to internet regular (simple and advanced) searches than to database, bolean or library standards.

Using simple html form elements, the same level of efficiency and precision can be achieved without sticking to library style queries and without compromising metadata structure of the papers.

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 21

3. COMPARISON OF THE PILOTS AGAINS OSI CRITERIA

This section evaluates SciX based on a set of criteria published by the Open Software Institute (OSI) Guide to Institutional Repository Software (http://www.soros.org/openaccess/software).

That document compares the following software:

- ARNO
- CDSWare
- Dspace
- Eprints
- Fedora
- I-Tor
- MyCoRe

We have used the tests there to evaluate the SciX software. It should be noted, however, that SciX is much more than an institutional repository and that it has numerous unique features.

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 22

	SciX	ARNO	CDSware	DSpace	Eprints	Fedora	i-Tor	MyCore
Technical Specifications								
1 Standards Information								
1.1 OAI_PMH version supported	OAI-PMH 2.0	OAI-PMH 2.0	OAI-PMH 2.0	OAI-PMH 2.0	OAI-PMH 2.0	OAI-PMH 2.0	OAI-PMH 2.0	OAI-PMH 2.0
1.2 Z39-50 protocol compliant	No	No	No	No	No	No	No	No
1.3 Open Source License	Yes	TBD	GNU-GPL	BSD	GNU-GPL	MPL	GNU-GPL	GNU-GPL
1.4 Latest Version release date	Feb 2003	Dec-03	Apr-02	Aug-03	Mar-02	Dec-03	Aug-03	Oct-03
1.5 Latest Version No.	1.1	1.0	0.0.9	1.1.1	2.2.1	1.2	1.1.4	1.0
2 Hardware								
2.1 Minimum Hardware requirements	No specific requirements	No specific requirements	No specific requirements	No specific requirements	No specific requirements	No specific requirements	No specific requirements	No specific requirements
2.2 SAN support	No			Yes	Yes	No		
3 Software								
3.1 Operating System	Unix/Linux/Windows	Linux/Solaris	Linux/Solaris	UNIX/MacOSX/Windows	GNU/Linux/Solaris	UNIX/MacOSX/Windows	Linux/Windows	AIX/Linux/Solaris/Windows
3.2 Programming Language	Perl	Perl	Python/PHP	Java	Perl	Jva	Java	Java
3.3 Database	WODA	Oracle 8i	MySQL	PostgreSQL	MySQL	MySQL/MCKoi/Oracle	MySQL/Oracle	MySQL,PostgreSQL,XML:DB compliant;commercial DB's
3.4 Web Server	Apache/Any	Apache	Apache/PHP,Python	Any	Apache 1.3	Tomcat 4.1	Jetty	Apache
3.5 Java servlet engine	N/A			Any	N/A	Tocat 4.1	Jetty	Any
3.6 Search Engine	WODA	Unix & SQL command line	Cdsware	Lucene	N/A	Database	Lucene	Via JDBC & XML:DB
3.7 Other			WML:Website meta language	OAICat	N/A			Apache Ant build tool

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 23

SCIX SCIENTIFIC INFORMATION EXCHANGE

	SciX	ARNO	CDSware	DSpace	Eprints	Fedora	i-Tor	MyCore
4 Clients Supported	All web browsers, Office 2003,Citation management software,SOAP Clients	Any browser	All HTML4.0 Clients	All web browsers	Netscape,Mozilla,IE,Lynx	Web browsers & SOAP Clients	All HTML4.0 Clients	All web browsers
5 Staff Requirements								
5.1 Unix Sys Admin	For setup	Yes	Yes	Yes	Yes	For setup	Recommended	recommended
5.2 Java Programmer	Not required	No	No	Recommended	No	Recommended	No	recommended
5.3 PERL Programmer	Yes	Recommended	No	No	Recommended	No	No	No
5.4 Python Programmer	No	No	No	No	No	No	No	No
6 Installed base								
6.1 No. of Installations	~15	7	7	15+	106	20	10	10
6.2 Geographic Coverage	Western , Central Europe	Netherlands	Eurpoe & US	Worldwide	Worldwide	Worldwide	Germany & Spain	Germany & Sweden

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 24

Repository & System Administration								
7 Set-up/Installation								
7.1 Automated installation script	Yes	Yes	Yes	Yes	Yes	Yes	Yes	yes
7.2 System update script	Yes	Yes	Yes	Yes	Yes	Yes	No	Via CVS Repository
7.3 Update sytem update w/o overwriting customized features	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
8 Module-level APIs	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
9 User registration,authentication & password administration								
9.1 Password administration	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
9.1.1 System Assigned passwords	No	No	Yes	Yes	No	No	no	
9.1.2 User selected passwords	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
9.1.3 Forgotten password function	Yes	Yes	Yes	Yes	Yes	No	no	
9.2 User registration/verification; Other security functions	User services	LDAP and/or ARNO registry	MySQL Table/Apache ACL	Email/X.509	MySQL Table	No	No	RDBMS table
9.2.1 Edit User Profile	Yes	No	Yes	Yes	Yes	No	yes	
9.3 Limit access by User Type	Yes	Yes	Yes	Yes	Yes	Yes	no	
9.4 Multiple Authentication Methods	No	Yes	Yes	Yes	No	Yes	no	
9.5 Limit Access at File/Object level	Yes	Yes	Yes	Yes	Yes	No	Yes	no
10 Content Submission Administration								

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 25

10.1 Define multiple collections within same instance of system	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
10.1.1 Set different submission parameters for each collection	No	Yes	Yes			yes		
10.1.2 Home page for each collection	Not called home page	No	Yes	Yes	No	Yes	no	
10.2 Submission Stages	Submit, Revise		Submit,modify,revise.approve etc	Assemble,pending,approved		Ingets,create,modify.activate,deactivate	Yes	no
10.2.1 Segregated submission workspace	Yes (if set up)	Yes	Yes	Yes	Yes	Yes	Yes	
10.2.2 Submission Roles	Submitter	Contributors,editors,administrators,site managers	Submitters,moderators,reviewers,approvers,administrators	Submitters,Reviewers,Approvers,Editors	User,Editor,Administrator	Administrator	Yes	
10.2.3 Configurable submission roles within collections	No	Yes	Yes	Yes		No	yes	
10.3 Submission support								
10.3.1 email notification for submitters	No	Only during registration	Yes	Yes	Yes	No	Yes	no
10.3.2 email notification for content administrators	No	Yes	Yes	Yes	Yes	No	Yes	no
10.3.3 Personalized system access for registered users	Yes	Yes	Yes	Yes	Yes	No	Yes	no
10.3.3.1 view pending content submissions	Yes	Yes	Yes	Yes	Yes	No	N/a	No
10.3.3.2 view approved content	No	Yes	Yes	Yes	Yes	No	N/a	No
10.3.3.3 View pending content administration tasks	Yes	Yes	Yes	Yes		No	N/a	no
10.3.4 Distribution licence								
10.3.4.1 Request distribution license	No	No	No	Yes	No	Yes	no	

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 26

10.3.4.2 Store distribution license with content	No	No	No	Yes	No	Yes	no	
11 System generated usage statistics and reports								
11.1 System-generated usage statistics	Yes	Yes	No	Yes	No	Yes	Yes	no
11.2 Usage reports	Yes	No	No	Yes	No	No	Yes	no

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 27

Content Management								
12 Content Import/Export								
12.1 Upload Compressed files	Yes	Yes	Yes	Yes	Yes	Yes	Yes	no
12.2 Upload from existing URL	No	Yes	Yes	No	Yes	Yes	Yes	no
12.3 Volume import for objects	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
12.4 Volume import for metadata	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
12.5 Volume export/content portability	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
13 Document/Object Formats								
13.1 Approved file format function	Setup	Yes	Yes	Yes	Yes	No	No	No
13.2 File formats ingested	All	All	All	All	All	All	All	all
13.3 Submitted items can comprise multiple files	No	Yes	Yes	Yes	Yes	Yes		yes
14 Metadata								
14.1 Basic metadata schema	Unqualified Dublin Core	Dublin Core	Standard Marc21	Qualified Dublin Core	Dublin Core	Dublin Core	Any	Qualified Dublin Core
14.2 Support for extended metadata	Yes	Yes	Yes	Custom	Yes	Yes	Any	any
14.3 Metadata review support	Yes, edit, modify	Yes	Yes	Yes	Accept,Edit,Bounce(require changes),Delete	No	No	no
14.4 Metadata export	Yes, several formats	Yes	OAI-Marc export	Custom XML Schema	Custom XML Schema	Yes	Yes	Yes
14.5 Disallow metadata harvesting	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
14.6 Add/delete metadata fields	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
14.7 set default values for metadata	Yes	Yes	Yes	Yes		No	Yes	

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 28

14.8 Support Unicode character set for metadata	Quoted	Partial	Yes	Yes	Yes	Yes	No	Yes
15 Real-time updating and indexing of accepted content	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 29

Dissemination								
16 User Interface								
16.1 Modify interface "look & feel"	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
16.2 Apply a custom header/footer to static or dynamic pages	Yes	No	Yes	No	Yes	Yes	Yes	Yes
16.3 Supports multiple language interfaces	yes	No	Yes	Yes	Yes	Yes	Yes	Yes
16.4 End user document folders	No	No	Yes	No	No	No	Yes	
16.5 Discussion forum support	About items	No	No	No	Yes	No	Yes	No
17 Search Capability								
17.1 Full Text	No	No	Yes	Yes	No	No	Yes	No
17.1.1 Boolean logic	Yes	No	Yes	No	No	No	Yes	
17.1.2 Truncation/wildcards	Yes	No	Yes	No	No	No	Yes	
17.1.3 Word stemming	Yes	No	No	No	No	No	No	
17.2 search all descriptive metadata	Yes	No	Yes	Yes	Yes	Yes	Yes	
17.2.1 Boolean logic	Yes	No	Yes	Yes		No	Yes	Yes
17.2.2 Truncation/wildcards	Yes	No	Yes	Yes		Yes	Yes	
17.2.3 Word stemming	Yes	No	No	Yes		No	Yes	Yes
17.3 Search selected metadata fields	No	No	Yes	Yes	Yes	Yes	Yes	Yes
17.4 Browse								
17.4.1 By Author	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
17.4.2 By title	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
17.4.3 By issue date	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
17.4.4 By subject term	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
17.4.5 By collection	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
17.5 Sort Search Results	Yes							
17.5.1 By Author	No	No	Yes	No	Yes	No	Yes	Yes
17.5.2 By title	No	No	Yes	No	Yes	No	Yes	Yes

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 30

17.5.3 By issue date	No	No	Yes	No	Yes	No	Yes	Yes
17.4.4 By relevance	No	No	No	No	No	No	Yes	
17.4.5 By other	Yes, by any field above	No	Any metadata field	No	Yes	No	Yes	Yes
18 Indexed by Google/other search engines	Yes	No	Possible	Yes		Possible	Yes	possible
Archiving								
19 Persistent document identification								
19.1 System-assigned identifiers	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
19.2 CNRI handles	No	No		Yes	Yes	No	No	Yes
20 Data preservation support								
20.1 Defined digital preservation strategy	No	No	Yes	Yes	No	Yes	No	No
20.2 Preservation metadata support	No	Yes	Yes	Yes	No	Yes	No	No
20.3 data integrity checks	No	No	No	MD5 checksum	MD5 checksum	SIP schema validation	No`	MD5 checksum
21 Object history/version control	Yes, metadata only	Versioning system for both objects & metadata	Versioning system	ABC harmony data model	Some	Linear	No	No

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 31

System Maintenance								
22 System support								
22.1 Documentation/manual	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
22.2 Listserv	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
22.3 Bug track/feature request system	No	No	Yes	Yes	No	Yes	Yes	No
22.4 Formal support/help desk	No	No	For fee	No	No	Yes	No	No

4. FULFILMENT OF OBJECTIVES

This section revisits the project goals and provides a brief answer (**in bold**).

4.1 ENABLE EFFICIENT ACCESS TO SCIENTIFIC RESULTS

- speed up the process from submission to final publication: **yes, as proven by ITcon.**

- allow a more rich content (multi-media**): SciX supports arbitrary files**.

- provide readers more efficient mechanisms for retrieving publications of interest: **yes, proved by the numerous DLs running on the SciX platform.**

- increase readership through the abolition of barriers such as subscriptions: **Yes, proved by the CUMINCAD and numerous other DLs.**

4.2 PROVIDE ACCESS TO NON-SCIENTISTS

- increase the number of readers from industry and from smaller universities:

- reduce time spent in searching and retrieving relevant information:

- **substantial percentage of users is coming from the .com domains; har to prove, however, they are not academics.**

4.3 RE-ENGINEER THE PUBLISHING PROCESS AND PERFORM A SOCIAL-ECONOMIC ANALYSIS

- scientific publications not as a commodity to be sold or archived as an essential part in a larger scientific communication process

- solutions based on the premise of globally free information on the world wide web, thus side-stepping some of the traditional intermediaries altogether

- **successful and well cited re-engineering done in WP1**

4.4 AUTOMATE REPOSITORY MANAGEMENT THROUGH SELF ORGANISATION

- create clusters of data automatically:

- clusters of similar papers, which should be of interest to the reader as well:

- **implemented in WP4, for example related to clustering conference papers into sessions.**

4.5 SIMPLIFY USE THROUGH INTELLIGENT PERSONALISED AGENTS

- user profiling system:

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 33

- automatic notification on new papers matching the profiles' interest:

- selective searches:

- **implemented through the agent interface in the 'works' service.**

4.6 INVESTIGATE LEGAL, SOCIAL AND PSYCHOLOGICAL ISSUES

- done, particularly in the deliverable D2 and D4.

4.7 DEVELOP BENCHMARKING METHODS FOR SCIENTIFIC JOURNALS

- done, particularly in the deliverable D2 and D4.

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 34

5. SWOT ANALYSIS OF SCIX

5.1 METHODOLOGY

SWOT analysis is a technique commonly used in business circles to assist in identifying strategic issues for a company or organisation. The predictive capabilities of the technique come about from the consideration of each system's strengths and weaknesses in the context of the environment, which is seen to present opportunities and threats. The intention is to determine how each product will fare in the light of changes taking place around it.

In the context of SciX, the SWOT analysis is relevant to understanding the issues affecting the future penetration of the platform, and the continuing contribution it makes to the main SciX objectives discussed in the previous section.

5.2 STRENGTHS

- A clear specification of the processes of scientific knowledge management, which covers the main elements of the processes of knowledge creation, dissemination and use, which has been subjected to public scrutiny.

- A clear analysis of the economics of open publishing.

- The SciX core platform has been defined and created in the context of the process models providing for a clear understanding of its requirements.

- The platform is easy to set up, reducing the barriers to implementation.

- The nature of the underlying WODA platform makes it relatively easy for a developer understanding WODA development to add new APIs and formats quickly.

- The platform is OAi compliant, allowing it to become part of the growing infrastructure of collaborating institutional publication repositories.

- SciX is now used to support a number of established services, including several repositories including CUMINCAD, ELPUB, ITC, and IAPS, an established electronic journal, ITCon, and several conferences, including IAPS2004.

- Support for journal and conference management is integrated with the core repository software. The ability to publish conference and journal papers within the same system that is managing the submission and review process makes SciX attractive for those purposes. Typical software offerings currently support one or other of the functions, but not both.

- The Value Added Publication facilities promise to enhance the exploitation and impact of scientific research

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 35

5.3 WEAKNESSES

- The process model does not cover well the industry use of scientific knowledge as supported by the Value Added Services.

- The current platform does not integrate well with main institutional IT infrastructures based on relational database technology, which require more rigorous normalization of data structures.

- SciX is built on WODA, which is a Perl-based platform. For other parties to contribute to the further development of SciX as a free open-source project, there will be a significant learning curve involved compared to a system based on a mainstream relational database product.

5.4 OPPORTUNITIES

- The process models provide the opportunity to make further research contributions to the areas of open publishing and scientific knowledge management.

- The SciX platform may be extended with more powerful knowledge management techniques and become the focus for future work in this area.

- Successful commercial services may be built on the Value Added Publication facilities of SciX.

- The integration of journal/conference management and publication repository can be used to encourage the adoption of the SciX platform.

- Because of its existing penetration in this construction research area, SciX can become the system of choice in this arena.

- The SciX platform could benefit from closer integration with work in the area of open citation indexing (such as the Open Citation project) in order to increase the chances of SciX-based journals and conferences being recognized.

5.5 THREATS

- There are several other solutions available addressing either the digital repository function or the journal and conference management functions. As interoperability improves, through standardization efforts including the OAi, combinations of such existing solutions may become popular.

- There are significant barriers to the adoption of open publishing generally, which are driven by the interests of the publishers and by the popular perception that particular established publishers are essential to the recognition of the quality of a journal.

- SciX does not yet have the same degree of penetration as DSpace and ePrints.Org in the institutional repository area.

D11: Overall assessment and evaluation report
version of 30-Mar-04
page 36

- The ISI citation index has a strong hold over the perception of the quality of journals, and presents a significant barrier to the creation of new, open-access journals.

- National research evaluation methods may fail to recognize the importance of open, electronic publishing in assessing the impact of research.