Draft Minutes of the Karstlink Round Table meeting

UIS Congress 2022

July 29th 2022

Attendees:

Eric Madeleine - FR
Joël Roy - FR
Fabien Hobléa - FR
Raphael Torquebiau - FR
Frédéric Urien - FR
Claude Alliod - FR
Ferdinando Didonna – IT/CR
Alex Pologea - RO
Wookey – UK
Spela Borko – SLO
Alain Gresse – FR
Stefan Naff – CH
Beat Heeb – CH
Matic Di Batista – SLO
Xiying Me - USA
Mike Lake - AUS

Links shared during the meeting:

License: http://opendatacommons.org/licenses/odbl/1.0/ . License equivalent to CC-BY-SA for non-copyrighted data

UIS Cave and Karst Data Fields List - Draft: http://uisic.uis-speleo.org/exchange/atenlist.html

HDF5 data format developed for storing and archiving large, complex, multi-dimensional scientific datasets: https://www.hdfgroup.org/

Presentation by Frédéric Urien

Collecting, producing, offering KarstLink compliant data with Grottocenter.

Grottocenter for personal use is like any other online website, you can look search and add data.

When you want to share your data each "entity" must have a permanent address (URI). Grottocenter offers a permanent URL to each cave, network, massif, organization, person, document. The USF also provides this type of service.

When we share data we need a license to share them. One has to choose a license for copyrighted data (eg CC license) and also a license for non-copyrighted data (like ODbL). It is important to respect the chosen licenses.

A first step to share data is to provide it with a csv. An improvement is to provide a csv in Karstlink format which facilitates sharing. It is possible to download Grottocenter data with a csv (Non Karstlink). Grottocenter can import data corresponding to documents and caves that are in the csv Karstlink

Next step to submit data is to use APIs, so that data are received automatically. An API to display the data and an API to search the database are the minimum to provide. On Grottocenter all actions can be done with APIs

When we have implemented APIs that provide data in Json format, it is easy to make a small transformation to provide data in Json-ld format which is a format used by Web-semantics applications for SPARQL requests.

Questions:

Alex Pologea: Is the use of SPARQL an option? SPARQL is very technical and should be an option.

- Answer: We should encourage web developers to restrict to one or two queries. For example, JSON-LD is compliant with Karstlink. There are different ways to communicate.

Xiying: She remarks that working flow is straightforward. She would like to know if it is possible to have an automatic that would harvest data once in a while (twice a year)? This way it would be much easier and your database always up to date.

- Answer: APIs allow to communicate with DB that create data, and see if the DB are up-to-date. CSV format does not allow that. But it is still important to have a CSV file for those who use Excel file.
-

Presentation by Ferdinando Didonna

Data exchange and protection of karst - legal issues

Data protection state of the platform

- Data subject to copyright
- Unless otherwise stated, all rights to the content of the Grottocenter site are granted under license CC-BY-SA. Unless otherwise stated in a particular document or in an appendix thereto, the information it contains under copyright is subject to the license CC-BY-SA.
- Data not covered by copyright
- The Grottocenter database structure is available under the Open Database license: http://opendatacommons.org/licenses/odbl/1.0/ . All rights to individual database content, which are not under copyright, are granted under the Database Content License: http://opendatacommons.org/licenses/dbcl/1.0/

You are free to:
Share — copy and redistribute the material in any medium or format
Adapt — remix, transform, and build upon the material
for any purpose, even commercially.
The licensor cannot revoke these freedoms as long as you follow the license terms.

The General Data Protection Regulation (GDPR) for personal data, the EU regulation

7 key principles of GDPR:

1. Lawfulness, fairness and transparency
2. Purpose limitation
3. Data minimization
4. Accuracy
5. Storage limitation
6. Integrity and confidentiality (security)
7. Accountability

Right of access to environmental information

What is the right of access to environmental information?
It is the right of the public guaranteed by Legislative Decree no. 195 of 19 August 2005 (ITALY) in implementation of EU Community Directive 2003/4/EEC, to access information on the environment held by Public Authorities. Public Authorities must be understood not only as state, regional and local Public Administrations, but also as autonomous and special companies, public bodies and concessionaires of public services, as well as any natural or legal person who performs public functions related to environmental issues or exercises administrative responsibilities under the control of a public body.

Is it possible to access any type of information?
No. Access may be deferred, excluded or limited for the cases referred to in art. 5 of Legislative Decree no. 195 of 19 August 2005.

INSPIRE directive and Geodata in Europe

1. More efficient management – data should be collected only once and managed where this can be done more efficiently;
2. Interoperability – it must be possible to combine data from different sources and share it between multiple users and applications;
3. Sharing – it must be possible to share information collected by different levels of government;
4. Abundance and usability – the geographical information necessary for good governance must exist and be truly accessible under conditions that do not limit its possible use;
5. Findability and access – it must be easy to identify what geographical information is available, assess its usefulness for your own purposes and the conditions under which it can be obtained and used.

Cave and Karst legal status
Basic assessment of nation-al, regional, and local regulations on caves, karst, and caving worldwide to strengthen common knowledge about karst and cave protection.
Results as of 31 May 2022
We received 37 responses from the 29 countries
The questionnaire to participate in the survey online is https://forms.gle/NZBKykNf2LDhFm2A7

Questions:
Mike L: In Australia, there are very few private caves. Data on the location of cave entrances is restricted. Clubs hold the exact location, the national database does not have the exact location. It is legally allowed to share location, but in protected areas and parks the authorities wouldn't be happy if it is shared.

Stephan: remarks that assets contain personal data without you knowing it. For example, initials on a map. It is subject to GDPR.

Frédéric: In French law, the name of the authors is mandatory on any documents. So we have to ask this information for each document we receive, which poses problems because this information is often difficult to collect. It also takes time depending to trace the authors. So many documents are not shared due to lack of this information.

Matic : In Slovania we have 3 levels of protection of the main DB gate.

1. Close
2. Close but ongoing monitoring
3. Open. 14 000 caves, about 99%.

Xiying Mi - University of South Florida Libraries

A Global Solution for Karst Metadata Sharing

Why share this information?

- Create wider access to the collection;
- Create richer user experience;
- Create collaboration opportunities with researchers in the field;
- Create datasets in linked data format

How?
Data mapping: Dublin Core KarstLink
Vocabulary mapping: LC terms + Getty terms. Karst ontology (Bibliography Commission of the International Union of Speleology)
Data format: Names, dates, encoding
Platforms:
https://digitalcommons.usf.edu/kip_talks/1/
https://grottocenter.org/ui/documents/135600

Management

- Data collecting: one staff member
- Data standardization: student workers
- KarstLink: one librarian

Lessons learned

- Metadata fields mapping
- Ontology exchange between library and KarstLink

Questions: It would be nice to have a read me document to share with us all so that we can see what has to be done to have the DB Karstlink compatible.

Xiying: Why do you use subjects for documents that are not standardized subjects like those offered by the Library of Congress?

Frédéric: the subjects were defined by the UIS bibliography commission. The subjects related to the Karst are specific and do not appear in the standard subjects but the reflection must be prolonged

Eric Madelaine

Eric's contribution still coming

Joël Roy (Karsteau)

A communication experience between 3 French databases

Karsteau's "develop" team has launched an experiment in database communication to federates 3 bases to make a first test, Karsteau, Grotto Center, Vmap. Manage the caving data with already existing tools that have been proven and used by a large community of experienced web developers who contribute to the development and improvement of the system. The tool must be reliable.

Work with PostgreSQL for the following reasons:

- This type of basic management tool is widely used by government departments and local authorities as the basis for their GIS.
- Main reference in France: IGN uses PostGIS to store high resolution French topographic maps. The BDUni has over 100 million features.

PostgreSQL

It is an open-source database management system. This database management system can handle many types of data (vectors, rasters). PostGIS is used to manage spatial data. (Spatial databases allow the storage and manipulation of spatial objects).

The ideal tool for manage caving data. PostGIS allows to store natively:

- •points (x, y, z): the entries ;
- •polygons: sectors and sub-sectors;
- •lines: topographic transfers.

Many SQL functions allow you to calculate: areas, lengths, intersections between two objects, buffers, etc.

A possibility is the interconnection of Postgres-PostGIS databases residing on different servers. Vmap has recently been interconnected with GrottoCenter and Karsteau as part of an open access cave sharing test.

The interconnection is via SQL views, parameterized by each database, which allows data to be distributed according to the requirements of each database.

The advantage using Postgres is that the implementation is simple and requires very little development and quickly.

The same issues remain as for Karstlink and Semantic Web

- Ontotology
- URI
- legal issues of data protection