



INSPIRE

Infrastructure for Spatial Information in Europe

Questionnaire on Unique Identifiers – Report

Title	Questionnaire on Unique Identifiers – Report
Creator	INSPIRE DT Data Specifications
Date	08.02.2008
Subject	Report summarising the results of the questionnaire
Status	Final
Publisher	INSPIRE DT Data Specifications
Type	Text
Description	Results, statistics and conclusion of the questionnaire – report on the current status of using unique identifiers in organizations, and to clarify if the proposed structure and restrictions in the lexical rules for unique identifiers within INSPIRE are feasible.
Contributor	Dominique Laurent, Clemens Portele, members of the Drafting Team Data Specifications, Darja Lihteneger
Format	Adobe PDF
Source	
Rights	
Identifier	INSPIRE-UID-report
Language	En
Relation	INSPIRE Data Specifications
Coverage	

These are Dublin Core metadata elements. See for more details and examples <http://www.dublincore.org/>.

Acronyms / Glossary

Used Term	Definition
EEA	European Environmental Agency
EU	European Union
INSPIRE	Infrastructure of Spatial Information in the European Community
LMO	Legally Mandated Organisation
SDI	Spatial Data Infrastructure
SDIC	Spatial Data Interest Community
UID questionnaire	Questionnaire on Unique Identifiers

Table of contents

Acronyms / Glossary	2
Purpose of this document.....	4
1. Basic statistics	5
2. Statistics on the results of the questionnaire.....	6
3. Conclusion.....	7

Purpose of this document

Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), was published in the Official Journal of the European Union on 25 April 2007 and it entered into force on 15 May 2007. INSPIRE lays down the legal framework for the establishment and operation of an Infrastructure for Spatial Information in Europe. The purpose of such an infrastructure is to assist policy-making in relation to policies that may have a direct or indirect impact on the environment.

Implementing Rules are being developed for technical arrangements for the interoperability and harmonisation of spatial data sets and services based on requirements from policies and based on the spatial data infrastructures in the Member States.

To support the development of the Implementing Rules, a questionnaire on the use of unique identifiers was published in May 2007 to solicit additional information on current practice beyond the information that could be extracted from the reference material submitted by stakeholders.

Article 8 (2) of the INSPIRE Directive requires that for datasets corresponding to the spatial data themes listed in the annexes I and II, the implementing rules shall address "a common framework for the unique identification of spatial objects, to which identifiers under national systems can be mapped in order to ensure interoperability between them."

The draft of the Generic Conceptual Model for INSPIRE data specifications contains a proposal for the approach regarding unique identifiers for spatial objects. The questionnaire was published in particular to support future decisions about the rules and lexical restrictions for unique identifiers in INSPIRE.

The technical rationale for a potential restriction of the lexical space (maximum length and/or allowed characters) of unique identifiers is that in implementations identifiers play an important role in referencing or retrieving items labeled with an identifier and, therefore, identifiers are used in indexing mechanisms in information systems of both the data provider and the data consumer. Since implementation platforms sometimes place restrictions on the lexical space for such keys, this has an impact on the implementation of INSPIRE. An example is the Web Feature Service (a candidate for the download service) that uses XML IDs to encode identifiers. As a consequence, identifiers are not constrained in length, but in the characters they can use. For example, they may not start with a digit and contain only a restricted set of allowed characters. Another related reason for considering lexical constraints is a potential performance impact on the indexing mechanism, if no constraints are specified.

This document informs about the results of the questionnaire about unique identifiers for spatial objects and the conclusions for the Implementing Rules on the interoperability of spatial data sets.

1. Basic statistics

Questionnaire on Unique Identifiers were published on the INSPIRE web site (<http://www.ec-gis.org/inspire/>) and it was available for registered SDICs and LMOs and others from the geo-information community interested in providing information. The answers were collected in a database.

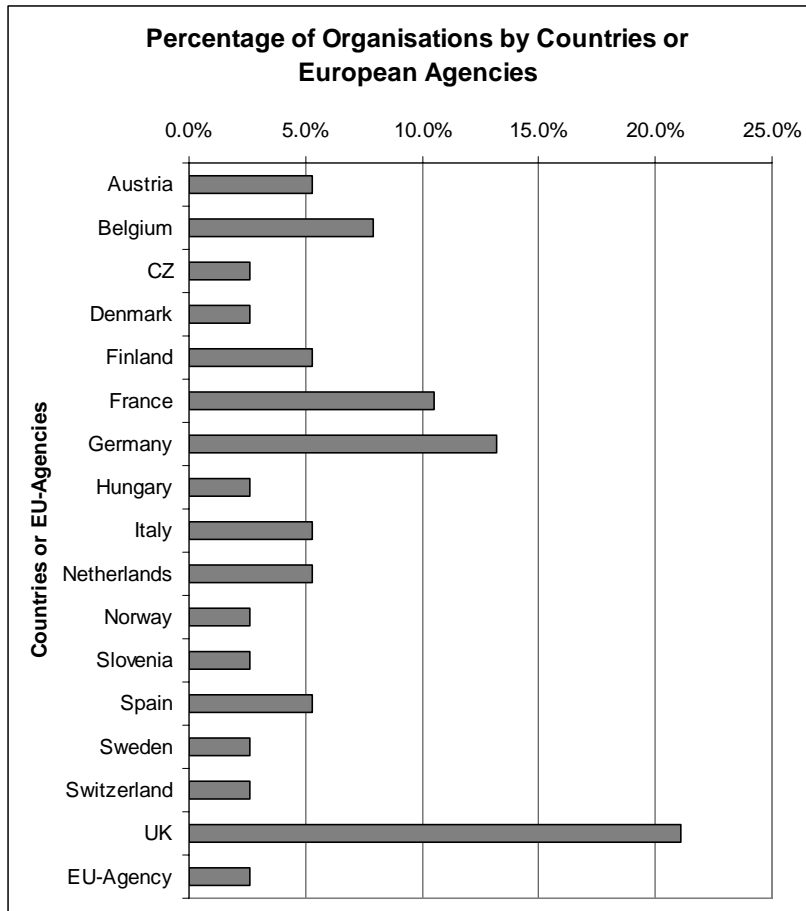
Basic statistics shows that 38 organizations submitted to the UID questionnaire. Based on the derived statistics, the organizations from 14 European Union countries, from Norway and Switzerland, and the European Environmental Agency supplied the answers to the questionnaire.

Detailed information is provided in Table 1 and Graph 1.

Country or Organization	Number of organizations	%
Austria	2	5.3%
Belgium	3	7.9%
CZ	1	2.6%
Denmark	1	2.6%
Finland	2	5.3%
France	4	10.5%
Germany	5	13.2%
Hungary	1	2.6%
Italy	2	5.3%
Netherlands	2	5.3%
Norway	1	2.6%
Slovenia	1	2.6%
Spain	2	5.3%
Sweden	1	2.6%
Switzerland	1	2.6%
UK	8	21.1%
EU-Agency (EEA)	1	2.6%
	38	100.0%

* EEA – European Environmental Agency

Table 1: Participation of Organizations in the UID Questionnaire, by Countries and by European Agencies



Graph 1: Percentage of Organizations, by Countries and by European Agencies, Participating in the UID Questionnaire

2. Statistics on the results of the questionnaire

As the INSPIRE Directive requires unique identifiers for spatial objects in themes of Annexes I and II, first question was about the use of unique identifiers for these themes.

Theme	Geographical names	Administrative units	Transport	Hydrography	Protected sites
Data with identifier	19	21	13	14	11
Data without identifier	9	8	6	7	5
No data	2	2	10	10	16

Theme	Elevation	Cadastral parcels	Addresses	Land cover	Geology
Data with identifier	9	13	14	14	3
Data without identifier	11	3	7	9	1
No data	8	14	10	8	23

There are generally more providers having data with identifiers than producers having data without identifiers, which is a positive point. Nevertheless, some data providers will have to initiate identifier management for INSPIRE.

3. Conclusion

The questionnaire aimed to test the feasibility of lexical rules proposed by DT DS about local identifiers for vector data:

- use of the following set of characters {"A"..."Z", "a"..."z", "0"..."9", "_", ".", "-", " "}
- the length of the local identifier is limited to 24 characters for vector data
- the length of the local identifier is limited to 128 characters for coverage data
- the length of the namespace is limited to 24 characters

The answers have shown that restriction of length was not feasible (even with a higher threshold) for some data providers. So, this restriction won't be adopted in the Generic Conceptual Model.

The answers have shown that the restriction to the proposed set of characters was considered as feasible by the data providers, either because they already use it (or a subset of it) or because they consider as possible to map their existing character set to the one proposed by the questionnaire. So, this restriction will be adopted in the Generic Conceptual Model.

The questionnaire also aimed to investigate if and how data providers identify different versions of a spatial object. The answers have shown a large range of cases: no identification of the different versions, use of start date/end date, use of versions (identified by number or letter), use of time stamp, historical table of identifiers, ... This diversity of answers confirm the relevance of requirements and recommendations given in the Generic Conceptual Model:

- allow publication of versioning but do not make it mandatory
- if versioning is provided, timestamps are recommended but not required.