EUROPEAN DISTRIBUTED INSTITUTE OF TAXONOMY (EDIT)

WP5: Internet Platform for Cybertaxonomy

INFORMATION SCIENCE & TECHNOLOGY COMMISSION (ISTC)

3rd Meeting, October 1-2, 2008

Botanic Garden and Botanical Museum Berlin-Dahlem

MINUTES

Participants: Agnes Kirchhoff (BGBM), Alain Empain (NBGB), Alexander Kroupa (SMNS), Andreas Hammer (MFN), Andreas Müller (BGBM, only 1st day), Andreas Kohlbecker (BGBM, only 1st day), Anke Hoffmann (BGBM), Anton Güntsch (BGBM), Chris Sleep (NHML), Christoph Häuser (SMHN, only 2nd day), Chuck Miller (MO), Eric Danon (RBINS), Gavin Malarky (NHML), Gerrit Stegehuis (CBS), Gregor Hagedorn (Key2Nature, only 1st day), Henri Michiels (MNHN), Jean Van Onacker (NBGB), Karol Marhold (IBSAS), Kèlètigui Toure (RMCA), Lutz Suhrbier (FUB-INF, only 1st day), Mark Jackson (RBGK), Martin Pullan (RBGE, until 2nd day 10:00), Patricia Mergen (RMCA), Paul Richards (NHML), Regine Vignes-Lebbe (UMPC), Robert Tolksdorf (FUB-INF), Simon Chagnoux (MNHN), Sven Kullander (NRM), Walter Berendsohn (BGBM), Yde de Jong (UvA)

Proxy participants: Kèlètigui Toure (RMCA) for André De Mûelenaere
Apologies: Eduard Stloukal (CUB)

DAY 1: October 1st 12:45 – 18:30

General Information:

All presentations of the ISTC meeting will be available on the WP5 webpage (link http://wp5.e-taxonomy.eu/blog/2007/04/04/istc)

1. Welcome; Introduction of Participants (Walter Berendsohn)

2. Integration (WP5.1)

a. Progress report (Anke Hoffmann)

A. Hoffmann gave an overview of the ISTC membership and the current status of the Memoranda of Understanding (see presentation). She invited the ISTC members to use the forum within the ISTC webpage as a communication tool (http://wp5.e-taxonomy.eu/blog//node/132; http://wp5.e-taxonomy.eu/blog/node/84).

b. Presentation of the platform security/Shibboleth (WP5.7: Lutz Suhrbier)

L. Suhrbier explained the architecture of the security system Shibboleth (see presentation).

Discussion:

(G. Hagedorn) One of the mayor issues was the question if and how a „low end contribution“to a Shibboleth federation will work. For example would it be possible for small non institutional groups of taxonomists, driving a simple Website (Drupal site, Wiki), to use somehow the EDIT Shibboleth federation and its IdP (Identity Provider) or can Drupal act as IdP?

(W. Berendsohn) In these scenarios two things are to consider: Applications like a Wiki or Drupal-site are always a bit insecure, so the user accounts in these applications are a target for potential manipulation. Users contributed by such an insecure IdP would be rejected by services in a federation. So a secure user management is crucial.

(P. Richards) A possible solution for “low end contribution” could be allowing an application like Drupal to register users to an external IdP, attributes would automatically be added which allow the user to access the service.

(W. Berendsohn) This requirement for a “low end contribution” will be put on the agenda of activity 5.7. Implications of this: users of the “connected” guest SP (application) would potentially have no access to the other SPs of the federation.

(L. Suhrbier) Requirements for setting up an IdP are a server running Tomcat & Apache, the Shibboleth IdP servlet, ports 443 and 8443 open, and a connector to a database providing users and attributes.

(R. Tolksdorf) The Freie Universität Berlin is using such system which is open to invite users to access services. Users can be granted access to a service without turning them into a member of the university. It is a big benefit to have this freedom of choice.

(R. Tolksdorf) Legal questions on national level concerning personal data: Do countries prohibit sending data about employees around?
(L. Suhrbier) Which user attributers are sent around is just a matter of configuring the IdP attribute policies and can be adapted to meet these restrictions.

The main issue was if EDIT member institutions are interested in integrating Shibboleth, i.e. if WP5.7 should continue its efforts along the present course. Statements:

NHML London (P. Richards): is positive about Shibboleth.
NBGB Meise (A. Empain): agrees in general to Shibboleth, but would like to experiment with a prototype.
RBINS Brussels (E. Danon): conceptually agrees to Shibboleth as possible software package for web single sign-on meanly because it's a open-source solution but a test (prototype) bundle is really required to make practical tests locally (in our computer environment) and to be able to give better technical feedback and advice.
CBS Utrecht (G. Stegehuis): Shibboleth can be realized.
IBSAS Bratislava (K. Marhold): sees no problems that IBSAS adapts to Shibboleth.
RMCA Tervuren, (K. Toure): cannot decide for his institution and will discuss it with the responsible persons at RMCA. He supports Shibboleth solution in general, but testing is requested.
NRM Stockholm (S. Kullander): needs to discuss it with his IT department.
MfN Berlin (A. Hammer): at the moment not able to make security installations, because no capacity available.
Key2Nature Project (G. Hagedorn): widely spread community participation, should be integrated into Shibboleth, but agrees in general.
UvA Amsterdam (Y. de Jong): agrees to use Shibboleth, should be included into WP3, system should align with databases with more than 500 users such as FaunaEuropaea.
RBGK Kew (M. Jackson): at present no position can be taken due to an ongoing IT review process at Kew.
MO St. Louis (C. Miller): Missouri has its own existing security system (CSSO), but agrees that integration with Shibboleth is possible and represents a positive development.
RBGE Edinburgh (M. Pullan): at the moment at RGBE the decision is open which system will be used (open id or other system), does not want to support two systems, and has to wait for institutional decision. Personally agrees with Shibboleth, but accentuates the need to extend the federation beyond EDIT [comment W. Berendsohn: this is certainly the intention, e.g. in PESI and LifeWatch].
MNHN / UPMC (R. Vignes-Lebbe): Specific situation at MNHM/UPMC must be considered, she personally supports Shibboleth.
MNHN Paris (H. Michiels): agrees to Shibboleth federation if it this is not in the way of joining other federations as well

Shibboleth is a widely accepted security system. By joining the Shibboleth federation institutions would save time concerning IT administration. Implementation of security systems is rather a political and legal issue than a technical issue.

Summary: positive response, prototyping needed to make further decisions, which should be incorporated into work plan of WP5.7. More documentation (especially for non-technical people) is needed. Role of EDIT: lead of development, demonstration for other institutions also beyond EDIT.

c. Further ideas for inter-institutional agreements (Memoranda of Understanding)
(W. Berendsohn) MoU are important as an easy proof for integration. MoU can have a broader context for all EDIT members or be specific for small groups. Until now we have received too little response for the already established MoUs (ABCD, Drupal), and we are missing suggestions for new agreements. Participants are requested to ask within their institutions for new themes for agreements. In other workpackages inter-institutional agreements are also available, e.g. WP7 in terms of use of data and obligations of participating institutions, WP1 is working on MoU concerning gender balance.

Two new themes for MoU had been proposed:
(1) Agreement on a long term, platform independent, document format
A. Empain presented ODF as a standard for long term conservation and potential interoperability of documents (see presentation).

Results of discussion:
The participants could not agree on ODF as the only standard for archiving. Participants recommended that institutions should define a policy for long-term archiving, but not only of documents, but also of data. Preferably, the solution should be defined in a larger framework (national data centres etc.), but presently EDIT is probably not able to provide such an infrastructure.

(2) Agreement on the utilization of available global name and classification services
P. Richards presented the proposal for a MoU on the utilization of available global name and classification services (see presentation).

Results of discussion:
The participants support this proposal, with some modifications, e.g. concerning the selection and citation of examples. Further recommendations should be sent to P. Richards, who volunteered to circulate a new draft.

3. Version 1 of the Internet Platform for Cybertaxonomy

a. Introduction: What will be delivered in November (Walter Berendsohn)

Version 1 of the Internet Platform for Cybertaxonomy is an EDIT project deliverable due by the end of November 2008. It will include both an overview of projected platform functionality and links, as well as a number of functional products and prototypes, that are demonstrated under points 3b. to 3h.

b. Platform presentation on the Internet (WP5.2: Andreas Müller)
Activity 5.2 has the task to join the different bits of the platform assembled in other activities. The central element for the development activity of WP5 is the Common Data Model (CDM), but the platform also aims at incorporating existing software solutions that are of use for taxonomists. The "EDIT Platform Cyber Gate" is introduced as a draft mock-up of the entry page for the Platform website (see presentation).

c. Bibliographic tools (WP5.3: Anton Güntsch for Julius Welby)
Severe problems were encountered with the commercial MetaLib software that had been selected for ViTaL (the Virtual Taxonomic Library, see http://taxonlib.org). However, services are now on-line, for example the ViTaL reference harvester search tool http://taxonlib.org/bibsearch. Digitisation services such as BHL (Biodiversity Heritage Library) have so far not reacted positively to the digitisation request service planned for ViTaL.

Summary of discussion:
ViTaL functionalities can be integrated with other platform components and will be used, for example, in the taxonomic editor software. The question of how to resolve duplicates of references remains open. Uptake of digitisation requests from BHL will certainly be better once the mass digitisation priorities (e.g. taxonomic journals) have been finished.

d. Geographic tools (WP5.4: Patricia Mergen)
GeoPlatform activities include, inter alia, making available maps and web services for map generation according to taxonomist's needs (see presentation). Developments have been integrated with tools developed for the SYNTHESYS project. Further development is somewhat hampered by restricted personal resources, especially at RMCA. The GeoPlatform/MapViewer functionalities were demonstrated in a video (see http://edit.csic.es/edit_geo/EDIT_video.html).

e. Taxonomic core: Taxonomic Editor (WP5.5: Anton Güntsch)
ISTC members reacted positively to a demonstration of the taxonomic editor software, which allows the editing of highly atomised taxonomic data in novel environment that includes word-processor like data entry.

f. Keys and Descriptions (WP5.6: Régine Vignes-Lebbe)
Activity 5.6 will use SDD (Structured Descriptive Data, a TDWG standard) as an exchange format to link existing descriptive tools to the CDM platform. This is done in close cooperation with the Key2Nature project. Existing software applications that can be used to generate identification keys and taxonomic descriptions are not well known among taxonomists. One aim

of WP5.6 is to make these programs better known, e.g. through training courses with in the EDIT School of Taxonomy. For more details please refer to the presentation.

g. Publication: Data Portal (WP5.8: Andreas Kohlbecker)
The Data Portal activity aims at a freely configurable web presentation of CDM content. Currently, the specifications provided by EDIT WP6 exemplar groups (Palms, Flies, Compositae) have been implemented using Data Portal software based on the Drupal content management system. See http://dev.e-taxonomy.eu/dataportal/palmae/ ; http://dev.e-taxonomy.eu/dataportal/diptera/ ; http://dev.e-taxonomy.eu/dataportal/cichorieae/

h. Specimen access: Specimen and observation explorer (WP5.9: Walter Berendsohn)
The EDIT Specimen and Observation Explorer (http://search.biocase.org/edit/) is a web portal for taxonomists that allow access to data provided by the GBIF infrastructure. Queries for names can be expanded using user-selected taxonomic thesauri (taxonomic checklists) and selected name relationships contained therein (e.g. synonyms). Development activity in activity 5.9 has been concluded. However, further taxonomic thesauri need to be included to fully harness the tool's potential. Talks are under way to include botanical databases with global scope, but zoological databases of that kind still need to be identified.

DAY 2: October 2nd 09:00 – 12:00

4. Preparing for the 4th Joint Programme of Activities (Outline & Activity Definition)

Summary:
The over-all plan was accepted by the ISTC as distributed and presented, with the above mentioned recommendations to Activity 5.7, the change of Activity 5.8 from active development to maintenance mode, and the potential reduction of Activity 5.6 as specified below. For further changes and suggestions please see the draft new version of the WP5 JPA-4 annexed to these minutes. No specific recommendations were issued as to where further cost reductions could be applied, if necessary.
The Network Steering Committee (NSC) will discuss the JPA-4 during its next meeting (20th Oct. 2008).
The following statements concerning the JPA-4 tasks within WP5 have been made:

ACTIVITY 5.1

5.1.4 Management, coordination, and liaison with other Work Packages (WPs):
Focus on Integration with other WPs: WP2 expert DB, Metadata collection; WP6 exemplar groups are going to work with WP5 products; WP6 Scratchpad development, Scratchpads should take over the community communication part; WP7 had some staffing problems, but with a new WP5 developer for WP7 starting 1. Oct. 2008 activities will be continued (Tasks: field tools for taxonomist, ATBI websites, integration with geo-platform); WP8: summer school support.

5.1.5 Promote scientific standards and standard tools:
CDM import and export tools, as well as the projected synchronisation interface will be developed based on – or in close collaboration with – TDWG standardisation efforts, where applicable.

ACTIVITY 5.2

5.2.14 Organise integration of software developments at partner institutions with the EDIT Platform
This task was discussed and it was agreed that it should be split into two tasks (please see JPA4 draft)

C5.xx Report on WP7 data management
C. Häuser commented regarding data capture tools in the field: WP 7 needs to start testing available hard-and software and methods for automatic data recording, (e.g. camera with GPS, use of cell phones in the field) immediately, because project time is limited. He calls on ISTC members to send information about existing tools to A. Kroupa or himself. P. Mergen will send some information around about a new communication tool (helmet combined with glasses delivering information) and noted that sponsoring might be possible from this company.

C5.xx Community store documentation for Platform releases
For this documentation one developer has been employed (3 levels: java source code, API / UML diagrams / WIKI on interaction of components and interfaces)

ACTIVITY 5.3

ViTaL: Product is now available; negotiations about possibilities for its long-term maintenance at the NHML will be resumed. In general, for platform sustainability commitments from institutions are necessary, responsibility for different platform components should be taken by different institutions.

ACTIVITY 5.4

C5.71 Organization of a joint Geospatial training course with WP5.4, WP7 and WP8
This task was further specified (please see JPA4 draft). However, WP 8 should take over organisation and should find additional funding to support training courses, also for the subject area of biodiversity informatics. This should be discussed in the next NSC meeting in October 2008.

C5.xx Implement phylogenetic module and C5.xx Implement geographical morphological distribution module.
Given the restricted resources available, these tasks were recommended to be modified or cancelled (please see JPA4 draft)

ACTIVITY 5.5

The timeline for the availability of input tools for structured data into CDM sites was discussed in the context of the Taxonomic Editor. It was clarified that although full input of structured data (taxonomic, bibliographic, distribution, and specimen data) is planned only for Platform version

3 in JPA-4, incremental increases in functionality will be made available throughout the forth and fifth year of EDID as soon as implemented and tested. At present, and with existing funding sources, descriptive data will be edited only via external tools, but visualisation in the CDM-based data portal is within scope.

ACTIVITY 5.6

This activity collaborates with the KeyToNature Project with respect to a joint meeting on identification with KeyToNature and to the further elaboration of an inventory of existing descriptive tools. A further cooperation exists with WP7 concerning taxonomy .applications for descriptive data (e.g. Diptera). It was suggested to muster further exemplar groups for testing the descriptive data tools.

C5.xx Technical integration of taxonomic tools into the platform of cybertaxonomy according to CDM2 , and C5.xx Extension of the existing CDM editor to descriptive data

These tasks were recommended to be modified (see comments under 5.5 above and attached JPA-4 draft)

ACTIVITY 5.7

The recommendations given by the ISTC during the extensive discussion about Shibboleth (see above) are to be integrated into JPA-4.

ACTIVITY 5.8

Current tasks in activity 5.8 will be concluded by the start of JPA-4; the issue of maintenance of the exemplar group portals will be coordinated by activity 5.2. A new task will focus on the production of printed output directly from the CDM, i.e. without last edits that do not enter the database.

ACTIVITY 5.9

This task is completed. Following future tasks depend on integration of non EDIT developments: provision of global thesauri (Tropicos, Kew checklist data). Zoological data are needed, please keep Walter informed. Annotations and feedback systems will be derived from SYNTHESYS developments.

5. Next steps

... were discussed under previous agenda items.

6. Other matters arising

Next ISTC meeting: in September/October 2009 in Berlin depending on the date of the TDWG conference. Suggestions will be circulated as soon as possible. It was suggested to offer guided tours to the collections at MfN and BGBM in conjunction with the next meeting.