OPEN UP! Opening Up the Natural History Heritage for Europeana

COMPONENT REPORT

Project Acronym: **OpenUp!**

Grant Agreement No: **270890**

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

C4.5.0 - Model for the integration of content from the areas of palaeontology and mineralogy

Revision: **4a [Final]**

Authors:

K. Zágoršek	NM	1st/ 2nd draft
J. Sejkora	NM	1st/ 2nd draft
J. Hoffmann	MfN	3rd/4th draft

Project co-funded by the European Commission within the ICT Policy Support Programme		
Dissemination Level		
P	Public	X
C	Confidential, only for members of the consortium and the Commission Services	

Revision History

Revision	Date	Author	Organization	Description
1	22.10.2011	K. Zágoršek	NM	
	24.10.2011	A. Güntsch	BGBM	Comments
	25.10.2011	Kroh	NHMW	Comments and suggestions
	25.10.2011	P. Mergen	MRAC	Suggestion
	2.11.2011	Berning	LANDOOE	Comments and suggestions
2	24.11.2011	K. Zágoršek & J. Sejkora	NM	
3	09.02.2012	J. Hoffmann	MfN	
	10.02.2012	J. Holetschek	BioCASe	Comments
	13.02.2012	A. Güntsch	BGBM	Comments
	01.03.2012	G. Koch	AIT	Comments and suggestions
	06.03.2012	F. Glöckler	GBIF-D	Comments and suggestions
4	08.03.2012	A. Güntsch, P. Danes, K. Zágoršek	BGBM, NM, NM	Comments and suggestions
4a	9.3.2012	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This deliverable contains original unpublished work except where clearly indicated otherwise.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
All participants	22.10.2011	1	
Work Package Lead	24.11.2011	2	
TMG / NM	09.02.2012	3 / Paleo	
TMG / WP lead (C. Häuser, MFN)	06.03.2012	4	YES
Project Coordinator (W. Berendsohn, BGBM)	09.03.2012	4a	YES

Definition of a data schema for paleontological and mineralogical data in OpenUp!

ABCD 2.06 (Access to Biological Collection Data) is a highly complex XML data schema that can be used to represent information on natural history specimen collection and observation data and thus for a wide range of biological collections/databases – for living and preserved specimens, observations, for zoological, botanical, bacterial and viral collections, marine or terrestrial, for herbaria, botanic and zoological gardens. For each of these special types, ABCD features special sections where information specific for the particular type can be assigned. It is the data schema typically used in conjunction with the BioCASe (Biological Collection Access Service) protocol and software implementations.

ABCD-EFG 1.0[1] (Access to Biological Collection Data Extended for Geosciences) is an extension of ABCD 2.06 developed by Kießling et al. for the use with paleontological, mineralogical and geological digitalized data. Thus, the schema provides a general format for data exchange and retrieval for geo-scientific collections. ABCD-EFG is also supported by the BioCASe Provider Software, which acts as a mapping tool and web-interface for providing local data sources in XML schemas to web services and portals i.e. GeoCASE (http://www.geocase.eu/) and GBIF (http://data.gbif.org). ABCD-EFG is intended to be proposed to the organisation of biodiversity information standards TDWG (Taxonomic Databases Working Group) as a data standard for geo-scientific collections. A detailed documentation of this schema can be found at http://www.geocase.eu/documentation/efg_schema.pdf.

ABCD-EFG is the proposed data schema for paleontological and mineralogical object data in the OpenUp! project, which aims at providing standardised content from the area of paleontology and mineralogy for Europeana (see Annex I – "Description of Work", Task 4.5). There are a number of ABCD-EFG elements mandatory for paleontological datasets (Table 1).

Model for integration of paleontological content in OpenUp! (Table 2)

Multimedia content provided to the OpenUp! project from the area of paleontology is represented by specimen images with associated metadata in various formats (see Annex I – "Description of Work", Underlying content, p. 14-23). Based on the most recent transformation schema of ABCD 2.06 to ESE elements (http://open-up.eu/sites/open-up.eu/files/u16/map_ABCD206-ESE-120202-result-man-p.pdf, Status 02 February 2012) a model for the integration of relevant paleontological data has been developed (Table 2). This model includes mandatory ABCD-EFG elements corresponding to the identified mandatory ESE elements and a selection of optional ABCD-EFG elements. The selection of ABCD-EFG elements is based on the currently used elements of GeoCASE providers (MfN, ETH, GIT, ELM, TUG) and highly desired information for the public users, e.g. PartofOrganism (describing the part or parts of the organism that have been preserved, e.g. shell, skeleton, skull, soft tissue etc.). In this model relevant ABCD and EFG elements for paleontological

[1] Kiessling, W., Copp, C., Rissoné, A., Döring, M., & Mewis, H. 2006. The EFG extension to the ABCD schema. In: L. Belbin, A. Rissoné and A. Weitzman (eds), Proceedings of TDWG: Abstracts of the 2006 Annual Conference of Biodiversity Information Standards (TDWG), 15-22 October 2006, St. Louis, Missouri, U.S.A., p. 6-7.

content are emphasized. They can be either found in the EFG elements or were already present as existing ABCD elements but of less importance for biological (botany, zoology) content.

There are two ABCD 2.06 elements in use which should be avoided if the integration model is applied for paleontological content, because they are redundant and less precise in ABCD 2.06 than in EFG:

- /DataSets/DataSet/Units/Unit/Gathering/Stratigraphy/ChronostratigraphicTerms/ChronostratigraphicTerm/Term, and
- /DataSets/DataSet/Units/Unit/Gathering/Stratigraphy/LithostratigraphicTerms/LithostratigraphicTerm/Term

The suggested EFG parent elements to be used are:

Chronostratigraphy (chronostratigraphic division and name):

- /DataSets/DataSet/Units/Unit/UnitExtension/**EarthScienceSpecimen**/UnitStratigraphicDetermination/ChronostratigraphicAttributions/ChronostratigraphicAttribution/
- /DataSets/DataSet/Units/Unit/UnitExtension/**UnitHostRock**/HostRockStratigraphy/ChronostratigraphicAttributions/ChronostratigraphicAttribution
- /DataSets/DataSet/Units/Unit/UnitExtension/**AllocthonousMaterial**/OriginalStratigraphy/StratigraphicAttributions/ChronostratigraphicAttributions/ChronostratigraphicAttribution

Lithostratigraphy (formation and member):

- /DataSets/DataSet/Units/Unit/UnitExtension/**EarthScienceSpecimen**/UnitStratigraphicDetermination/LithostratigraphicAttributions/LithostratigraphicAttribution
- /DataSets/DataSet/Units/Unit/UnitExtension/**UnitHostRock**/HostRockStratigraphy/LithostratigraphicAttributions/LithostratigraphicAttribution
- /DataSets/DataSet/Units/Unit/UnitExtension/**AllocthonousMaterial**/OriginalStratigraphy/StratigraphicAttributions/LithostratigraphicAttributions/LithostratigraphicAttribution

Model for integration of mineralogical content in OpenUp! (Table 3)

The model for integration of mineralogical content was developed following the same strategy as for the paleontological data. It contains all elements relevant for mineralogical content mapped by GeoCASE providers so far (Table 3, grey boxes). Mineralogical objects are not biological objects and thus most of the ABCD 2.06 (Access to Biological Collection Data) elements do not apply. However, here the EFG elements provide very useful concepts perfectly suitable for mineralogical objects. Although there is very little experience with mineralogical content, the model includes all relevant information for the transformation of ABCD-EFG elements into ESE elements.

Outlook

In a next step, the model will be thoroughly tested with paleontological and mineralogical test datasets. The datasets are in preparation. It is planned that all partners with paleontological and

mineralogical content will have functional mappings according to the developed model for D12/MS09/C4.2.0 – *Local zoological provider software and metadata mapping functional for all content data sources* in August.