TAPIR Meeting
November 2005 (7 to 11)
CSIC Computer Center, Madrid

1.Summary

The TDWG Access Protocol for Information Retrieval (TAPIR) has a first data provider software implemented as a proof of concept (PyWrapper). PyWrapper was presented during the TDWG 2005 meeting, and although it was considered fully functional, some issues with dynamic processing of TAPIR views were identified and deserved attention. A second implementation is planned to be developed on top of the DiGIR2 platform. This meeting brought together developers from both data provider sides (DiGIR and BioCASe) to discuss TAPIR views and implementation issues, and it also tried to define a strategy for a TAPIR Interest Group as well as future plans for the protocol.

List of participants:

- Donald Hobern, GBIF (Global Biodiversity Information Facility), dhobern [at] gbif . org

- Javier de la Torre, MNCN (Museo Nacional de Ciencias Naturales), jatorre [at] mncn . csic . es

- Markus Döring, BGBM (Botanischer Garten und Botanisches Museum Berlin-Dahlem), m.doering [at] bgbm . org

- Renato De Giovanni, CRIA (Centro de Referência em Informação Ambiental), renato [at] cria . org . br

- Robert Gales, KU (University of Kansas Natural History Museum & Biodiversity Research Center), rgales [at] ku . edu

- Steven Perry, KU (University of Kansas Natural History Museum & Biodiversity Reasearh Center), smperry [at] ku . edu

For more information please contact Renato De Giovanni: renato [at] cria . org . br

More details about the meeting can be found at:
http://ww3.bgbm.org/protocolwiki/Madridresults2005

2.Goals

- Discuss the implementation of TAPIR views in PyWrapper and DiGIR2.

- Discuss additional protocol features.

- Revise the protocol.

- Identify software components that either exist or could be developed to improve the usability of the protocol (clients, basic portals, configuration tools, etc.).

- Discuss the relationship between TAPIR and TDWG, in particular if TAPIR should try to become an official subgroup.
- Propose future plans.
- Test, experiment and develop TAPIR wrappers and other TAPIR related software.

3.Out Come

- The issues with dynamic processing of TAPIR views arise from the fact that they are only mapping leaf nodes against class properties, exactly as it happens with DiGIR record structures. This can bring unexpected results if the underlying database doesn't match the schema of a TAPIR view, especially considering cardinality. However, TAPIR dynamic views can still be used in a more controlled conceptual schema world, like the flat DarwinCore schema with extensions, and in this case its flexibility is still desirable. Therefore, the **decision was to keep dynamic processing of TAPIR views, although it is now an optional feature advertised by capabilities response**.
- To better clarify the idea behind a TAPIR view, it was decided to split it into two new things called **"output model"** and **"query template"**. Essentially, an "output model" is the combination of a response structure (schema) mapped against one or more concepts. And a query template combines an output model with filtering conditions.
- It was decided that TAPIR should be able to validate individual documents representing output models and query templates, and that these could also be referenced by URIs inside TAPIR messages.
- It was agreed that **three main levels of provider implementations should be possible**: TAPIRLite providers (completely based on query templates), providers based on fixed output models, and providers capable of dynamically parsing output models (custom output models).
- It has been agreed that initially DiGIR2 will not implement dynamic parsing of output models and query templates.
- The **entire protocol was revised** and many changes were suggested, including a change in how concepts are referenced from the protocol (still using fully qualified identifiers but now decoupled from XML namespace prefixes), a new filter encoding allowing search and inventory operations to be invoked with key-value pairs on HTTP requests, filter extensibility and many others. Changes will be discussed in the mailing list before updating the current schema in the subversion repository.
- A **decision making process**[1] was suggested for TAPIR.
- A **list of TAPIR related software**[2] was prepared.
- **Examples of requests and responses for all TAPIR operations** were included in the wiki.
- A page with **frequently asked questions**[3] was prepared and made available on the wiki. It also clarifies the relationship between TAPIR and TDWG and TAPIR and GBIF.
- It has been agreed that TAPIR should try to **become an official TDWG subgroup**, and Stan Blum was contacted in that respect.
- It was agreed that in principle no additional technical requirements are needed by the subgroup, since a mailing-list is already being provided by TDWG, and both a wiki and a subversion repository are being provided by BGBM.
- A **roadmap**[4] was proposed and made available on the wiki.

---

[1] http://ww3.bgbm.org/protocolwiki/DecisionMakingProcess?action=recall&rev=5
[2] http://ww3.bgbm.org/protocolwiki/TapirSoftware?action=recall&rev=8
[3] http://ww3.bgbm.org/protocolwiki/FrequentlyAskedQuestions?action=recall&rev=6
[4] http://ww3.bgbm.org/protocolwiki/RoadMap?action=recall&rev=4

4.Conclusions

Although there was no time left for the development of TAPIR related software, the whole protocol has been revised with improvements and clarifications in many parts. As next steps, it was suggested to prepare an official specification covering all aspects of TAPIR, and to develop a full featured regression test suite.