E

WIPO

SCIT/SDWG/3/6
ORIGINAL: English
DATE: April 29, 2003

WORLD INTELLECTUAL PROPERTY ORGANIZATION
GENEVA

STANDING COMMITTEE ON INFORMATION TECHNOLOGIES

STANDARDS AND DOCUMENTATION WORKING GROUP

Third Session
Geneva, May 5 to 8, 2003

PROGRESS REPORT BY THE IPDL ELECTRONIC TASK FORCE ON STANDARDS DEVELOPMENT

Document prepared by the Secretariat

Introduction

1. Over the past several years, many developments have taken place within the Intellectual Property Digital Library (IPDL) community. Various organizations have made new or next-generation IPDL systems available on the Internet. Existing systems are becoming more and more user-friendly, and intellectual property Offices are also expected to publish public data using the Internet.

2. The International Bureau (IB) of WIPO has completed a project to publish certain data in an Internet-based environment aimed at facilitating the development of voluntarily accepted policies and standards concerning the interoperability of IPDLs established by the WIPO Member States. As a result of this effort, the IB has organized two IPDL Standards Workshops and has published several documents. These documents not only outline a generalized role for the IB in the global IPDL environment but also make certain recommendations concerning standards suitable for use in the IPDL environment.

3. During its second session in December 2002, the Standards and Documentation Working Group (SDWG) of the Standing Committee on Information Technologies (SCIT), the IB was requested to research certain areas of interest, and present recommendations to the SDWG.

4. It has not proved possible to organize the third IPDL Standards Workshop during the short period of time between the second and third SDWG sessions. However, the IB has undertaken the task of researching the subjects as requested, and hereby presents its findings in this document, including, for informational purposes only, some of the options available in these areas.

5. Several specific recommendations are also included in this document. Those recommendations are made independently by the IB since they are primarily procedural and concern matters to be considered by the Task Force.

Task Force Procedural Matters

6. As approved at earlier sessions by the SDWG and the SCIT Plenary, the IB continues to:

(a) coordinate the efforts of related Task Forces established by the SCIT (in particular, the SCIT IPDL Electronic Task Force);

(b) develop preliminary working proposals related to IPDL standards and policies for consideration by Task Forces and the SCIT; and

(c) become engaged in the development and monitoring of standards and policies within the general Internet community that are relevant to IPDL systems.

7. In light of the ongoing need for this work, it is proposed that the SDWG recommend to the SCIT Plenary to extend the mandate of the IPDL Electronic Task Force for a period of one year, and that the Task Force continue to deliberate upon items requested by the SCIT Plenary and the SDWG.

Report of Task Force Activities

8. At the December 2002 session of the SDWG, the IPDL Electronic Task Force was requested to begin work on a set of basic policies for presentation to the SDWG at its third session. The Task Force was also requested to produce a sample statement concerning organizational commitments to the persistence of published electronic data objects. The IB will prepare preliminary materials on these topics, and submit them to the Task Force for review. Some of these materials are elaborated in the Annex to this document.

9. The SDWG is invited:

(a) to note the contents of this document and its Annex; and

(b) to consider and recommend to the SCIT Plenary to extend the mandate of the IPDL Electronic Task Force, as proposed in paragraph 7, above.

[Annex follows]

ANNEX

PRELIMINARY FINDINGS

Matters Relating to the Development of IPDL Policies

1. The objective of developing a set of basic IPDL policies is to ensure systems-level interoperability. The outcome of this exercise will be to provide a set of standards, selected or developed through the collaborative efforts of the IB and WIPO Member States, which intellectual property Offices may adopt in their development of new systems or the enhancement of their existing systems.

2. The IB has begun its work to develop introductory materials for the IPDL Electronic Task Force as a replacement for the work originally proposed to be done by the third IPDL Standards Workshop, in particular, focussing on matters relating to policies concerning organizational commitment regarding the permanence of electronic information.

Policy Recommendations Concerning the Permanence of Electronic Information

3. In keeping with the goal of developing simple and non-intrusive policies, the IB has researched several proposed permanence policies (primarily in conjunction with the adoption of permanent identifiers) adopted by various library and publishing organizations. In this connection, IPDL Standards Workshops have expressed particular concern regarding the ability of the intellectual property publishing community to inform users whether an electronic resource cited today will remain available in the future, retrievable from the same address, and unchanged in content. The issue at hand is to develop a policy that addresses these concerns, while remaining sufficiently flexible for use with any permanent identifier scheme.

4. The simplest scheme most suited to the particular needs of the intellectual property community was that put forward by the United States National Library of Medicine. This approach identifies three core categories of permanence for electronic resources:

(a) Identifier Validity – the extent to which a user can be assured that a given name, number or other identifier will not be changed or assigned to another resource;

(b) Resource Availability – the extent to which a given resource will remain accessible; and

(c) Content Invariance – the extent to which the content of a given resource and the links it contains will remain unchanged.

5. This approach requires the policy-based rating of intellectual property office materials, along with some shared policies or understandings concerning the management of hardware systems that support resources rated as “Permanent.” The different ratings for the three core categories are given below.

Identifier Validity (IV)

(a) *Undefined or Transient* (Either no rating has been undertaken or the identifier could be changed or reassigned.)

(b) *Guaranteed* (The identifier will not be changed or reassigned to another resource, example: a published patent document.)

Resource Availability (RA)

(a) *No Guarantee* (The resource may become unavailable at any time. Example: temporary announcements.)

(b) *Permanently Available* (Accessibility is guaranteed. This rating implies a commitment to archive the resource, example: Published patent documents.)

Content Invariance (CI)

(a) *Unrated* (No rating has been undertaken or no guarantee has been made.)

(b) *Dynamic* (The content may be replaced, corrected and revised. Internal and external links could change, example: WIPO WWW home page.)

(i) Open
(ii) Closed

(c) *Stable* (The content is subject only to correction and minor additions. Internal links will be updated, example: PCT Applicants Guide.)

(i) Open
(ii) Closed

(d) *Invariant* (The content is static, example: Published patent documents.)

The optional sub-elements “Open” and “Closed” may be assigned to “Dynamic” and “Stable” resources to indicate whether the resource will grow in regular increments [open] or is no longer growing [closed].

6. This scheme is simple, and permits the simple categorization of data products on a local basis. Furthermore, a default condition of either IV=Undefined, RA=No Guarantee, CI=Unrated, can be automatically applied immediately to the backlog of data products, with further categorization at a later date.

7. A proposal will be made to the Task Force that an overall policy concerning permanence be drafted with published guidelines for the categorization of intellectual property information. It will also be proposed to the Task Force that a condensed rating scheme based upon the model mentioned above be developed, which should be suitable for inclusion in electronic metadata records of the type found in most proposed permanent identifier schemes. This will allow permanence information associated with intellectual property information resources to be easily retrievable by interested users.

Matters Relating to the Development of IPDL Technical Standards

8. The situation for the development of IPDL Technical Standards is the same as that for IPDL Policies. The IB is developing basic materials for submission to the Task Force. It is anticipated that discussions will soon begin on matters related to:

(a) support for one or two common data formats;

(b) support for a common, easily implemented, language-independent search protocol;

(c) support for common search attributes;

(d) support for a common, easily implemented data retrieval protocol;

(e) support for common data retrieval attributes;

(f) optionally, support for shared user-authentication strategies and services;

(g) optionally, support for 'grid-based' computing architectures, such as GridIR (see http://www.grid.org/ and http://www.gridir.org/).

9. Item (a) listed in the preceding paragraph is currently being addressed through other initiatives in the IB and Offices of WIPO Member States. It will, however, be submitted to the Task Force for general comments. The IB has concentrated on items (b) through (f), and has been considering item (g) as part of SDWG Tasks Nos. 10 and 18. No discussion of item (g) will be presented in this document; however, basic information will be provided to the Task Force for its consideration in due course.

Search and Retrieval

10. It should be noted that Items (b) through (e) are generally considered together in the context of a search and retrieval protocol. The clear trend in the Information Retrieval (IR) community is foreseen towards customized degrees of search and retrieval functionality supported by the Web Services model. During the second session of the SDWG, discussions took place on the Web Services model as part of the presentation of document SCIT/SDWG/2/10. It became clear during these discussions that there is significant interest among WIPO Member States to promote some sort of simple interoperability between IPDL systems without resorting to complex, difficult-to-implement communications protocols. A useful solution would be one that utilizes the Web Services infrastructure while providing the power of custom search and retrieval protocols such as Z39.50. It seems desirable, in particular, to support the establishment of common data elements and criteria to achieve search interoperability across multiple servers that may use different data models, data managers and search engines.

11. The IB has researched this matter and will present to the Task Force information and specifications concerning Search/Retrieve Web Service (SRW) protocol, which will integrate access to various networked resources and promote interoperability between distributed databases by providing a common utilization framework. SRW is a Web Service-based

protocol whose underpinnings are formed by bringing together more than 20 years of experience from the collective implementers of the Z39.50 Information Retrieval protocol with recent developments in the web technologies arena.

12. SRW uses the Simple Object Access Protocol (SOAP) as the information exchange mechanism, and the Web Service Description Language (WSDL) for record description.

13. In contrast to the 18 native and extended services supported under classic Z39.50, SRW supports only one service (SearchAndRetrieve).

14. SRW is semantically equivalent to classic Z39.50, which preserves the experience of the Z39.50 community without the overhead of the standard aspects of Z39.50 requests and responses.

15. SRW defines a Web Service combining several Z39.50 features, most notably, the Search, Present, and Sort Services. Additional features/services may be added later or defined later as new Web Services.

16. SRW specifications may be found at:
http://www.loc.gov/z3950/agency/zing/srw/specifications.html

Permanent Identifiers

17. The second IPDL Workshop made a recommendation for the use of Archival Resource Keys (ARK) as a permanent identifier scheme for intellectual property offices. The SCIT Plenary and the SDWG indicated interest in a comparison between Digital Object Identifiers (another model of potential interest to IP offices) and the ARK scheme.

18. Rather than engaging in its own detailed review of these proposed mechanisms, the IB has researched the topic and discovered an excellent comparison report issued by the National Library of Australia at http://www.nla.gov.au/initiatives/persistence/PIcontents.html.

19. The IB shall present this report to the Task Force for its consideration.

[End of Annex and of document]