D09 – DELIVERABLE 7.1.4

Project Acronym: **OpenUp!**

Grant Agreement No: 270890

Project Title: **Opening up the Natural History Heritage for Europeana**

Guidelines for users and content providers v. 1

D09 – Deliverable 7.1.4

Revision: Final

Authors:

Franck Theeten	**MRAC**
Pere Roca Ristol	**MNHN**
Boris Jacob	**MRAC**
Janno Jõgeva	**UT-NHM**

0 REVISION AND DISTRIBUTION HISTORY AND STATEMENT OF ORIGINALITY

Revision History

Revision	Date	Author	Organisation	Description
1	2011-07-12	B. Jacob	MRAC	Publishing first post in the forums of the OpenUp! Helpdesk
2	2012-02-01	F. Theeten, B. Jacob	MRAC	Compiling information from forum posts and other communication to one document
2	2012-02-01	P. Mergen	MRAC	1st review by WP leader
3	2012-02-23	F. Theeten, B. Jacob, P. Roca Ristol, J. Jõgeva	MRAC MNHN UT-NHM	Update
3	2012-02-23	P. Mergen	MRAC	2nd review by WP leader: ok to send this document to partners for comments and further contributions, revisions and more input needed before it can be accepted as a deliverable
4	2012-02-27	B. Jacob	MRAC	Update, incorporating comments by WP lead
4	2012-02-27	P. Mergen	MRAC	3rd review by WP leader: minor typos, then ok
4a	2012-02-27	B. Jacob	MRAC	Minor editorial corrections
4b	2012-02-27	Coordination Team	BGBM	Minor editing

Statement of Originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Distribution

Recipient	Date	Version	Accepted YES/NO
Work Package Leader (P. Mergen, MRAC)	2012-02-27	4	YES
Project Coordinator (W. Berendsohn, BGBM)	2012-02-28	4b	

TABLE OF CONTENTS

1   INTRODUCTION

This is version 1 of the "OpenUp! Guidelines" and it is based on questions raised and answers provided in the OpenUp! Helpdesk Forums[1] over the last seven months. A list of questions in the Appendix refers to the chapters with the answers. Answers to these questions have been provided by the key staff of the Helpdesk Network[2] in close collaboration with other OpenUp! partners and with colleagues from other projects and networks. This work will be continued in the second year of the project and these guidelines will be updated accordingly.

For already existing User guides and documentation, in order not to duplicate efforts and guarantee up to datedness of the content, the reader is kindly asked to consult the guides and manuals of those projects and initiatives. Those are mentioned in chapter 3.4, which is based on OpenUp! *Component 7.2.1 Identification of existing documentation.*

Two developments on side of Europeana are important to mention here. First, the current data model ESE (Europeana Semantic Elements) will in the future be replaced by the much richer Europeana Data Model. Until then, all data will be mapped to ESE. Secondly, as of beginning of this year, the old agreements between the content providers and aggregators on the one hand and Europeana on the other hand will be replaced by a new Europeana Data Exchange Agreement (DEA), which allows Europeana to publish the provided metadata under the CC0 license. Documentation on this can be found in the Appendix in chapter 3.4.2. Also in the Appendix, there are two lists of mappings between ABCD/ESE provided by the OpenUp! Technology Management Group (TMG), one for the full data set, one for the restricted data set. The restricted data set being a compromise towards content providers that do not wish to provide the full data set under the new Europeana Licensing Framework.

The authors would like to thank their colleagues within the OpenUp! project who helped answering questions and thus also contributed to these guidelines, especially Gerda Koch and Jiri Frank. Furthermore we would like to thank our colleagues from our partner projects BioCASE, GBIF and Europeana, who helped answering questions and of which documentation and technology the OpenUp! project builds upon.

Efforts have been made to give accurate and up to date information in these guidelines. However, the authors cannot take responsibility in case some guidelines may not work or achieve the expected result. In case of comments or questions, please contact the OpenUp! Helpdesk at
http://openup.helpdesk.africamuseum.be

---

[1] OpenUp! Helpdesk Forums: http://openup.helpdesk.africamuseum.be/forum

[2] See OpenUp! Component 7.1.1 for reference:
http://open-up.eu/sites/open-up.eu/files/u2/C711_Key_staff_helpdesk_Final%20Version_Updated.pdf

2 GUIDELINES[3]

2.1 Basic technical information

2.1.1 What is an IP address

Question: What is an IP address?[4]

An Internet Protocol address (IP address) is a unique number, identifying the devices belonging to the same computer network. It is used both in LAN (Local Areas Networks), WAN (Wide Area Networks) and on the Internet. In the current version of the IP protocol (version 4), addresses have a fixed length of four octets (32 bits), but a new version of the protocol (version 6), using 128 bits is expected to be introduced at wider scale anytime soon to overcome the depletion of available addresses on the Internet.

Typically an Internet address has the following form: 192.168.0.1 (Decimal representation of 4 octets). This address is divided in 2 parts having variable size: The first part of the address identifies the number of a secondary network into the main network, while the second one identifies the number of the machine inside of this network.

There are 3 possible subnets determining how much octets are used to identify the network (1 to 3) and how much are used to identify the device in the network (the remaining one). An IP address can be either static (permanent) or dynamic (valid for a limited time and reinitialized, in order to reuse deprecated addresses).

Installation of a BioCASe provider is made easier if the data provider has an Internet static IP and/or a permanent domain at his disposal.

2.1.2 What does DNS mean

Question: DNS in practice: what is it and how does it work?[5]

The Domain Name System (DNS) is an Internet service linking an Internet IP address with one or several alphabetical and human readable domain names. Domain names are hierarchically organized from right to left. In the domain name *openup.helpdesk.africamuseum.be* for example *openup* is a subdomain of *helpdesk*, which is a subdomain of *africamuseum*, while *africamuseum* is a subdomain of *be*. Each of these subdomains has its own IP.

Theoretically, addresses are resolved sequentially in order of their organisation. In practice, many DNS cache the IP of several root domains and subdomains placed at different hierarchical levels to alleviate the traffic on the Internet.

---

[3] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/forum/620

[4] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/164

[5] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/165

2.1.3 Definition of SQL databases

Question: What is an SQL database? [6]

SQL (Structured Query Language) is a programming language for relational databases standardized by ISO as an international norm. It is used to:

- define the structure of databases (SQL-DDL: Data Definition Language).
- query data in the database with complex search criteria linked together in Boolean conditions (SQL-DQL: Data Query Language).
- modify the data: insert new values, update old ones, delete deprecated records, etc. (SQL-DML: Data Manipulation Language).
- handle the security and integrity of the databases: allocate hierarchical rights and privileges to users, rollback aborted transactions, etc. (SQL-DCL: Data Control Language).

A lot of database software is compatible with SQL (especially those intended for Internet access to the data). But they often vary from the ISO norm and are extended with additional programming functionalities specific to each implementation, thus allowing the use of a language similar to C in database statements.

MySQL and PostgreSQL[5] are two well-known and widely used open-source database management systems compatible with SQL. The main drawback of SQL is that it is a text-based programming language without graphical interface. MySQL and PostgreSQL feature nonetheless modules and extensions helping the user to generate SQL statements from a graphical interface (e.g. 'phpMyAdmin' for MySQL and 'PGAdmin3' for PostgreSQL). PostgreSQL features a plug-in called 'PostGIS' that extends the original SQL statement with specific instructions for geographical data, allowing calculations of distances and areas, creation of base data for maps, reprojections, etc.

2.1.4 Explanation and glossary in regard to relational databases

Question: What is a relational database and how does it work? [7]

A relational database system is software that stores data in a multi-tabular structure. All information referring to the same real world or logical object is supposed to be stored in a specific table in order to ease performance, to avoid redundancy of information, and to ease the check of the integrity of the system. A specific jargon is associated to relational databases. The definitions below are given for information purposes in order to encourage using a relational database for the first time.

- **Table**: a bi-dimensional container of information with rows and named columns
- **Record**: an object that is described in a single line of a table
- **Field**: a named column in a table that stores information in a specific 'type' or format (text, integer, float, bytes, etc.).
- **Normalization**: The process of defining a multi-tabular structure by logically grouping the object described in the system in separated tables.
- **Transaction**: a set of instructions modifying a batch of rows in a database (deleting, adding, or modifying data). Transactional databases are able to revert the data back to their initial state if a transaction fails and inform the user of this failure.

---

6 See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/171

7 See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/169

- **Primary key**: Unique and non-repeatable identifier specific to each object in one table (usually a unique index number).
- **Foreign key**: Repetition of the primary key in another table to represent the link existing between the two tables.
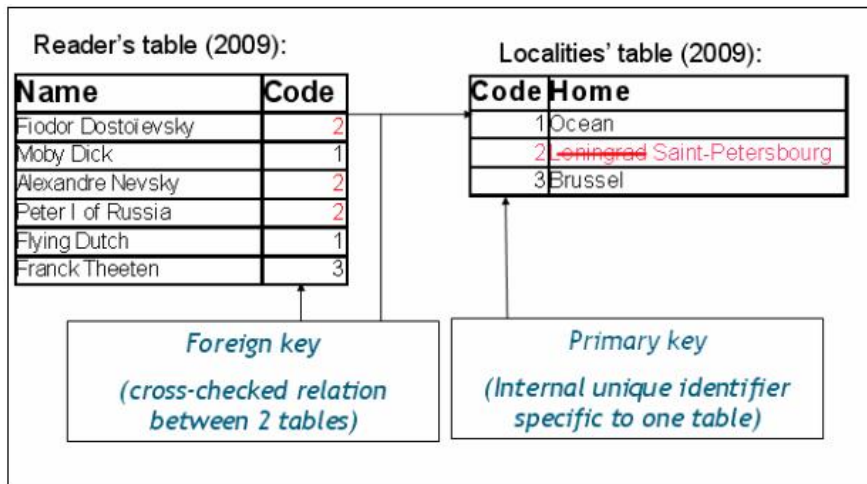
Fig. 1 Illustration of the relation between primary keys and foreign keys

- **Referential integrity**: a process that ensures that data stored in different tables remains consistent, e.g., the deletion of data in one table still referred by another table is made impossible unless the user deletes those references first.
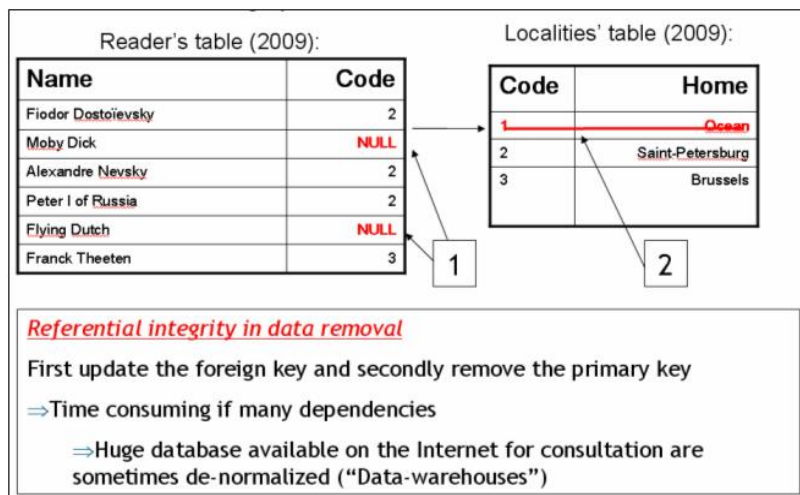
Fig. 2 Illustration of the concept of relational integrity

- **Query**: a set of instructions to retrieve the content of one of several tables with optional search criteria. The return of rows is technically called 'resultset' or 'recordset'. If you need to query information located in several tables, the reference to the relation between a primary key and a foreign key in a query is called a Join. Queries with Inner Join will extract the data that contains reference in all the requested tables; queries with Outer Join will extract all the data (including the values that belong only to one table). A query corresponds to the 'DCL' part of the SQL language (see pages on SQL).

- **View**: a view is a query made permanent in a database. This is especially useful for queries that are very complex and contain several joins (i.e.: several interlinked tables). The SQL syntax to query tables and views is the same (except that the user cannot modify, add, or delete data in a view while it is possible in a table). In most database systems, it is also relatively easy to create a new table from the results of a view. So a view is a convenient way to increase the security of your database when connecting it to an Internet service and to simplify the handling of complex data.

2.2 Data exchange schemata and protocols

2.2.1 OpenUp! A first technical approach

Question: What's the big picture of OpenUp! and its different actors?[8]/[9]

One of the major technical issues OpenUp! has to solve is the provision of links to multimedia files coming from multiple scientific institutions across Europe. Textual, non-multimedia information has been harvested for a long time by different projects like GBIF or BioCASE, using a variety of software and protocols (BioCASe, GBIF Darwin Core, DIGIR, TAPIR).

These projects use (among others) the ABCD schema (Access to Biological Collection Data), an XML schema for curatorial metadata and for data used to map data coming from SQL systems to the network for future harvesting. The ABCD schema with its about 1200 concepts[10] is quite large.

ABCD features an XML element allowing to link the description of a specimen to a multimedia document that cannot be directly harvested by Europeana as its structure is different from the Europeana data schema. The reference object of ABCD is a specimen in a collection or an observation while in Europeana it is the multimedia document itself.

One of the aims of OpenUp! is to make data in ABCD format compliant with the Europeana data schema in order to provide public access to the complete data, metadata, and associated content.

OpenUp! will use the established technical infrastructures of GBIF including the Biological Collection Access Service (BioCASe) to create a single access point to distributed multimedia content in the natural history domain for Europeana.

This means that OpenUp! uses the BioCASe provider software and the ABCD metadata schema to feed Europeana. It is important to note that BioCASE can also handle other schemata like DarwinCore but in OpenUp! ABCD was chosen for its greater complexity and usability. The aim of OpenUp! is then to create a network of BioCASE providers that can be accessed and harvested by Europeana through the Harvesting and Indexing Toolkit (HIT) developed by GBIF, AIT, and other partner institutions of OpenUp!.

In order to keep this metadata enrichment compliant with Europeana standards, a mapping from ABCD to ESE (Europeana Semantic Elements) and later on to the Europeana Data Model (EDM) that will replace ESE, is required.

---

[8] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/forum/623

[9] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/204

[10] XML document listing all concepts: http://www.bgbm.org/TDWG/CODATA/Schema/ABCD_2.06/ABCD_2.06.XSD

2.2.2 BioCASe and the XML schemas for biodiversity metadata

Question: Which XML schemas do I use with the BioCASe provider? What is the ABCD schema?[11]

When submitting data through the BioCASe provider, you need to define a mapping between your source database and the ABCD XML schema to exchange data with the network[12].

ABCD is the schema being used in the context of OpenUp!. Its current version is 2.06. ABCD is a hierarchical schema (with a tree structure) containing about 1200 fields (or 'concepts'). It features an extension mechanism. Most of these fields are optional, only 5 fields related to the identification of a specimen or object in a collection and to its scientific name are mandatory.

ABCD has its structure, syntax and semantic being discussed, controlled and approved by TDWG Biodiversity Information Standards, also known as Taxonomic Databases Working Group. TDWG is a not for profit scientific and educational association that is affiliated with the International Union of Biological Sciences.

2.3 Installation and configuration

2.3.1 Checklist installation and configuration of the BioCASe web service

Checklist: Installation of BioCASe		
1	Install an Apache HTTP Server (v. 2.2).	
2	Install Python.	
3	Download and unpack BioCASe.	http://wiki.bgbm.org/bps/index.php/Installation
4	Run the installation script of BioCASe from the console ("python.setup.py").	
5	Adapt the config file of Apache (httpd.conf or /site/<sitename>) with the output of the configuration script (setting of the alias and web directory).	BioCASe should already be visible from the web at this point, but additional libraries need to be installed.
6	Install the additional Python Packages and libraries (like DB drivers).	Local BioCASe > Utilities > Library test
7	Change the main password.	Config tool > System administration

---

11 See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/198

12 See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/197

Checklist: Registration and mapping of a new dataset		
8	Create a new DSA.	Config tool > System administration > Datasources > Create DSA
9	Refine the database connexion.	Config tool >Datasource administration > Database Connection > Edit Connection parameters
10	Register aliases, primary keys and foreign keys for the tables and views to be mapped.	Config tool >Datasource administration > Database Structure > Edit DB Structure Click "Save" +"Refresh" after the definition of a new table in order to access the Primary Key/Foreign Key parts.
11	Choose the mapping schema.	Config tool > Datasource administration > Schemas > List "Map your DB against a new schema.
12	Enter the mapping form.	Config tool > Datasource administration > Schemas > Link with schema name

2.3.2 BioCASe as a web service wrapping databases

Question: Does BioCASe store the data? Is it a database system? [13]

BioCASe is a web application installed on top of a database and exposes the data of the source database to the web. It is intended for databases containing curatorial, taxonomic, and ecological information. BioCASe does not store any data itself (unlike Access, Excel, or BRAHMS) but it is dependent on the external source database system that contains the data (SQL or ODBC).

All queries and results containing data are exchanged on the Internet and both are standardized into a common XML format. By exchanging queries and data in the XML format, applications using BioCASe are independent from the technical implementation and structure of the original database. BioCASe is thus an XML over HTTP application that enables the client to ask the data to be 'structure-agnostic' towards the structure of the server-side database containing the source data. BioCASe:

- enables users to define a correspondence between the structure of their database and the structure of the XML schema,
- provides an XML over HTTP interface to interpret queries (send in the BioCASe format as HTTP parameters or as XML) and return the results as XML document,
- features a graphical search interface (called 'QueryTool') that converts the XML into a human-readable representation.

---

[13] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/197

Several BioCASe providers can provide data to a network whose portal acts as a central gateway to all source databases. BioCASe is already being extensively used as a data provider for the GBIF network[14] and the BioCASe network[15]. These networks are semi-decentralized: part of the data are harvested and "cached" into a central database belonging to the network, other information is queried via live requests sent to the provider. Special interest networks using BioCASe include the Australian herbaria network (using the HISPID extension of ABCD) and the DNA Bank Network (using ABCD-DNA).

2.3.3 BioCASe installation (prerequisites and procedure)

Question: How do I install BioCASe?[16]

In order to install the BioCASe provider software (also known as "PyWrapper") you need to have:

- Python installed on your computer,
- an HTTP server (preferably HTTP Apache) with its Python module,
- an SQL database system (MySQL, PostgreSQL, Oracle, MS SQL server, ODBC, etc...),
- the Python libraries for your SQL database (the PyWrapper can guide you for their installation).
- Some other Python libraries are needed (e.g., to use a an internal search engine called the QueryTool) but can be installed after the installation of the main BioCASe package

BioCASe can be installed on Linux, Windows, and Mac. The installation process is divided into three major steps:

- run the script "setup.py" located at the root of the web folders,
- update the configuration of your Apache HTTP server (the installation script generates a sample configuration to be copied and pasted),
- use a module of the BioCASe provider software that checks that all the needed libraries are installed.

For a more complete description of the installation process, please see the documentation at the PyWrapper Wiki[17].

---

[14] GBIF Data Portal: http://data.gbif.org

[15] BioCASE Network: http://www.biocase.org/

[16] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/IntroInstallation

[17] PyWrapper Wiki: http://wiki.bgbm.org/bps/index.php/Installation

2.3.4 Installing the BioCASe provider/PyWrapper on your computer

Question: Can I install the BioCASe provider/PyWrapper on my own computer? [18]

It is technically possible to install BioCASe on your own computer but it is suggested to do so only for testing purposes. As the BioCASe provider is a web service that must be connected to a network, it should be installed on a server with a permanent Internet connection, static IP, and domain name.

If you do not have direct access to the servers of your institution, it is suggested to:

1. install a running BioCASe on the server, and leave it empty at this stage,
2. ensure that you also have a database server ready on your computer,
3. install
     i. BioCASe pywrapper,
     ii. HTTP Apache,
     iii. a copy of your SQL database on your own computer. Eventually you can directly connect the BioCASe provider installed on your local computer to the database of the institution if your network administrator allows you to do so.
4. Define a first ABCD mapping on your local computer.
5. Once the mapping is functional on your own computer (it is available through http://localhost/<myBioCASeWebfolder>):
     i. You can copy your database to the server of your institution or ask your colleagues in charge of the server to do so. If possible do not rename your database.
     ii. Copy the config folder of your own BioCASe dataset to the one located on your institution server:<BiocaseFolder>/config/datasources/<Name_of_the_datasource>)
6. You will probably have to correct the connection settings (user names and password) from your local database to the server database (see Fig. 3) as last step.
7. If the two databases have the same name and structure and they run on the same SQL server, BioCASe should be functional at this stage.

---

[18] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/199
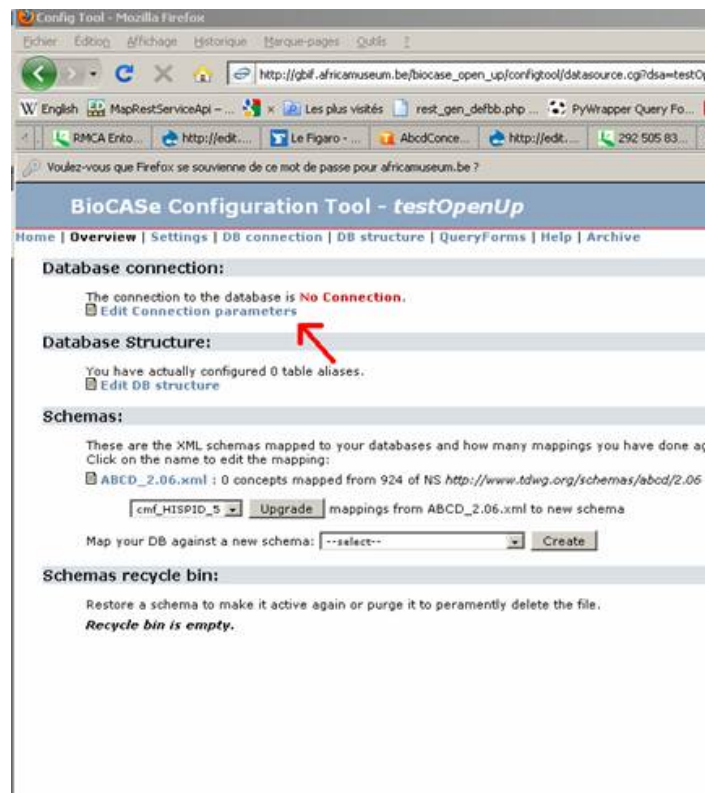
Fig. 3 Editing PyWrapper connection parameters

2.3.5 BioCASe on a virtual machine

Question: Can I install the BioCASe provider/PyWrapper on my own computer? [19]

As BioCASe requires the installation of several low-level libraries (Python database drivers, XML libraries, GraphViz to visualize your database structure), it could be interesting to install BioCASe as a guest of a virtual machine. This virtual machine can be moved between your local computer and the server of the institution as a single package.

By doing so the installation of the libraries, the database, the configuration of Apache and definition of mapping occur only once. However, you have to check the network configuration (definition of IP and of machine name) when you move your virtual machine from one server to another. A server for virtual machine requests a substantial amount of resources but can be installed on modern PCs and laptops. Virtual machines greatly ease backup, migration of data, and copying between several servers.

Several servers for virtual machines are available, amongst others:

- VirtualBox (ex Sun, now Oracle, free, features a graphical interface): http://www.virtualbox.org/

- OpenVZ (free and Open Source but no native Graphical interface, although plug in exists): http://wiki.openvz.org/Main_Page

---

[19] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/199

- VMWare (commercial with a graphical interface to administer. Some major components are freely available): http://www.vmware.com/

Most of these virtual servers recognize the OVF (Open Virtualization format) to allow cross-platform replication of a virtual machine.

2.3.6 BioCASe and its back-end database

Question: How do I connect an SQL database to BioCASe?[20]

BioCASe is intended to work only with SQL databases and it connects these databases by means of a low-level software called 'driver' (that must be installed on the server as a prerequisite).
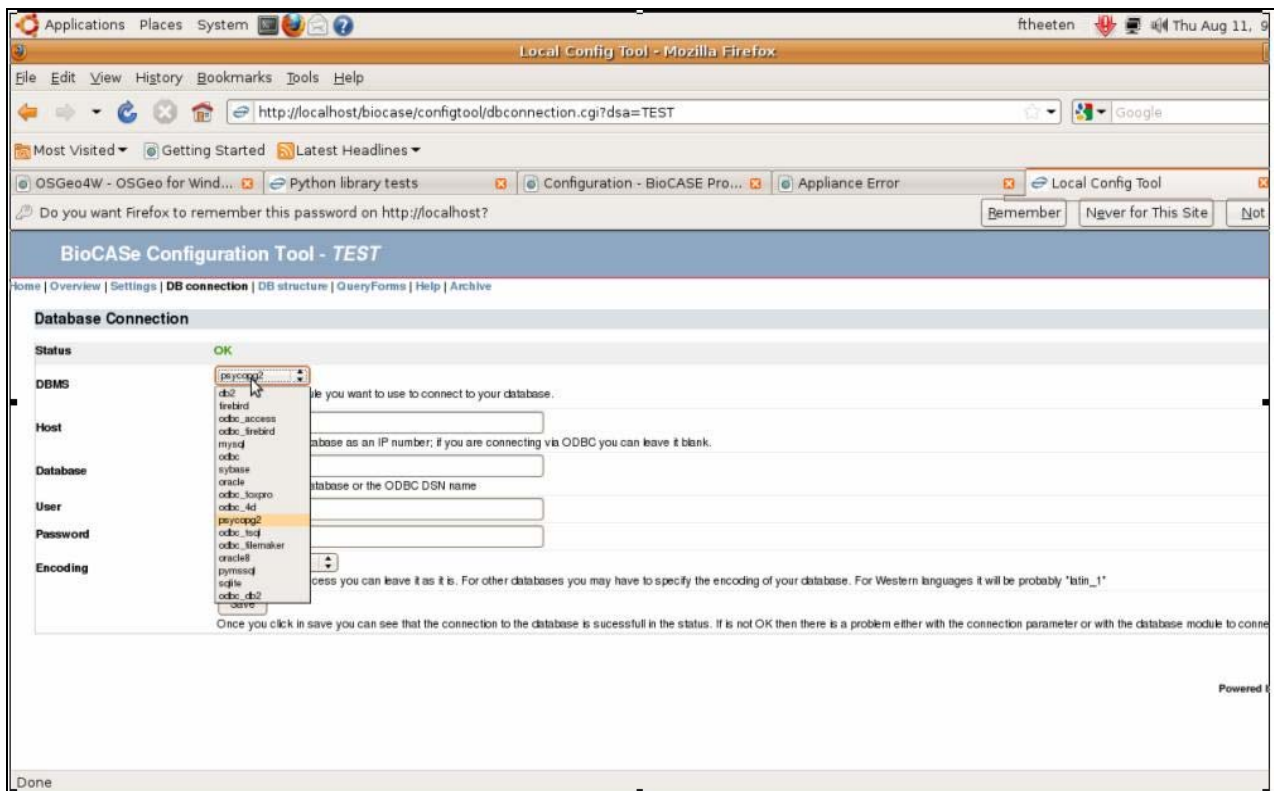
Fig. 4 Configuration page of the BioCASe (version 2.6.0) provider for connecting to an SQL database by generating the appropriate connection string. The drop-down list presents the recognized drivers for SQL databases.

The database management system MS Access is compatible with SQL but unfortunately this is not as easy with MS Excel. It is possible, however, to connect an ODBC driver to Excel and thus make it compatible with SQL (performance in the network, however, will still be very poor). Several desktop database applications intended for specific biological work (curatorial management, definition of taxonomic keys) exist that are not compatible with SQL in their native state. These often have a very good interface that allows an easy visualisation of the scientific content. . To use these (or other non-SQL databases) and publish your data with

---

[20] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/172

BioCASe, you will have to build a replica of your database (or of the part you want to publish) in a SQL system, and transfer the data between the two systems, e.g. in tab-delimited or CSV-format[21].

The difference between the encoding used in the source database, the intermediate CSV document, and the destination collected to BioCASe is a classic source of technical problems in this process (data can be published with unreadable characters replacing diacritic marks). Please ensure that the CSV documents and the database share the same encoding format (UTF-8, LATIN-1, etc.)
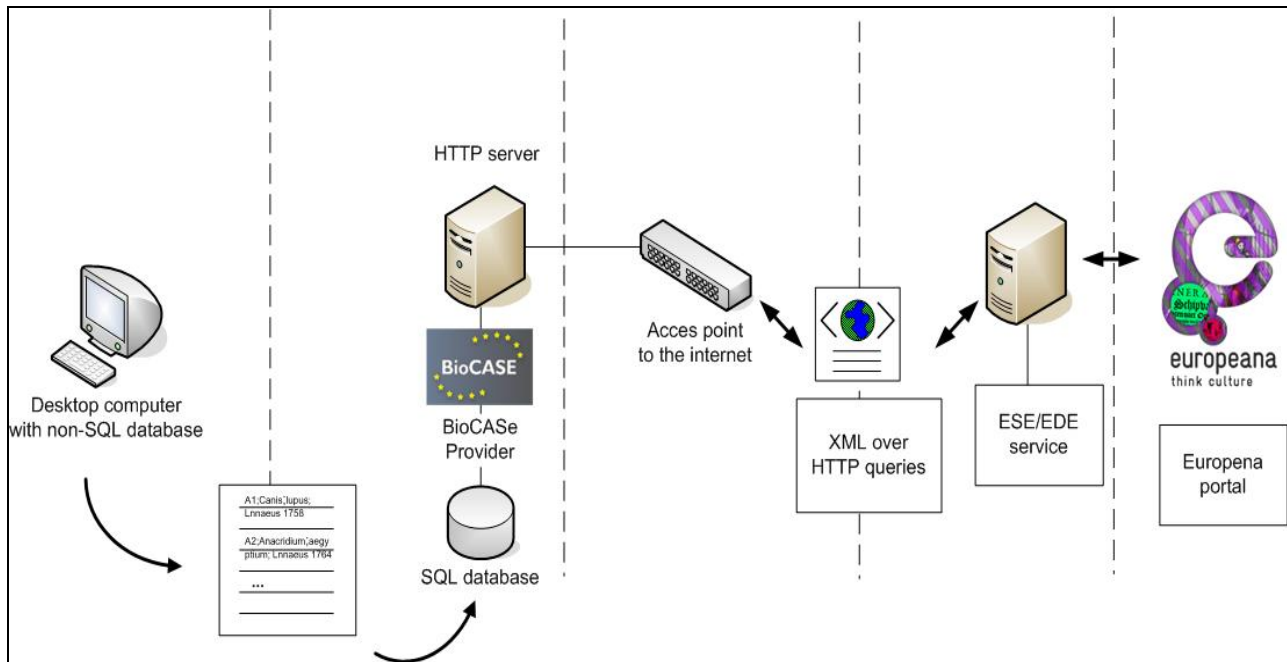

Fig. 5 Full workflow of the submission of data to OpenUp! from a source desktop database to the portal. The BioCASe provider intervenes in the 2nd step.

2.3.7 Preparation of the database by creating dedicated SQL views

Question: Should I use database tables or views for publishing data with BioCASe? [22]

When publishing data with BioCASe, we strongly recommend that users creates views based on the data they want to submit to BioCASe as a preparatory step and declare this view in the BioCASe provider. Ideally the user should create 3 views, one for the scientific information provided to the network, one for the metadata of the dataset, and a third one with the contact coordinates of the scientific curators and technicians working on the data.

If the database system you plan to publish via BioCASe is not the same as your reference database for your daily work but a copy of it, it will be relevant to create only three tables in the copy that have the same structure of the views defined in the reference database, in order to simplify the import of data from the

---

[21] CSV (comma-separated values) and tab-delimited formats are two ways to format structured content in order to store tabular information (field-delimited sheet of data featuring rows and columns) into a simple text document that can be opened by a simple text editor. Commas, semicolon, tabs or other characters are used to represent columns in the document). Most of database systems and spread sheet software can export and import data in this format.

[22] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/170

original database to the copy (via the exchange of data in CSV format). Exporting data from a view in CSV format is a simple task in most of relational database systems.

2.3.8 IP, DNS and the connection of the BioCASe provider to networks

Question: How do I configure my server to let users access BioCASE? [23]

Once a user has configured a dataset with the BioCASe provider, he receives what is called an "access point". This access point is a URL identifying both the server containing the provider software, and the name of the resource. There is one different access point for each different collection registered in the provider. This URL is the gateway to the collection and must be communicated to portals and networks so that they can access your data together with those of the other providers by means of a common search interface (like GBIF).

The access point can be accessed by clicking on the link containing the dataset name from the first page of BioCASe. In the example (see Fig. 6) the access point is:
http://gbif.africamuseum.be/biocase_rmca/dsa_info.cgi?dsa=demo_openup

This access point is supposed to be permanent (or 'as permanent as possible'). For this reason we suggest that users of the BioCASe provider willing to contribute their data to a federated network have at least an Internet domain name and a static IP at their disposal, and use the domain rather than the IP (that may change more often in time) in the access point URL.

Please remember that two services are intervening when resolving the URL of a BioCASe access point (like for most of the websites):

1. The DNS of your institute that must first resolve the domain of the provider. This is the left part of the URL before the first single '/', in our example: http://gbif.africamuseum.be
2. The HTTP server configured inside of this server that resolves the remaining part of the URL (/biocase_rmca/dsa_info.cgi?dsa=demo_openup). This server will most likely be Apache - an open source HTTP server for webpages. The HTTP server redirects the right part of the URL to the appropriate website made available within the domain.

The BioCASe provider contains an installation script that also helps to configure the Apache HTTP server in an appropriate way. This script is called 'setup.py' and must be run in DOS console (Microsoft system) or shell in Linux-based systems. Beforehand, the DNS needs to be configured (normally this is the case if the server you use already provides webpages to the public).

NB: it is possible to link several domains to a single IP address. That situation requires that the system administrator pays special attention to the good synchronization between the DNS and the Apache HTTP server. The domain names used in the configuration of both services should be the same.

---

[23] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/168
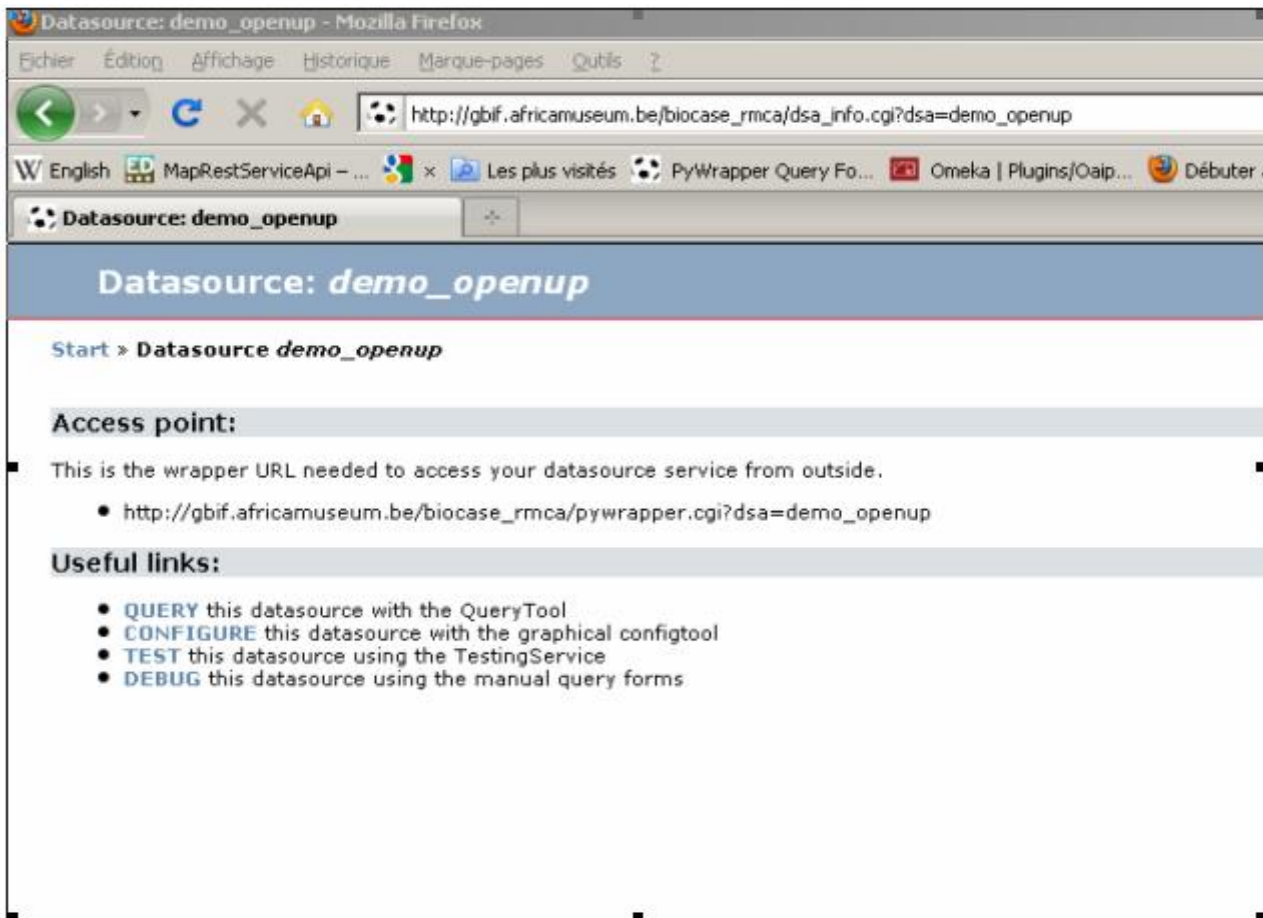
Fig. 6 The BioCASe page providing the Access point to a dataset (i.e. the URL to be communicated to the network operating the portal where data are provided)

2.3.9 Editing and copy/pasting the configuration file

Question: Can I replicate datasets between several implementation of the BioCASe provider? [24]

In case you plan to map several databases, it is suggested to create smaller datasets, where you can refine the mapping, instead of a big mapping file containing all the tables. If your database tables and views all share the same structure and field names, you can relatively easy generate smaller datasets from a big one.

1. In your file system go to **<BiocaseFolder>/config/datasources**
2. Copy the main subfolder with the name of your original dataset into another folder having the name of the new dataset you want to create (see Fig. 7).

---

[24] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/200
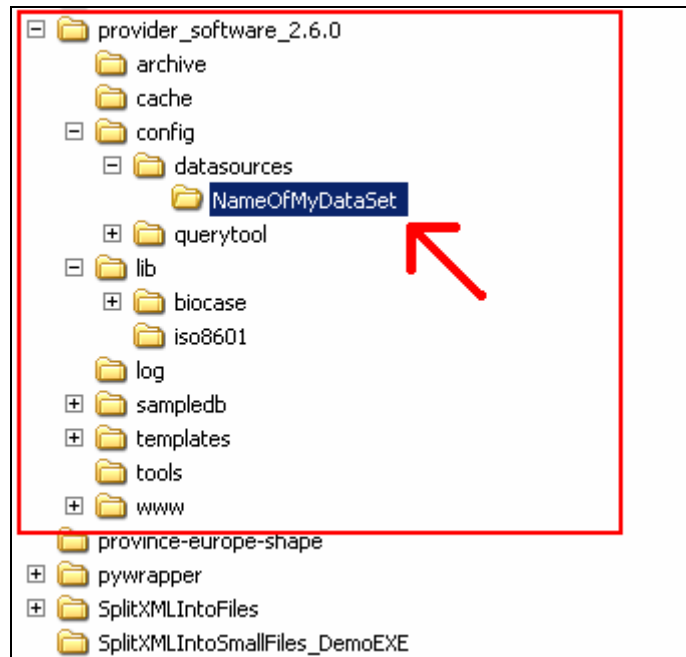
Fig. 7 Folder where a datasouce is located

3. The folder "NameOdMyDataset" matches the name of the dataset in your Access point URL (e.g.: http//localhost/biocase/pywrapper.cgi?dsa=NameOfMyDataset) Copy/paste the subfolder will create a valid BioCASe dataset with the same structure and the name of the subfolder.

4. in your SQL database, crate a view which has the same structure (same number of fields and field names) as the table (or view) you connected previously, with the appropriate filter in its "WHERE" clause

5. Open your new dataset in BioCASe, go to "configure" and then "database structure"-> "edit database structure"

6. in the column 'Table', replace the name of your previous table with the one of your new view and save your configuration (see Fig. 8).

This new dataset should be accessible and visible in the list of DSA of the provider.

The same procedure can be applied to **move or replicate** the **configuration** of a dataset from one instance of the BioCASe pyWrapper to another one (don't forget to move the data and/or to update the setting of the connection to the database in parallel).
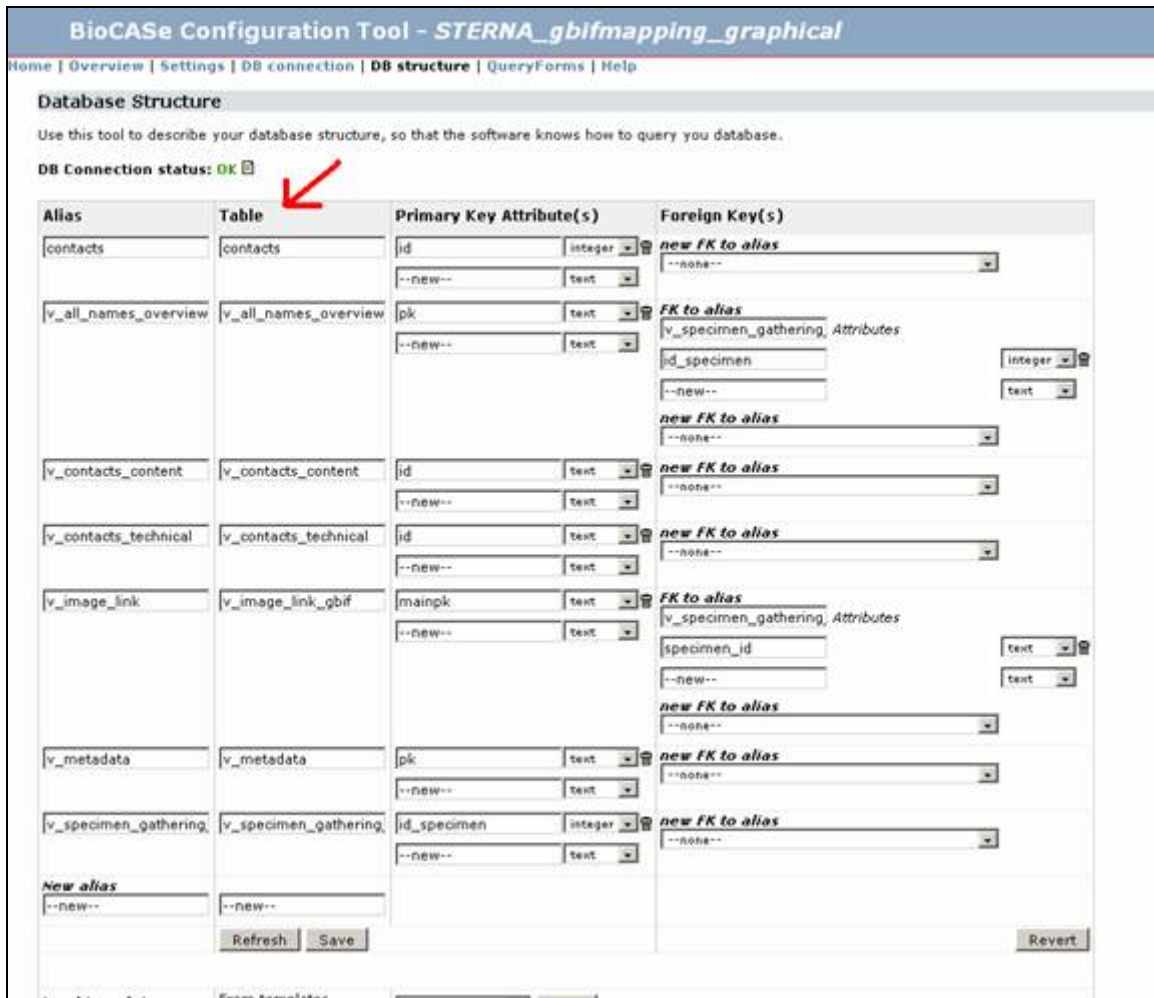
Fig. 8 Editing database structure

2.4 ABCD and ESE mappings

2.4.1 Reference documentation of ABCD concepts

Question: Where can I find a complete list of ABCD concepts? [25]

A concept explorer displaying the complete tree structure of ABCD is available at:
http://gbif.africamuseum.be/biocaseABCDexplorer/configtool/ratingsweb/
Select the option **"show documentation" -> "extended"** on the top right to see the documentation of the concepts. Please note that the concepts of ABCD extensions (like ABCDEFG for geosciences) are not listed there.

The developers wiki at BGBM lists the most commonly used fields (recommended for a first attempt of mapping): http://wiki.bgbm.org/bps/index.php/CommonABCD2Concepts

---

[25] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/ListABCDConcepts

The list of concepts provided by the TMG represents a very important reference for the mapping between ABCD in the prospect of ESE visualisation (the Europeana data model):
http://open-up.cybertaxonomy.africamuseum.be/Forum_ABCD2ESEmapping

The full list of ABCD concepts is also available on the TDWG website:
http://wiki.tdwg.org/twiki/bin/view/ABCD/AbcdConcepts

2.4.2 Correspondence between ABCD and ESE (Europeana Semantic Elements) fields

Question: What are the minimal required metadata fields for content provision in OpenUp!? [26]

Question: How do I map ABCD to ESE? [27]

An updated mapping between ABCD and ESE (the current Europeana data model) has been released in the beginning of 2012. There are two template documents related to this mapping:

1   A complete list of ABCD fields available for mapping
2   A "restricted data" list, containing the minimal set of ESE data fields that need to be supplied in order to have your metadata ingested into the Europeana Portal.

Three important notes on this mapping:

- ESE will be progressively replaced by the new EDM (European Data Model) in 2012, a fact that may lead to modification of these documents. The first ingest of OpenUp! data into Europeana will take place in February 2012, still via ESE.

- Europeana needs data either in **dc:title** or in **dc:description** as key identifier of the data. These two elements are actually used by Europeana for the indexing objects and to check their uniqueness in the system. **dc:title** corresponds to the ABCD fields containing the scientific name of a specimen, other elements can be mapped in ESE via **dc:europeana**

- The elements belonging to the DublinCore namespace (marked with the prefix dc:) can be potentially repeated. This allows the implementation of "one to many" relations between one multimedia document and several attributes depicting it, for instance several scientific names describing one object.

Attached documents to this deliverable:

- Complete ABCD/ESE mapping: map_ABCD206-ESE-120202-result-man-p.pdf
- Restricted ABCD/ESE mapping: map_ABCD206-ESE-120124-man-restricteddata_0.pdf

Important remark on metadata at general level:

Many providers map the general presentation of their data set into:
/DataSets/DataSet/Metadata/Description/Representation/Coverage and/or
/DataSets/DataSet/Metadata/Description/Representation/Details. But these fields are currently not mapped into ESE.

---

[26] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/Forum_ABCD2ESEmapping
[27] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/NewESEMappingFeb2012

2.4.3 Bugs and recommendations related to BioCASe and ABCD mapping [28]

2.4.3.1 Integer primary key on metadata

The metadata table storing metadata information at dataset level (name of the collection, contact coordinates of collection manager, address of the institution, etc.) should be declared with an integer primary key in the ACBD mapping, even if this table often contains just one line of record. This has to be done via the DB Structure tab of BioCASe (see Fig. 9).
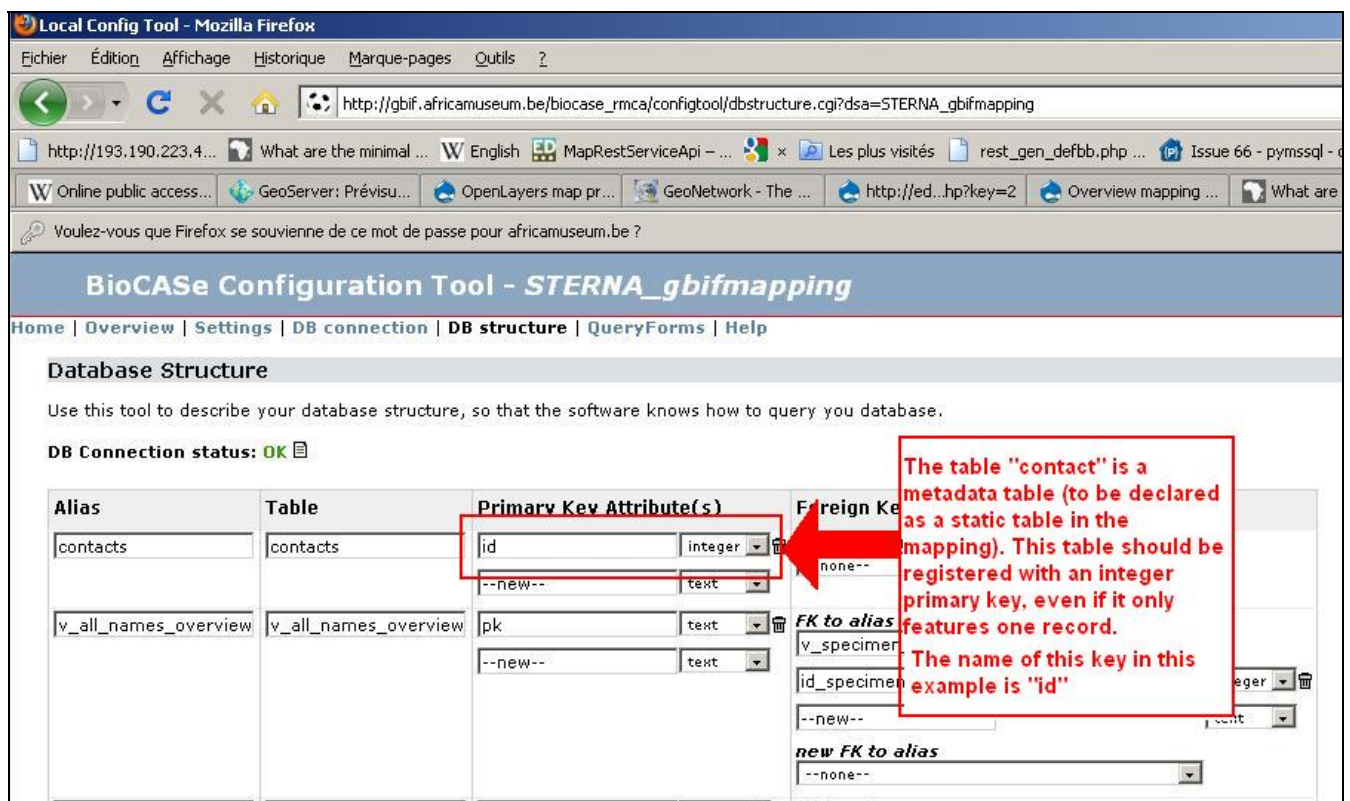
Fig. 9 Declaring primary key

Database managers could be tempted to define a "text" primary key on a field also containing data for convenience reasons (for instance the acronym of an institution), as this table contains a limited amount of records and is not bound via a foreign key to the specimen table. But this solution may generate problems and bugs at the level of the SQL drivers of the database, especially if the column contains diacritic signs. We experienced this issue with an MS SQL server database using the ODBC driver that could not retrieve data from a table when their primary key contained diacritic characters ("é","è", "ê","ï"). Using an integer primary key is a workaround to overcome this bug.

2.4.3.2 Version of pymssql driver for MS SQL Driver

If you use an MS SQL server in combination with pymssql[29], please ensure that you have at least the version 2.0.0 installed. We faced bugs with prior versions that cannot open a cursor on a remote server and browse the data via "fetchall".

---

[28] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/biocase_mapping_fixes

2.4.3.3 Declaration of several static tables

The SQL tables (or views) containing metadata must be declared inside of the "Static table aliases" drop down windows when you are mapping a schema (see Fig. 10).
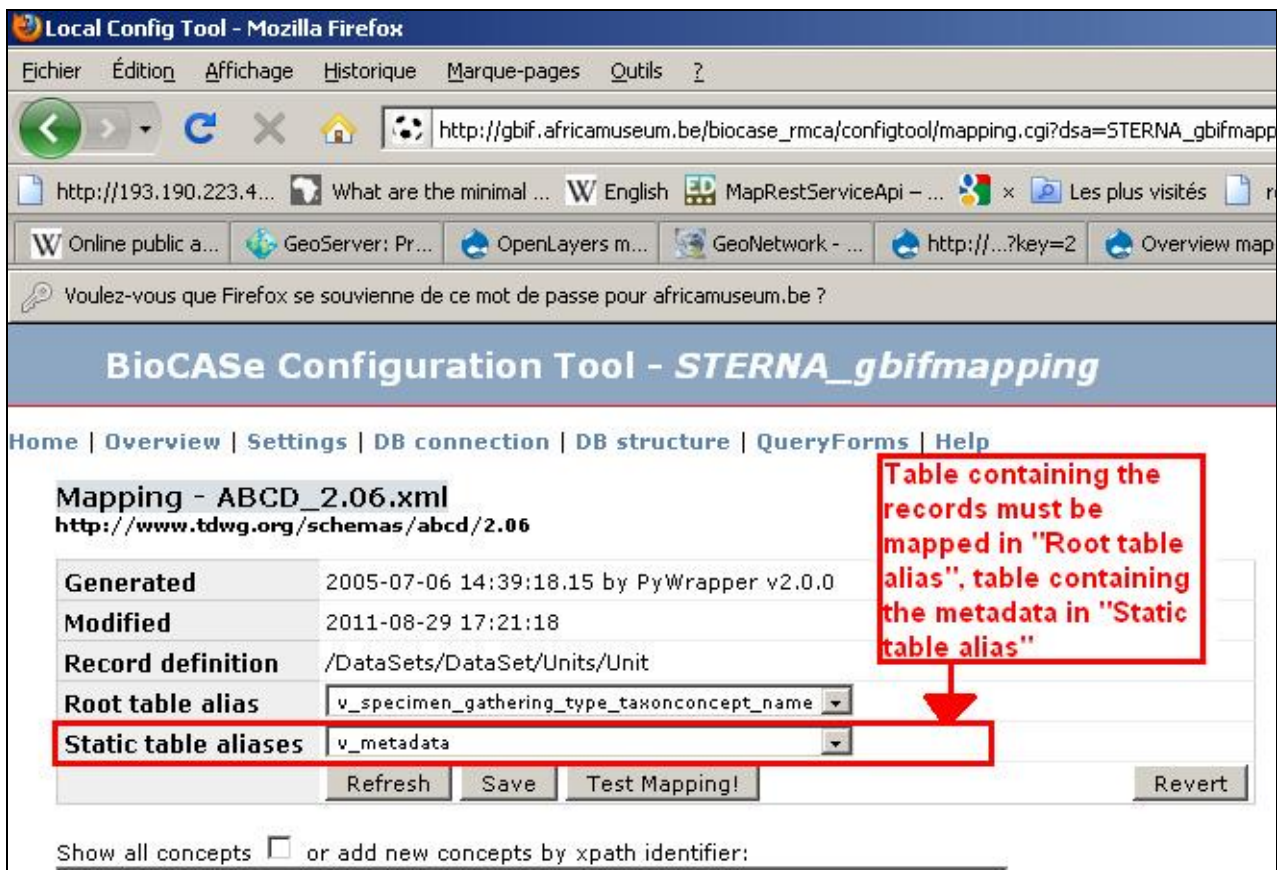
Fig. 10 Determining metadata alias

This windows provides space to declare only one static table, but it is actually possible to declare more than one metadata table just by editing the source XML of the mapping. On BioCASe 3.0 and for a mapping into ABCD 2.06, this XML document is located at: <BioCASe installation directory>/config/datasources/<name of the dataset>/cmf_ABCD_2.06.xml.

The static tables are defined inside of the cmf:settings/cmf:staticTableAlias element defined on top of the page (see Fig. 11).

Fig. 11 Static tables

If you manually added these additional static tables into this XML file in a BioCASe version prior to 2.6, the graphical interface of BioCASe allowing you to define the ABCD mapping of your database will not be synchronized anymore with this XML file. When you use this interface to modify an already existing ABCD mapping, it will just copy into the XML file the name of the first metadata table in cmf:staticTableAlias and erase the additional one. You will have to manually correct the XML file and re-add the missing 'cmf:StaticTableAlias' element after each modification to get your provider running.

2.4.3.4 Static string fields and concatenated elements cannot be handled as searchable elements by the BioCASe query engine

Only the fields coming directly from an SQL database can be searched with a text pattern or a value and used in a scan query. Those written as static string values are just displayed in the result set of a search query. Literal or concatenated fields are marked with the attribute "searchable="0"" in the presentation page of the mapping (when requesting the access point of the provider). See "Adding Mandatory ABCD Elements" in http://wiki.bgbm.org/bps/index.php/ABCD2Mapping (under the screenshot of the MappingEditor).

2.4.3.5 Problem with the escape character in static strings

When defining a static text value in the mapping (a text directly entered in the mapping form and not a database field), you should also avoid the sign "\" that may be wrongly handled by the XML parser of BioCASe (replace it by "\\" if you use it, for instance in file paths).

2.5 Preparation and mapping of multimedia content

2.5.1 Image requirements[30]

Question: Which image formats can be used in the OpenUp! Project? [31]

---

[30] See Europeana Portal Image Policy:
http://version1.europeana.eu/c/document_library/get_file?uuid=6b52d4be-6a4d-443a-842a-ab991bca2b1f&groupId=10602
[31] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/ImageConfig4OpenUp

Europeana generates thumbnails during their ingest process, and these generated thumbnails are cached in the Europeana system. Here are the preconditions for providing images to Europeana for thumbnail production:

The URL for the image file must link directly to an object which

1. is an **image file** (e.g. http://www.server.org/image.jpeg) and NOT an image embedded within a webpage,
2. has a width of at least **200px,**
3. is ideally a **jpg** file (or another image format supported by **ImageMagick**),
4. alternatively is a **pdf**, in which case the images will be created from the first page of this pdf. (Providers should ensure that the first page is a suitable image and not a blank page or a page containing the color scales etc.).

Providers should not supply a link to an image that is itself already the size of a thumbnail as this produces very poor quality results when it is used in the portal functions described. Similarly, the source image should not have a watermark nor should it be a local default thumbnail image as these do not give good results.

2.5.2 Placement of images to be mapped by BioCASe via ABCD

Question: Do the images submitted to OpenUp merely have to be in an accessible and stable webfolder?[32]

Images can be placed in a web folder, but the relation between images and reference objects (species, observation, or collection specimen) must be documented through an ABCD record whose identifier will be used to index the image at harvester level. It is suggested that the ABCD record stores the URL of the image, which should be as permanent as possible.

The most commonly used concepts to store links to images are located under the ABCD node: **/DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaObject**.

2.5.3 The license statement on thumbnails must be mapped as an URI in ABCD

Question: How do we communicate the kind of license to apply to our content to Europeana?[33]

Europeana requires data providers to explicitly mention the license they wish to apply to the content. "Content" stands for a **multimedia** document made available on the Internet by the **contributing institutions** (i.e. the URLs provided in the ESE elements **'europeana:isShownBy'** or **'europeana:isShownAt'**) as well as the small images, like **thumbnails** and previews generated to be displayed on the **Europeana** portal. The rights on metadata are defined at another level, in the Data Provider agreement.

The mention of this license has the form of an URL linking to the appropriate statements. Within the framework of OpenUp!, this statement must be notified as an URI into the ABCD elements **/DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaObject/IPR/Licenses/License/URI** which corresponds to the ESE element: **europeana:rights**

There are 12 possible types of a license statement[34], each one having its URL (only one statement can apply for each object): eight from Creative Commons[35]. Note: Europeana accepts license URLs with versions and

---

[32] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/MultimediaTag

[33] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/ese_license_in_abcd

[34] See Europeana Licensing Framework p17f.: http://version1.europeana.eu/c/document_library/get_file?uuid=3f32e3b6-b69b-40b5-92a6-32654571087f&groupId=10602

language flags (e.g.: http://creativecommons.org/licenses/by-nd/2.0/es for 'CC BT-ND version 2.0 in Spanish').

License examples	URL
Public Domain Mark	http://creativecommons.org/publicdomain/mark/1.0/
CC – Zero (copyright waiver)	http://creativecommons.org/publicdomain/zero/1.0/
CC BY (Attribution)	http://creativecommons.org/licenses/by/3.0/
CC BY-SA (Attribution, Share Alike)	http://creativecommons.org/licenses/by-sa/3.0/
CC BY-NC (Attribution for non commercial use only, others can apply different license on derivatives)	http://creativecommons.org/licenses/by-nc/3.0/
CC BY-NC-SA (Attribution and share alike for non-commercial use only)	http://creativecommons.org/licenses/by-nc-sa/3.0
CC BY-ND (Attribution without derivative)	http://creativecommons.org/licenses/by-nd/3.0
CC BY-NC-ND (Attribution for non commercial use only and without derivative))	http://creativecommons.org/licenses/by-nc-nd/3.0

Four from the Europeana rights statement

Europeana rights statements	URL
Rights Reserved - Free Access	http://www.europeana.eu/rights/rr-f/
Rights Reserved - Paid Access	http://www.europeana.eu/rights/rr-p/
Rights Reserved - Restricted Access	http://www.europeana.eu/rights/rr-r/
Unknown	http://www.europeana.eu/rights/unknown/

We strongly **recommend** the use of Creative Common licenses as this is the kind of licensing mentioned in the data provider agreement for OpenUp!

Note: The article **Creative Commons licenses and the non-commercial condition: Implications for the re-use of Biodiversity information** by Gregor Hagedorn, Daniel Mietchen, Robert A. Morris, Donat Agosti, Lyubomir Penev, Walter G. Berendsohn and Donald Hobern also provides very valuable information on the Creative Common license applied to biodiversity content.[36]

---

[35] See http://creativecommons.org/licenses/
[36] See www.pensoft.net/journals/zookeys/article/2189/creative-commons-licenses-and-the-non-commercial-condition-implications-for-the-re-use-of-biodiversity-information

2.5.4 Thumbnails for Europeana from non-publicly available images

Question: What if I want to provide thumbnails but protect the original document? Do I need a firewall for OpenUp!? [37]

You can use a single instance of the BioCASe provider to publish multimedia documents along with their metadata. This is actually the most common scenario. The URLs of the multimedia document are in this case referenced in the appropriate ABCD element.

These documents can be stored on the same physical server as well as the BioCASe provider or on a different one. In this case, like for any other website you could host, you must ensure that your firewall and Proxy allow any incoming connection on the port 80 (standard HTTP port) as well for the BioCASe provider and the images.

Providing thumbnails by protecting the original document:

Alternatively, you may wish to provide thumbnails to Europeana and OpenUp! by preventing at the same time the access to the original high-resolution material.

It is suggested that you set up two parallel instances of the BioCASe provider during the time of the harvesting of data by Europeana:

- the main one containing the text metadata, and a link to the records of the second BioCASe provider. This would be the permanent access point to the data.
- the second one containing the links to the high resolution images, that would be only available from a limited number of service and for a limited duration. By configuring the firewall or Apache HTTP the appropriate way, you can restrict to a limited scope the IP addresses used for incoming connection. This technique can be used to ensure that only the Europeana harvesting system can access the original document.

---

[37] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/ProtectionAndFirewall
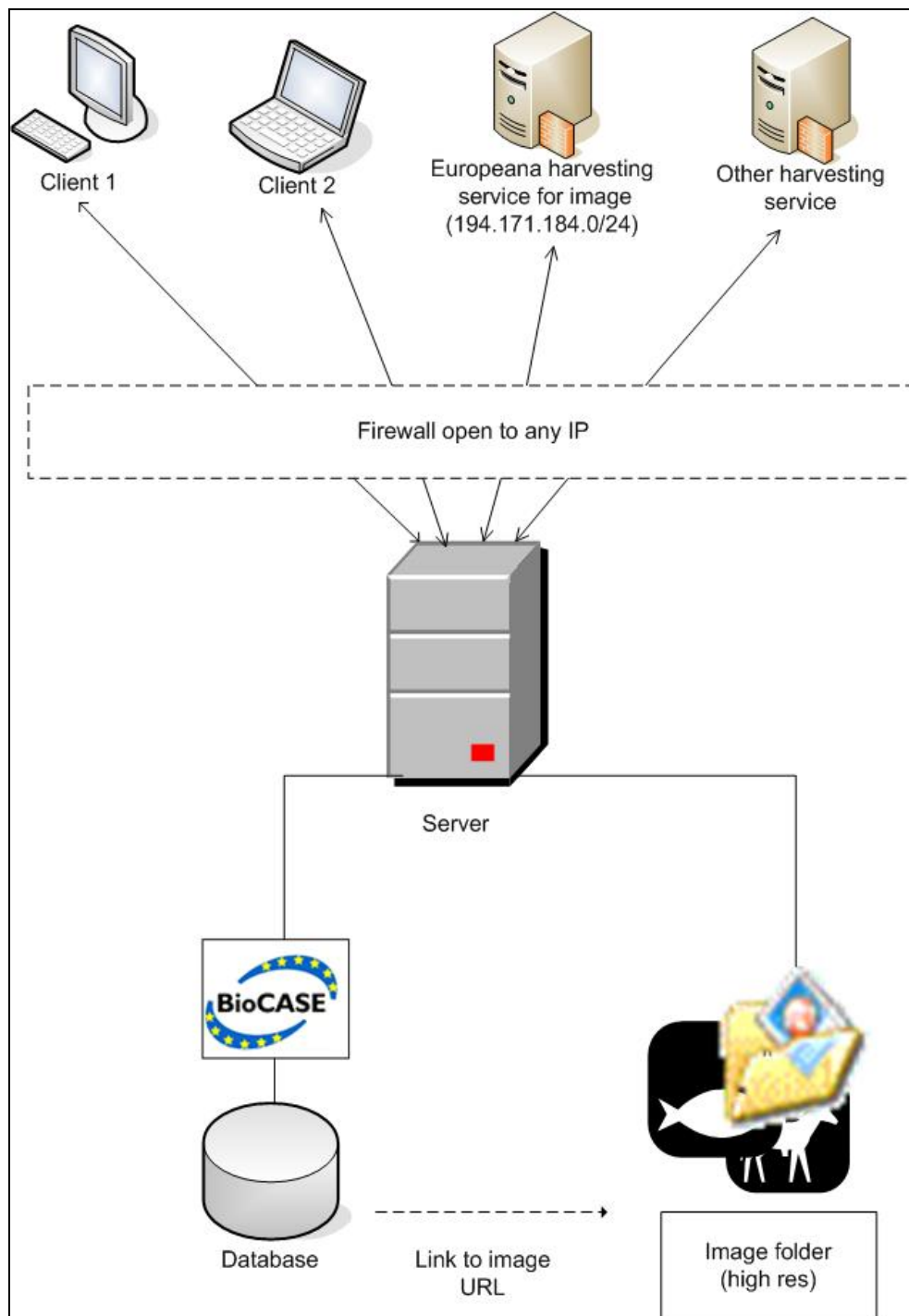
Fig. 12 Firewall and proxy configuration for a single instance of the BioCASe provider (common for metadata and large images)

The deny and allow directives of Apache HTTP can be used to define two different policies for each BioCASe installation if they run on the same server (it is also possible to filter IPs through the proxy module of Apache):

- http://httpd.apache.org/docs/2.3/en/mod/mod_access_compat.html#allow

- http://httpd.apache.org/docs/2.3/en/mod/mod_access_compat.html#deny
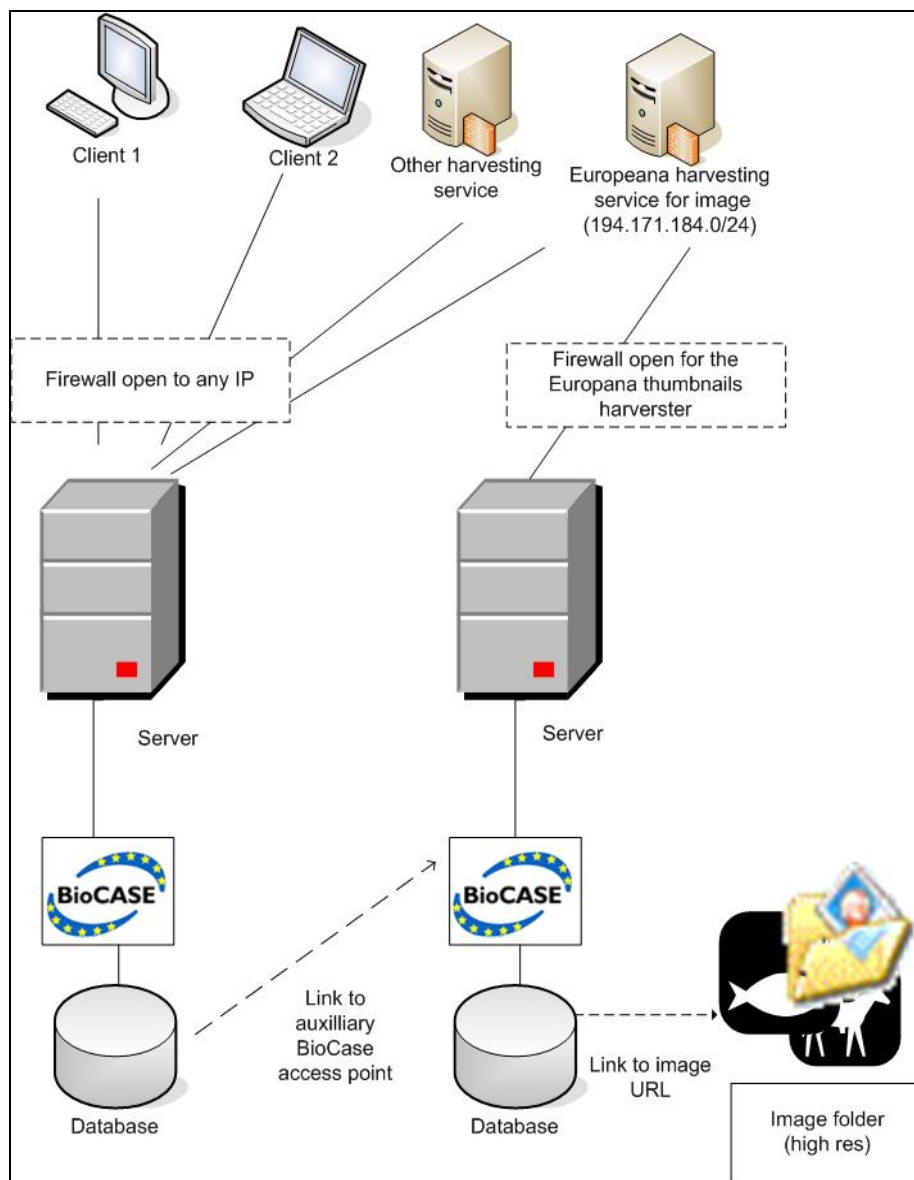
Fig. 13 Firewall and proxy configuration for two instance of the BioCASe provider (the second being intended to prevent direct access to original multimedia document by also allowing harvesting of images by Europeana)

Note: this figure describes two different servers but it is also possible to install two different BioCASe on the same server and limit the access to the second one by defining rule in the modules of Apache HTTP.

The primary installation of the BioCASe provider should be open to any IP, but the BioCASe provider responsible for delivering images for thumbnail-creation (that can be either the primary provider software installation or a separate one for the images only) needs to be accessible from the IP number **194.171.184.0/24**, which is the Europeana server taking care of the thumbnail generation. Please ensure that the firewall and proxy servers of you institution allow incoming connection to the BioCASe provider serving thumbnail to this range of IP, at least for the server containing the image.

2.5.5 Relation between images and objects (representation of cardinality)

Question: Can I provide several images describing a single unit to OpenUp!? [38]

BioCASe can handle several images for each specimen, but **ESE** (the current data model of Europeana) can currently handle only displays of one image per specimen. However, if you have several images for each specimen we suggest to already define a complete mapping in the prospect of the implementation of the next **EDM** (the new Europeana Data Model) that shall be able to ingest several images per metadata record.

Once the mapping of images has been defined, the local query tool of the BioCASe provider can also display the images in its query tool on-line interface, e.g.:

http://gbif.africamuseum.be/biocase_rmca/querytool/details.cgi?dsa=STERNA_gbifmapping&detail=unit&wrapper_url=http://gbif.africamuseum.be/biocase_rmca/pywrapper.cgi?dsa=STERNA_gbifmapping&schema=http://www.tdwg.org/schemas/abcd/2.06&inst=RMCA&col=Aves&cat=RMCA%20A.8972

The detailed steps are:

1. The URL of the image should be mapped under the
   /DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaObject/FileURI element
2. Currently the harvester can only ingest one image into the Europeana portal. If you select a "preferred" image to be displayed in Europeana, it should be put first in the list of displayed images in the ABCD XML. The BioCASe provider currently does not have any mechanism to sort elements in a list, but this can be implemented at database level by using an SQL view with an ORDER BY clause on a flagged column. If you cannot create immediately this SQL view, you can first focus on the basic configuration before working on this. The sort order can be defined later on with minimal changes in the configuration of an ABCD dataset.
3. Important: Do you have a website where all the images are presented together with the specimen information? If so, then please map the URL to this web site to
   /DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaObject/ProductURI. This URL can already be provided as a link detailing the image with the current ESE model.

2.5.6 Interaction with image servers

Question: Can images be served from an image server like Morphbank installation (e.g. http://morphbank.digitarium.fi/)? [39]

For thumbnail creation, the Provider must register a URL pointing to a raw file and not the URL of a viewer-application based e.g. on Flash, nor of a webpage containing the image. However submitting images with Morphbank should not be a problem as it seems that this tool gives a link to the source image (in different formats), just by adding a suffix to the URL:

- HTML page: http://morphbank.digitarium.fi/?id=3500032
- Original TIFF: http://mbimages.digitarium.fi/?id=3500032&imgType=tiff
- High Res JPEG : http://mbimages.digitarium.fi/?id=3500032&imgType=jpeg
- Medium resolution JPEG: mbimages.digitarium.fi/?id=3500032&imgType=jpg

One of the images links should be documented in the ABCD mapping.

---

[38] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/OpenUpMultipleImages

[39] See OpenUp! Helpdesk Forum: http://open-up.cybertaxonomy.africamuseum.be/node/212

3 APPENDIX

3.1 List of figures

3.2 List of questions

Question	Chapter
What is an IP address?	2.1.1 What is an IP address
DNS in practice: what's and how it works?	2.1.2 What does DNS means
What is a SQL database?	2.1.3 Definition of SQL databases
What is and how do relational databases work?	2.1.4 Explanation and Glossary in regard to relational databases
What's the big picture of OpenUp! and its different actors?	2.2.1 OpenUp! A first technical approach

Question	Chapter
Which XML schemas do I use with the BioCASe provider? What is the ABCD schema?	2.2.2 BioCASE and the XML schemas for biodiversity metadata
Does BioCASe store the data? Is it a database system?	2.3.2 BioCASe as a web service wrapping databases
How do I install BioCASe?	2.3.3 BioCASe installation (prerequisites and procedure)
Can I install the BioCASe provider/PyWrapper on my own computer?	2.3.4 Installing the BioCASeprovider/PyWrapper on your computer
Can I install the BioCASe provider/PyWrapper on my own computer?	2.3.5 BioCASe on a virtual machine
How do I connect a SQL database to BioCASe?	2.3.6 BioCASe and its Back-end database
Should I use database tables or views for publishing data with BioCASe?	2.3.7 Preparation of the database by creating dedicated SQL views
How do I configure my server to let users access BioCASE?	2.3.8 IP, DNS and the Connection of the BioCASe provider to networks
How do I install BioCASE?	2.1.1 Installation of BioCASe
How do I register and create a mapping for a new dataset?	2.1.2 Registration and mapping of a new dataset
Can I replicate datasets between several implementation of the BioCASe provider?	2.3.9 Editing and copy/pasting the configuration file
Where can I find a complete list of ABCD concepts?	2.4.1 Reference documentation of ABCD concepts
What are the minimal required metadata fields for content provision in OpenUp?	2.4.2 Correspondences between ABCD and ESE (Europeana data model) fields
How do I map ABCD to ESE?	2.4.2 Correspondences between ABCD and ESE (Europeana data model) fields
Which image formats can be used in the OpenUp! project?	2.5.1 Image requirements
Do the images submitted to OpenUp! merely have to be in an accessible and stable webfolder?	2.5.2 Placement of images to be mapped by BioCASE via ABCD
How do we communicate the kind of license to apply to our content (e.g.: thumbnails) to Europeana?	2.5.3 The license statement on thumbnails must be mapped as an URI in ABCD
What if I want to provide thumbnails but protect the original document? Do I need a firewall for OpenUp!?	2.5.4 Thumbnails for Europeana from non-publicly available images
Can I provide several images describing a single unit to OpenUp!?	2.5.5 Relation between images and object (representation of cardinality)
Can the images be served from an image server like Morphbank installation?	2.5.6 Interaction with image servers

3.3 Glossary

Keyword	Short Description
ABCD (Access to Biological Collections Data)	The XML hierarchical schema mostly used to exchange data within a network using BioCASE software. It has about 1200 fields (or 'concepts'). ABCD is the schema being used in the context of OpenUp!.
BioCASE (Biological Collection Access Service Europe)	Web application installed on top of the database, that exposes the data of the source database to the web. It is intended for databases containing curatorial, taxonomical and ecological information. Used by OpenUp! to feed Europeana with biodiversity data.
DarwinCore	Another XML schema that BioCASE can handle. Designed by GBIF, its flat structure contains about 120 fields and can handle several extra extensions.
DNS (Domain Name System)	Internet service linking an Internet IP address with one or several alphabetical and human readable domain names
ESE (Europeana Semantic elements)	Metadata set used to describe cultural heritage objects in Europeana using its particular XML schema. OpenUp! provides the tools to convert Metadata from ABCD to ESE. ESE is expected to be progressively replaced by the new EDM (Europeana Data Model) in 2012.
Europeana	Europeana is "Europe's digital library, archive and museum". It incorporates metadata and content from European cultural and natural history institutions, enabling people to explore those digital resources
GBIF (Global Biodiversity Information Facility)	International organisation that is working to make the world's biodiversity data accessible everywhere in the world. GBIF and its many partners work to mobilise the data, and to improve search mechanisms, data and metadata standards, web services, and the other components of an Internet-based information infrastructure for biodiversity.
IP address	An Internet Protocol address (IP address) is a unique number, identifying the devices belonging to the same computer network.
Mapping	Process of defining a correspondence between the structure of a database and the structure of the XML schema. In our context, it means to assign what field of my database is equivalent to the different concepts in the ABCD schema. BioCASE enables this mapping through its interface.
OpenUp!	European project that will make multimedia data related to biodiversity available to the general public through Europe's digital library, archive and museum: Europeana. OpenUp! is one of the projects that serves as an aggregator for content to be included into Europeana.
Relational database	A software storing data in a multi-tabular structure. Its information can be defined, updated and queried using SQL language.
SQL (Structured Query Language)	Programming language for relational databases, used to define its table structure, query and modify its data, as well as handle the security and integrity

Keyword	Short Description
XML (Extensible Markup Language)	Markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable. Both ABCD and DarwinCore are XML schemas used by BioCASE (and others) to share biodiversity information on the web.

3.4 Additional information resources on OpenUp!, Europeana, GBIF, BioCASE and the TDWG standard ABCD

3.4.1 OpenUp!

The technical documentation of the project *Opening up the Natural History Heritage for Europeana – OpenUp!* will be updated on the OpenUp! Heldpesk[40]. All public components and deliverable are available from the OpenUp! project website[41]. Further information on activities of the project and texts about it can be accessed in the documents[42] section and news[43] section.

3.4.2 Europeana

On the Europeana v1.0 project documents website[44], users can browse the documentation by thematic tabs for material on business and technical documents, as well as presentations and white papers. There are also usages guides for public domain works and a whole section on material on the new Europeana data exchange agreement, which also features a long Q and A section and a summary of the consultation process[45]. Three documents should be highlighted in particular:

- *Europeana Aggregators' Handbook[46]* describes technical, organisational and financial aspects of being/becoming a Content Provider for Europeana.
- The New *Europeana Data Exchange Agreement[47]* (DEA) is the legal document describing under which conditions data is made available to Europeana and under which conditions Europeana will make these data available from July 2012 onwards.
- *The Europeana Licensing Framework[48]* explains DEA and the Creative Commons Zero license, The Data Use Guidelines and the edm:rights field of the Europeana Data Model (EDM).

---

[40] OpenUp! Helpdesk: http://openup.helpdesk.africamuseum.be/forum

[41] OpenUp! Project Website Public deliverables and components: http://open-up.eu/content/deliverables-and-components-pu

[42] OpenUp! Project Website Documents: http://open-up.eu/category/menu/outcomes/documents

[43] OpenUp! Project Website News: http://open-up.eu/category/menu/news

[44] Europeana v1.0 documents: http://version1.europeana.eu/web/europeana-project/documents

[45] DEA consultation process: http://version1.europeana.eu/web/europeana-project/newagreement-consultation/

[46] Aggregators' Handbook http://www.version1.europeana.eu/c/document_library/get_file?uuid=94bcddbf-3625-4e6d-8135-c7375d6bbc62&groupId=10602

[47] Europeana DEA: http://version1.europeana.eu/c/document_library/get_file?uuid=deb216a5-24a9-4259-9d7c-b76262e4ce55&groupId=10602

3.4.3 GBIF

The centralized source of documentation for the *Global Biodiversity Information Facility* is the GBIF Online Research Center[49] (ORC). Users can browse the documentation by thematic priorities: GBIF Welcome Box, BIF Start Up Kit and Advanced Kit, Training Resources, Tools, as well as Strategic Documents. A second way of accessing those documents is by categories: Recently Added, Featured, Most downloaded and Best rated. On GBIF Resources[50] a set of technical resources intended for machine read can be found, and GBIF Tools[51] provides an index of digital tools and software.

3.4.4 BioCASE

Biodiversity Collection Access System Europe refers both to a former European project and a suite of related open source XML web-services for publishing biodiversity information. The BioCASe provider software, represents the backbone of the technical infrastructure of OpenUp! and has to be installed by the content providing institutions. The PyWrapper Wiki[52] of the Botanischer Garten und Botanisches Museum Berlin-Dahlem (BGBM) is the primary source of information for partners in the process of installing the BioCASe software in their institution. On the BioCASe Provider Software[53] (BPS) website, the latest version of the software can be downloaded. BPS is closely related to TDWG's ABCD schema.

3.4.5 TDWG

The centralized source of documentation for TDWG's biodiversity information standards is the TDWG Wiki[54] where the documentation on the ABCD[55] schema (Access to Biological Collection Data) can be found. This resource also links to ABCD extension, the most relevant for OpenUp! being probably the extension for earth science and palaeontology ABCDEFG[56] (Access to Biological Collection Databases Extended for Geosciences) of the GeoCASE project.

A list of all TDWG standards can also be viewed at http://www.tdwg.org/standards/ . This is also a reference point for the status of the standard - draft/current/prior.

3.5 ABCD/ESE mapping lists

Attached as separate files are - as provide by the OpenUp! Technology Management Group (TMG) - the

- Mapping of the full data set: map_ABCD206-ESE-120202-result-man-p.pdf
- Mapping of the restricted data set: map_ABCD206-ESE-120124-man-restricteddata_0.pdf

---

[48] Europeana Licensing Framework: http://version1.europeana.eu/c/document_library/get_file?uuid=3f32e3b6-b69b-40b5-92a6-32654571087f&groupId=10602

[49] GBIF Online Research Center: http://www.gbif.org/nc/orc/

[50] GBIF Resources: http://rs.gbif.org/

[51] GBIF Tools: http://tools.gbif.org/

[52] PyWrapper Wiki: http://wiki.bgbm.org/bps/index.php/Main_Page

[53] BioCASe Provider Software: http://www.biocase.org/products/provider_software/

[54] TDWG Wiki: http://wiki.tdwg.org/

[55] ABCD Schema documentation: http://wiki.tdwg.org/ABCD/

[56] ABCDEFG Schema documentation: http://www.geocase.eu/efg.asp

	ESE	ABCD 2.06	Apr.11	Option (or	Comment	Obligation
1	dc:title	/DataSets/DataSet/Units/Unit/Identifications/Identification/Result/TaxonIdentified/ScientificName/FullScientificNameString	135x		take the value with: PreferredFlag	Mandatory to provide one of these two (dc:title or dc:description)
	dc:title	/DataSets/DataSet/Units/Unit/SpecimenUnit/NomenclaturalTypeDesignations/NomenclaturalTypeDesignation/TypifiedName/FullScientificNameString	12x		Append: (Typus) repeatable	
2	dc:description	/DataSets/DataSet/Units/Unit/Notes	47x			
	dc:description	/DataSets/DataSet/Units/Unit/SpecimenUnit/Preparations/Preparation/PreparationType	44x			
	dc:description	/DataSets/DataSet/Units/Unit/Sex	38x		evt. Append (sex)	
	dc:description	/DataSets/DataSet/Units/Unit/KindOfUnit	30x			
	dc:description	/DataSets/DataSet/Units/Unit/Gathering/Notes	13x		check what is usually used in this field!	
	dc:description	/DataSets/DataSet/Units/Unit/Age	9x			
	dc:description	/DataSets/DataSet/Units/Unit/SpecimenUnit/Marks/Mark/MarkText	6x			
3	dc:type	/DataSets/DataSet/Units/Unit/RecordBasis	109x		(PreservedSpecimen, LivingSpecimen, FossileSpecimen, OtherSpecimen, HumanObservation, MachineObservation, DrawingOrPhotograph, MultimediaObject)	Mandatory to provide one of these two (dc:type or dcterms:spatial)
4	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/LocalityText	113x			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/SiteCoordinateSets/SiteCoordinates/CoordinatesLatLong/LatitudeDecimal	100x		EDM Preferred display sequence: date, method, coordinates (lat+long), errordistance, accuracystatement (alternative if no coordinates used: coordinatetext)	
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/SiteCoordinateSets/SiteCoordinates/CoordinatesLatLong/LongitudeDecimal	100x		EDM	
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Country/Name	81x			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Country/ISO3166Code	75x			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Altitude/MeasurementOrFactAtomised/LowerValue	61x			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/NamedAreas/NamedArea/AreaName	48x			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/SiteCoordinateSets/SiteCoordinates/CoordinatesLatLong/CoordinateErrorDistanceInMeters	44x		put it after longitude	
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Altitude/MeasurementOrFactAtomised/UpperValue	39x			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/NamedAreas/NamedArea/AreaClass	30x		qualifies Area (put it in brackets after the Area Name)	
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Altitude/MeasurementOrFactAtomised/UnitOfMeasurement	18x		if not there put m as default value	
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/AreaDetail	18x			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Biotope/Text	18x			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Depth/MeasurementOrFactAtomised/LowerValue	13x			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/SiteCoordinateSets/SiteCoordinates/CoordinateMethod	6x			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/SiteCoordinateSets/SiteCoordinates/CoordinatesLatLong/SpatialDatum	6x		put it in brackets after the longitude	
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Biotope/Name	5x			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Stratigraphy/ChronostratigraphicTerms/ChronostratigraphicTerm/Term	4x			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Depth/MeasurementOrFactAtomised/UnitOfMeasurement	3x		default m (meter) if no value here	
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Aspect/Ordination	2x			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Aspect/Text	2x			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Height/MeasurementOrFactAtomised/UnitOfMeasurement	2x		default m (meter) - if not used	
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Height/MeasurementOrFactAtomised/UpperValue	2x			

Update 24.1.2012

	ESE	ABCD 2.06 Apr.11	Option (or	Comment	Obligation
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/NearNamedPlaces/Na medPlaceRelation/NearNamedPlace	2x		
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Height/MeasurementO rFactAtomised/LowerValue	1x		
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Altitude/Measurement OrFactAtomised/LowerValue/accuracy		not yet in use	
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Altitude/Measurement OrFactAtomised/UpperValue/accuracy		not yet in use	
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Stratigraphy/Lithostrati graphicTerms/LithostratigraphicTerm/Term			
	dcterms:spatial	/DataSets/DataSet/Units/Unit/Gathering/Stratigraphy/Biostratig raphicTerms/BiostratigraphicTerm/Term			
5	europeana:dataProvider	/DataSets/DataSet/Metadata/Owners/Owner/Organisation/Na me/Representation/Text	106x	if no unit owner is given	mandatory
	europeana:dataProvider	/DataSets/DataSet/Units/Unit/Owner/Organisation/Name/Repr esentation/Text	5x	this is the preferred ABCD field for dataprovider if not used than take the value from /DataSets/DataSet/Metadata/Owners/ Owner/Organisation/Name/Represent ation/Text	
6	europeana:isShownAt	/DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaOb ject/ProductURI	7x	mandatory to provide either ProductURI or FileURI - in any case the Format information for the digital object has to be provided even if there is only a ProductURI in your ABCD data (see below)!!	mandatory to provide either europeana:isShownBy or europeana:isShownAt
7	europeana:isShownBy	/DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaOb ject/FileURI	43x		
8	europeana:rights	/DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaOb ject/IPR/Licenses/License/URI		License.URI [Priority: 1.MultimediaObject.IPRStatements, 2. Unit.IPRStatements and 3. Dataset.Metadata.IPRStatements)	mandatory
9	europeana:type	/DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaOb ject/Format		for correct object type mapping in ESE the Format field has to be filled in! (you may use one of these values: image, video, audio, text or another standard mimetype) NB: *Europeana, at the time of the creation this mapping, does need a qualification of any object as: **image, video, audio, text***	mandatory
10	dc:identifier	/DataSets/DataSet/Units/Unit/SourceInstitutionID 135x		all three together: unique identifier	Mandatory
	dc:identifier	/DataSets/DataSet/Units/Unit/SourceID 135x		all three together: unique identifier	
	dc:identifier	/DataSets/DataSet/Units/Unit/UnitID 135x		all three together: unique identifier	
11	dc:contributor	/DataSets/DataSet/Units/Unit/Gathering/Agents/GatheringAge nt/AgentText	22x		optional
	dc:contributor (role: identifier)	/DataSets/DataSet/Units/Unit/Identifications/Identification/Ident ifiers/IdentifiersText	2x	preferred flag	
	dc:contributor (role:collector)	/DataSets/DataSet/Units/Unit/Gathering/Agents/GatheringAge nt/Person/FullName	88x	Append: (/DataSets/DataSet/Units/Unit/Collect orsFieldNumber), (/DataSets/DataSet/Units/Unit/Gather ing/Code)	
	dc:contributor (role:collector)	/DataSets/DataSet/Units/Unit/Gathering/Agents/GatheringAge ntsText	9x	all contrib utors should be include d here	attribute: primary etc. most important

	ESE	ABCD 2.06	Apr.11	Option (or	Comment	Obligation
	dc:contributor (role:identifier)	/DataSets/DataSet/Units/Unit/Identifications/Identification/Identifiers/Identifier/PersonName/FullName	70x		only for identifiaction with preferred flag after (role:identifier) put: (/DataSets/DataSet/Units/Unit/Identifications/Identification/Date/ISODateTimeBegin(if not available put: (/DataSets/DataSet/Units/Unit/Identifications/Identification/Date/DateText)	O
	dc:contributor	/DataSets/DataSet/Units/Unit/CollectorsFieldNumber	20x		Append after: dc:contributor field number(collector)	
	dc:contributor	/DataSets/DataSet/Units/Unit/Identifications/Identification/Date/ISODateTimeBegin	8x		dc:contributor (role:identifier)	
	dc:contributor	/DataSets/DataSet/Units/Unit/Gathering/Code	4x		dc:contributor field number(collector)	
12	dc:date	/DataSets/DataSet/Units/Unit/Gathering/DateTime/DateText	82x		EDM if the recordbasis is NOT FossilSpecimen use> dcterms:temporal	Optional
	dc:date	/DataSets/DataSet/Units/Unit/Gathering/DateTime/ISODateTimeBegin	74x		EDM if the recordbasis is NOT FossilSpecimen use> dcterms:temporal	
	dc:date	/DataSets/DataSet/Units/Unit/Gathering/DateTime/ISODateTimeEnd	36x		EDM if only timebegin: put timebegin (just one date); if timeend available put: timebegin - timeend if the recordbasis is NOT FossilSpecimen than use> dcterms:temporal	
		/DataSets/DataSet/Units/Unit/Identifications/Identification/Date/DateText	24x		dc:date if timebegin not used	
13	dc:identifier	/DataSets/DataSet/DatasetGUID	20x			
	dc:identifier	/DataSets/DataSet/Units/Unit/UnitGUID	4x			
14	dc:relation	/DataSets/DataSet/Units/Unit/Associations/UnitAssociation/AssociatedUnitID	2x		EDM	Optional
	dc:relation	/DataSets/DataSet/Units/Unit/Associations/UnitAssociation/AssociatedUnitSourceInstitutionCode	2x		EDM	
	dc:relation	/DataSets/DataSet/Units/Unit/Associations/UnitAssociation/AssociatedUnitSourceName	2x		EDM	
	dc:relation	/DataSets/DataSet/Units/Unit/Associations/UnitAssociation/AssociationType	2x		EDM	
	dc:relation	/DataSets/DataSet/Units/Unit/Associations/UnitAssociation/Comment	2x		EDM	
	dc:relation	/DataSets/DataSet/Units/Unit/Assemblage/UnitAssemblage/AssemblageID			EDM	
	dc:relation	/DataSets/DataSet/Units/Unit/Assemblage/UnitAssemblage/AssemblageName			EDM	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/TermsOfUseStatements/TermsOfUse/Text	92x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/Copyrights/Copyright/Text	56x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/Disclaimers/Disclaimer/Text	52x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/Acknowledgements/Acknowledgement/Text	44x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/Copyrights/Copyright/URI	40x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/Citations/Citation/Text	37x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/Disclaimers/Disclaimer/URI	34x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/IPRDeclarations/IPRDeclaration/Text	34x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/Licenses/License/Text	26x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/Acknowledgements/Acknowledgement/Details	20x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaObject/IPR/Copyrights/Copyright/Text	18x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaObject/IPR/Licences/License/Text			Concatenate Statement.Text + Statement.Details + Statement.URI	Update 02.02.201
	dc:rights	/DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaObject/IPR/TermsOfUseStatements/TermsOfUse/Text	15x		Concatenate Statement.Text + Statement.Details + Statement.URI	

	ESE	ABCD 2.06	Apr.11	Option (or	Comment	Obligation
15	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/Acknowledgements/Acknowledgement/URI	11x		Concatenate Statement.Text + Statement.Details + Statement.URI	Optional
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/IPRDeclarations/IPRDeclaration/URI	11x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/TermsOfUseStatements/TermsOfUse/Details	10x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/Disclaimers/Disclaimer/Details	9x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/TermsOfUseStatements/TermsOfUse/URI	9x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/Citations/Citation/URI	8x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/Copyrights/Copyright/Details	8x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/IPRDeclarations/IPRDeclaration/Details	8x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/Licenses/License/Details	8x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Metadata/IPRStatements/Citations/Citation/Details	7x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Units/Unit/IPRStatements/Copyrights/Copyright/Text	2x		overrules metadata rights Concatenate Statement.Text +	
	dc:rights	/DataSets/DataSet/Units/Unit/IPRStatements/Citations/Citation/Text	1x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Units/Unit/IPRStatements/Licenses	1x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Units/Unit/IPRStatements/Licenses/License/Text	1x		Concatenate Statement.Text + Statement.Details + Statement.URI	
	dc:rights	/DataSets/DataSet/Units/Unit/IPRStatements/TermsOfUseStatements/TermsOfUse/Text	1x		Concatenate Statement.Text + Statement.Details + Statement.URI	
16	dc:source	/DataSets/DataSet/Metadata/Description/Representation/Title	133x			Optional
17	dcterms:provenance	/DataSets/DataSet/Units/Unit/SpecimenUnit/History/PreviousUnitsText	1x			Optional
18	europeana:object				used for thumbnail creation possible solutions: extension for ABCD Europeana:object, or encryption within unit:notes	Optional
19	europeana:unstored			[could be used for description items]		Optional

	ESE	ABCD 2.06	Apr.11	Option (or	Comment	Obligation	
1	dc:title	/DataSets/DataSet/Units/Unit/Identifications/Identification/Result/TaxonIdentified/ScientificName/FullScientificNameString	135x		take the value with: PreferredFlag	Mandatory to provide one of these two (dc:title	*organism name*
	dc:title	/DataSets/DataSet/Units/Unit/SpecimenUnit/NomenclaturalTypeDesignations/NomenclaturalTypeDesignation/TypifiedName/FullScientificNameString	12x		Append: (Typus) repeatable		*organism name*
3	dc:type	/DataSets/DataSet/Units/Unit/RecordBasis	109x		(PreservedSpecimen, LivingSpecimen, FossileSpecimen, OtherSpecimen, HumanObservation, MachineObservation, DrawingOrPhotograph, MultimediaObject)	Mandatory to pro	*kind of object*
5	europeana:dataprovider	/DataSets/DataSet/Metadata/Owners/Owner/Organisation/Name/Representation/Text	106x		if no unit owner is given	mandatory	*institutional metadata*
	europeana:dataprovider	/DataSets/DataSet/Units/Unit/Owner/Organisation/Name/Representation/Text	5x		this is the preferred ABCD field for dataprovider if not used than take the value from /DataSets/DataSet/Metadata/Owners/Owner/Organisation/Name/Representation/Text		*institutional metadata*
6	europeana:isShownAt	/DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaObject/ProductURI	7x		mandatory to provide either ProductURI or FileURI - in any case the Format information for the digital object has to be provided even if there is only a ProductURI in your ABCD data (see below)!!	mandatory to provide either europeana:isShownBy or europeana:isShownAt	*URL that links to the multimedia content*
7	europeana:isShownBy	/DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaObject/FileURI	43x				*URL that links to the multimedia content*
8	europeana:rights	/DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaObject/IPR/Licenses/License/URI			License.URI [Priority: 1.MultimediaObject.IPRStatements, 2. Unit.IPRStatements and 3. Dataset.Metadata.IPRStatements)	mandatory	*IPR statements regarding the linked object*
9	europeana:type	/DataSets/DataSet/Units/Unit/MultiMediaObjects/MultiMediaObject/Format			for correct object type mapping in ESE the Format field has to be filled in! (you may use one of these values: image, video, audio, text or another standard mimetype) NB: *Europeana, at the time of the creation this mapping, does need a qualification of any object as: **image, video, audio, text***	mandatory	*Type of the multimedia object*
10	dc:identifier	/DataSets/DataSet/Units/Unit/SourceInstitutionID	135x		all three together: unique identifier	Mandatory	*institutional metadata*
	dc:identifier	/DataSets/DataSet/Units/Unit/SourceID	135x		all three together: unique identifier		*institutional metadata*
	dc:identifier	/DataSets/DataSet/Units/Unit/UnitID	135x		all three together: unique identifier		*institutional metadata*