Component Report

Project Acronym: **OpenUp!**

Grant Agreement No: **270890**

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

C6.1.1 Definition of the metadata standard and controlled vocabularies

Revision: 0.3a - Final

Authors:

Wolfgang Koller (NHMW)

Heimo Rainer (NHMW)

REVISION HISTORY, STATEMENT OF ORIGINALITY AND DISTRIBUTION

Revision History

Revision	Date	Author	Organisation	Description
0.1	09/08/11	Koller	NHMW	Initial structure of component
0.2	06/09/11	Rainer / Koller	NHMW	Metadata for geographical place names
0.3	26/09/11	Koller	NHMW	Administrative metadata for common-names; Service providers & names checking for geographical place names
0.3a	22/11/11	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This deliverable contains original unpublished work except where clearly indicated otherwise.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
Coordination	06/09/2011	0.2	NO
TMG	08/09/2011	0.2	NO
TMG	18/11/2011	0.3	YES
WP Leader (same as author)	18/11/2011	0.3	YES
Project Coordinator	22/11/2011	0.3a	YES

TABLE OF CONTENTS

METADATA FOR PERSON NAMES

It was the original intention to use the DublinCore metadata standard as the metadata standard for person names. However, after reviewing the DublinCore application profile guidelines[1] it was revealed that it does not contain its own definition for person but rather incorporates the FOAF (Friend Of A Friend) standard[2]. Therefore it was decided to use the FOAF standard directly for sharing person name data.

A complete definition of the available metadata fields for a "Person" can be found at the FOAF website[3].

External resources like VIAF[4] (Virtual International Authority File) provide webservices for use as name reference systems. These webservices can be utilized in order to provide both metadata enrichment (e.g. different spellings of the same person) and quality checking (e.g. by checking a name for existence).

METADATA FOR COMMON NAMES

As a first basis for common names metadata the DarwinCore (DwC) standard was analysed[5]. However, it only contains a very lightweight definition for common names (handled as "vernacular names" in DwC) which does not provide the level of detail which is required to cover all possible types of common names and their application.

Therefore, it was decided to create a model for handling all required information for common names. That model at the same time summarizes the requirements for common names in order to derive a schema for data exchange (and harvesting). The minimum requirement remains a mapping of the scientific to the common name.

The model for common names is defined by the Entity-Relationship-Diagram displayed in Figure 1.

The actual implementation of the data model (as a schema) will be part of setting up the harvesting infrastructure for common names (which is due in M18).

The implementation will also contain additional administrative metadata in order to be able to quote the original source of the information. This will also be addressed during the setup of the harvesting infrastructure.

---

[1] http://dublincore.org/documents/profile-guidelines/

[2] http://xmlns.com/foaf/spec/

[3] http://xmlns.com/foaf/spec/#term_Person

[4] http://viaf.org/

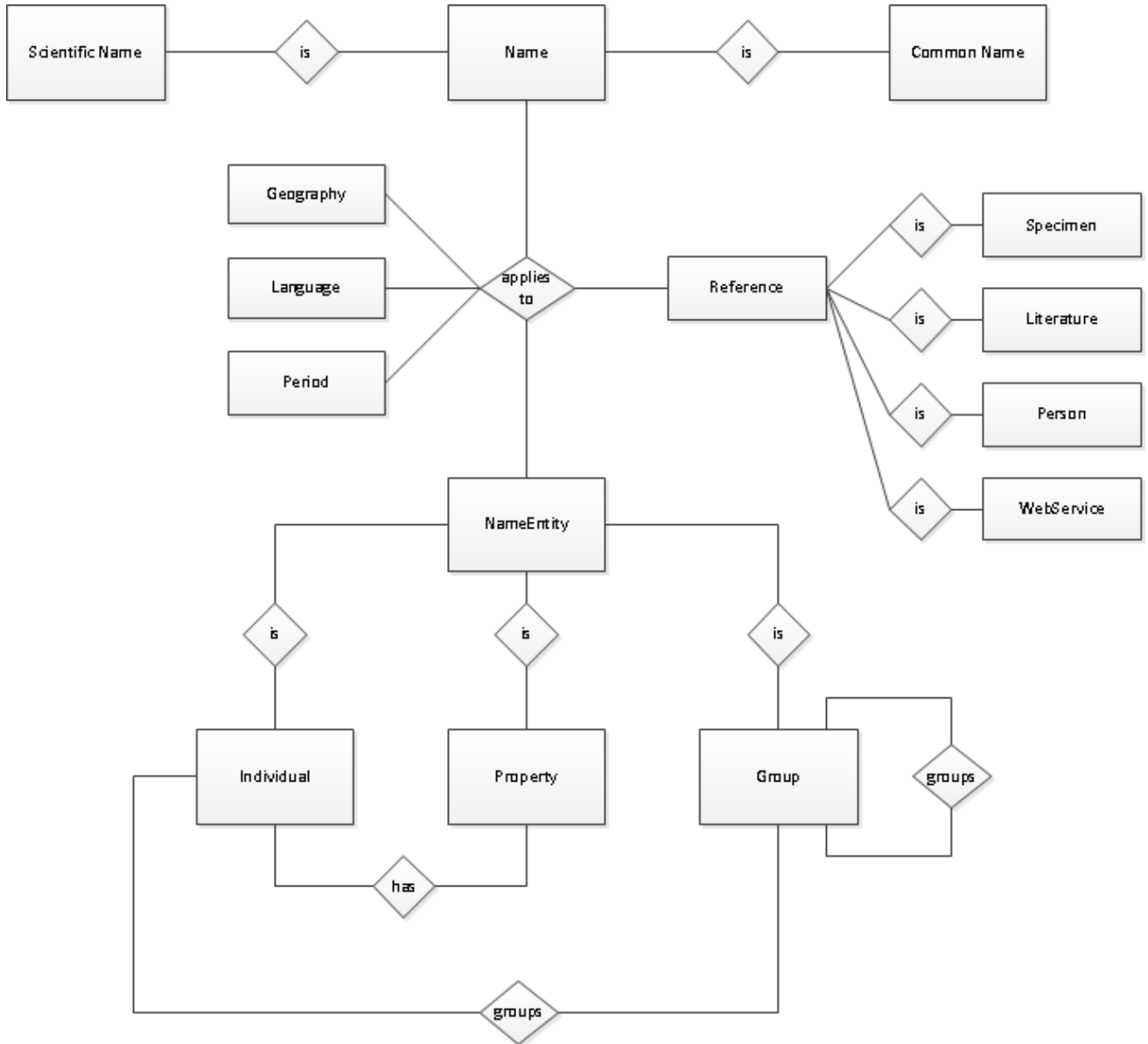[5] http://vocabularies.gbif.org/node/127063

Figure 1: Common names data model

METADATA FOR GEOGRAPHICAL PLACE NAMES

As geographical place names are sometimes ambiguous it will not be possible to clearly identify a name as a certain place. The amount of false positives will be too high for an automated process. Therefore it should be considered to exclude geographical place names from an automated service architecture and let content providers directly fill in the information (as accurate as possible). Geographical names checking can be done using existing service providers (see below for details).

Metadata format

As a metadata format for geographical information several standards were evaluated. First the standards provided by the Open Geospatial Consortium (OGC)[6] have been evaluated. However those standards provide way more information than required for place names in the case of OpenUp!.

A quite simple but yet robust implementation for geographical place names is provided by the GeoNames[7] ontology definition[8]. It provides a generic way to exchange place name information and is provided via GeoNames webservices as well.

It utilizes latest semantic web technologies (rdf-s) to model geographical place names and relations between them.

Service providers

As mentioned above GeoNames provides several webservices which can be used for checking and identifying geographical place names within OpenUp!. The GeoNames project offers several widely used webservices for access to their database. These services even provide fuzzy searching capabilities which would allow a more satisfying names checking mechanism.

Therefore, it is not suggested to maintain an own infrastructure for geographical place names but instead leverage the GeoNames services for use within the OpenUp! architecture.

---

[6] http://www.opengeospatial.org/

[7] http://www.geonames.org/

[8] http://www.geonames.org/ontology/documentation.html