



INSPIRE

Infrastructure for Spatial Information in Europe

D2.8.III.10 Data Specification on Population Distribution - Demography – Draft Guidelines

Title	D2.8.III.10 INSPIRE Data Specification on <i>Population</i> Distribution - Demography – Draft Guidelines
Creator	INSPIRE Thematic Working Group <i>Population</i> Distribution - Demography
Date	2011-07-13
Subject	INSPIRE Data Specification for the spatial data theme <i>Population</i> Distribution - Demography
Publisher	INSPIRE Thematic Working Group <i>Population</i> Distribution - Demography
Type	Text
Description	This document describes the INSPIRE Data Specification for the spatial data theme <i>Population</i> Distribution - Demography
Contributor	Members of the INSPIRE Thematic Working Group <i>Population</i> Distribution - Demography
Format	Portable Document Format (pdf)
Source	
Rights	Public
Identifier	D2.8.III.10_v2.0.1
Language	En
Relation	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)
Coverage	Project duration

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page II

Foreword

How to read the document?

This document describes the “INSPIRE data specification on *Population Distribution - Demography – Guidelines*” version 2.0.1 as developed by the Thematic Working Group (TWG) **Error! Unknown document property name.** using both natural and a conceptual schema language. This version is now available for the public consultation. Based on the results of the consultation (received comments and the testing reports), the final version 3.0 will be prepared by the TWGs.

The data specification is based on a common template used for all data specifications and has been harmonised using the experience from the development of the Annex I data specifications.

This document provides guidelines for the implementation of the provisions laid down in the draft Implementing Rule for spatial data sets and services of the INSPIRE Directive.

This document includes two executive summaries that provide a quick overview of the INSPIRE data specification process in general, and the content of the data specification on *Population Distribution - Demography* in particular. We highly recommend that managers, decision makers, and all those new to the INSPIRE process and/or information modelling should read these executive summaries first.

The UML diagrams (in Chapter 5) offer a rapid way to see the main elements of the specifications and their relationships. The definition of the spatial object types, attributes, and relationships are included in the Feature Catalogue (also in Chapter 5). People having thematic expertise but not familiar with UML can fully understand the content of the data model focusing on the Feature Catalogue. Users might also find the Feature Catalogue especially useful to check if it contains the data necessary for the applications that they run. The technical details are expected to be of prime interest to those organisations that are/will be responsible for implementing INSPIRE within the field of *Population Distribution - Demography*.

The technical provisions and the underlying concepts are often illustrated by examples. Smaller examples are within the text of the specification, while longer explanatory examples and descriptions of selected use cases are attached in the annexes.

In order to distinguish the INSPIRE spatial data themes from the spatial object types, the INSPIRE spatial data themes are written in *italics*.

The document will be publicly available as a ‘non-paper’. It does not represent an official position of the European Commission, and as such cannot be invoked in the context of legal procedures.

Legal Notice

Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page III

Interoperability of Spatial Data Sets and Services – General Executive Summary

The challenges regarding the lack of availability, quality, organisation, accessibility, and sharing of spatial information are common to a large number of policies and activities and are experienced across the various levels of public authority in Europe. In order to solve these problems it is necessary to take measures of coordination between the users and providers of spatial information. The Directive 2007/2/EC of the European Parliament and of the Council adopted on 14 March 2007 aims at establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) for environmental policies, or policies and activities that have an impact on the environment.

INSPIRE will be based on the infrastructures for spatial information that are created and maintained by the Member States. To support the establishment of a European infrastructure, Implementing Rules addressing the following components of the infrastructure are being specified: metadata, interoperability of spatial data themes (as described in Annexes I, II, III of the Directive) and spatial data services, network services and technologies, data and service sharing, and monitoring and reporting procedures.

INSPIRE does not require collection of new data. However, after the period specified in the Directive¹ Member States have to make their data available according to the Implementing Rules.

Interoperability in INSPIRE means the possibility to combine spatial data and services from different sources across the European Community in a consistent way without involving specific efforts of humans or machines. It is important to note that “interoperability” is understood as providing access to spatial data sets through network services, typically via Internet. Interoperability may be achieved by either changing (harmonising) and storing existing data sets or transforming them via services for publication in the INSPIRE infrastructure. It is expected that users will spend less time and efforts on understanding and integrating data when they build their applications based on data delivered within INSPIRE.

In order to benefit from the endeavours of international standardisation bodies and organisations established under international law their standards and technical means have been utilised and referenced, whenever possible.

To facilitate the implementation of INSPIRE, it is important that all stakeholders have the opportunity to participate in specification and development. For this reason, the Commission has put in place a consensus building process involving data users, and providers together with representatives of industry, research and government. These stakeholders, organised through Spatial Data Interest Communities (SDIC) and Legally Mandated Organisations (LMO)², have provided reference materials, participated in the user requirement and technical³ surveys, proposed experts for the Data Specification Drafting Team⁴ and Thematic Working Groups⁵.

¹ For all 34 Annex I,II and III data themes: within two years of the adoption of the corresponding Implementing Rules for newly collected and extensively restructured data and within 5 years for other data in electronic format still in use

² Number of SDICs and LMOs on 8/6/2011 was 461 and 249 respectively

³ Surveys on unique identifiers and usage of the elements of the spatial and temporal schema,

⁴ The Data Specification Drafting Team has been composed of experts from Austria, Belgium, Czech Republic, France, Germany, Greece, Italy, Netherlands, Norway, Poland, Switzerland, UK, and the European Environmental Agency

⁵ The Thematic Working Groups of Annex II and III themes have been composed of experts from Austria, Belgium, Bulgaria, Czech Republic, Denmark, Finland, France, Germany, Hungary, Ireland, Italy, Latvia, Netherlands, Norway, Poland, Romania, Slovakia, Spain, Sweden, Switzerland, Turkey, UK, the European Commission, and the European Environmental Agency

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page IV

This open and participatory approach was successfully used during the development of the data specification on Annex I data themes as well as during the preparation of the Implementing Rule on Interoperability of Spatial Data Sets and Services⁶ for Annex I spatial data themes.,

The development framework elaborated by the Data Specification Drafting Team aims at keeping the data specifications of the different themes coherent. It summarises the methodology to be used for the data specifications and provides a coherent set of requirements and recommendations to achieve interoperability. The pillars of the framework are four technical documents:

- The Definition of Annex Themes and Scope⁷ describes in greater detail the spatial data themes defined in the Directive, and thus provides a sound starting point for the thematic aspects of the data specification development.
- The Generic Conceptual Model⁸ defines the elements necessary for interoperability and data harmonisation including cross-theme issues. It specifies requirements and recommendations with regard to data specification elements of common use, like the spatial and temporal schema, unique identifier management, object referencing, a generic network model, some common code lists, etc. Those requirements of the Generic Conceptual Model that are directly implementable will be included in the Implementing Rule on Interoperability of Spatial Data Sets and Services.
- The Methodology for the Development of Data Specifications⁹ defines a repeatable methodology. It describes how to arrive from user requirements to a data specification through a number of steps including use-case development, initial specification development and analysis of analogies and gaps for further specification refinement.
- The “Guidelines for the Encoding of Spatial Data”¹⁰ defines how geographic information can be encoded to enable transfer processes between the systems of the data providers in the Member States. Even though it does not specify a mandatory encoding rule it sets GML (ISO 19136) as the default encoding for INSPIRE.

Based on these framework documents and following the successful development of the Annex I Data specifications (Technical Guidelines) and the Implementing Rules, the new Thematic Working Groups have created the INSPIRE data specification for each Annex II and III theme. These documents – at the version 2.0.1 – are now publicly available for INSPIRE stakeholders for consultation. The consultation phase covers expert review as well as feasibility and fitness-for-purpose testing of the data specifications.

The structure of the data specifications is based on the “ISO 19131 Geographic information - Data product specifications” standard. They include the technical documentation of the application schema, the spatial object types with their properties, and other specifics of the spatial data themes using natural language as well as a formal conceptual schema language¹¹.

A consolidated model repository, feature concept dictionary, and glossary are being maintained to support the consistent specification development and potential further reuse of specification elements. The consolidated model consists of the harmonised models of the relevant standards from the ISO

⁶ Commission Regulation (EU) No 1089/2010 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards interoperability of spatial data sets and services, published in the Official Journal of the European Union on 8th of December 2010.

⁷ http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.3_Definition_of_Annex_Themes_and_scope_v3.0.pdf

⁸ http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.5_v3.3.pdf

⁹ http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.6_v3.0.pdf

¹⁰ http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.7_v3.2.pdf

¹¹ UML – Unified Modelling Language

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page V

19100 series, the INSPIRE Generic Conceptual Model, and the application schemas¹² developed for each spatial data theme. The multilingual INSPIRE Feature Concept Dictionary contains the definition and description of the INSPIRE themes together with the definition of the spatial object types present in the specification. The INSPIRE Glossary defines all the terms (beyond the spatial object types) necessary for understanding the INSPIRE documentation including the terminology of other components (metadata, network services, data sharing, and monitoring).

By listing a number of requirements and making the necessary recommendations, the data specifications enable full system interoperability across the Member States, within the scope of the application areas targeted by the Directive. They will be published (version 3.0) as technical guidelines and will provide the basis for the content of the Amendment of the Implementing Rule on Interoperability of Spatial Data Sets and Services for data themes included in Annex II and III of the Directive. The Implementing Rule Amendment will be extracted from the data specifications keeping in mind short and medium term feasibility as well as cost-benefit considerations. The Implementing Rule will be legally binding for the Member States.

In addition to providing a basis for the interoperability of spatial data in INSPIRE, the data specification development framework and the thematic data specifications can be reused in other environments at local, regional, national and global level contributing to improvements in the coherence and interoperability of data in spatial data infrastructures.

¹² Conceptual models related to specific areas (e.g. INSPIRE themes)

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page VI

Population Distribution - Demography – Executive Summary

Population distributions are defined as “datasets of statistical information describing how some phenomenon is spread within some part of the 2D space”. This document presents data specifications for population distributions. It is based on the following principles:

Separation between statistical data and statistical units

A statistical data is defined as “any numerical representation of a phenomenon”. A statistical unit informs on the statistical data location. This document addresses only the statistical data specification: The specification of statistical unit is out of the scope and statistical data refers to statistical units through their identifier. Linking statistical data to spatial object is not only linking a database object to another one. The challenge is to improve the interaction between the statistics and the GIS worlds. This document recommendations aims at improving the integration of spatial and statistic analyses.

Genericity

There are many different kinds of statistical data about human population: about people, dwellings, people at their work place, etc. This document does not intend to provide specifications for all these. Common characteristics have been extracted and represented into a generic data model.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page VII

Acknowledgements

Many individuals and organisations have contributed to the development of these Guidelines.

The Thematic Working Group on Statistical Units and Population Distribution (TWG-SU-PD) included: Alina Kmiecik (TWG Facilitator until 15/12/2010) and Udo Maack (TWG Facilitator from 15/12/2010), Pieter Bresters, Ian Coady, Mary Haldorson, Jean-Luc Lipatz (TWG Editor), Miroslaw Migacz, Susanne Schnorr-Baecker, Julien Gaffuri (European Commission contact point).

Contributors:

From Eurostat: Beatrice Eiselt, Ekkehard Petri and Daniele Rizzi
 From CBS (NL): Niek van Leeuwen, Olav ten Bosch

Participants on GISCO / Geostat meetings:

Hugo Poelmann (DG Regio) – Requirements for Change Information
 Ingrid Kaminger (Statistic Austria) – User Survey according Grids
 Marjan van Herwijnen (ESPO-Coordination Unit) - Requirements for Change Information

Other contributors to the INSPIRE data specifications are the Drafting Team Data Specifications, the JRC data specifications team and the INSPIRE stakeholders - Spatial Data Interested Communities (SDICs) or Legally Mandated Organisations (LMOs).

Contact information

Vanda Nunes de Lima
 European Commission Joint Research Centre
 Institute for Environment and Sustainability
 Spatial Data Infrastructures Unit
 TP262, Via Fermi 2749
 I-21027 Ispra (VA)
 ITALY
 E-mail: vanda.lima@jrc.ec.europa.eu
 Tel.: +39-0332-7865052
 Fax: +39-0332-7866325
<http://ies.jrc.ec.europa.eu/>
<http://ec.europa.eu/dgs/jrc/>
<http://inspire.jrc.ec.europa.eu/>

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page VIII

Table of contents

1	Scope	1
2	Overview	1
2.1	Name	1
2.2	Informal description	1
2.3	Normative References	2
2.4	Terms and definitions	3
2.5	Symbols and abbreviations	3
2.6	Notation of requirements and recommendations	3
2.7	Conformance	4
3	Specification scopes	4
4	Identification information	4
5	Data content and structure	4
5.1	Basic notions	5
5.1.1	Stereotypes	5
5.1.2	Placeholder and candidate types	5
5.1.3	Voidable characteristics	6
5.1.4	Code lists and Enumerations	7
5.2	Application schema Population Distribution - Demography	7
5.2.1	Description	7
5.2.2	Feature catalogue	14
6	Reference systems	22
6.1	Coordinate reference systems	22
6.1.1	Datum	22
6.1.2	Coordinate reference systems	22
6.1.3	Display	23
6.1.4	Identifiers for coordinate reference systems	23
6.2	Temporal reference system	23
6.3	Theme-specific requirements and recommendations on reference systems	24
7	Data quality	24
7.1	Data quality elements and measures	24
7.1.1	Completeness – Commission	25
7.1.2	Completeness – Omission	25
7.1.3	Completeness – Population not counted	26
7.1.4	Rough location – Conventionally located population	26
7.1.5	Rough location – Approximately located population	27
7.2	Minimum data quality requirements and recommendations	27
8	Dataset-level metadata	28
8.1	Common metadata elements	28
8.1.1	Coordinate Reference System	30
8.1.2	Temporal Reference System	30
8.1.3	Encoding	31
8.1.4	Character Encoding	32
8.1.5	Data Quality – Logical Consistency – Topological Consistency	32
8.2	Metadata elements for reporting data quality	33
8.3	Theme-specific metadata elements	34
8.3.1	Maintenance Information	35
8.3.2	Data Quality – Completeness – Commission	35

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page IX

8.3.3	Data Quality – Completeness – Omission	36
8.4	Guidelines on using metadata elements defined in Regulation 1205/2008/EC	36
8.4.1	Conformity.....	36
8.4.2	Lineage	36
8.4.3	Temporal reference	37
9	Delivery	37
9.1	Delivery medium	37
9.2	Encodings	38
9.2.1	Default Encoding(s)	38
9.2.2	Alternative Encoding(s).....	38
10	Data Capture.....	38
11	Portrayal	39
	Bibliography.....	42
	Annex B (normative) Abstract Test Suite	43
	Annex A (informative) Use cases	44

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 1

1 Scope

This document specifies a harmonised data specification for the spatial data theme *Population Distribution - Demography* as defined in Annex III of the INSPIRE Directive.

This data specification provides the basis for the drafting of Implementing Rules according to Article 7 (1) of the INSPIRE Directive [Directive 2007/2/EC]. The entire data specification will be published as implementation guidelines accompanying these Implementing Rules.

2 Overview

2.1 Name

INSPIRE data specification for the theme Population Distribution - Demography.

2.2 Informal description

Definition:

Geographical distribution of people, including population characteristics and activity levels, aggregated by grid, region, administrative unit or other analytical unit [Directive 2007/2/EC].

Description:

Population = human population

A population is any entire collection of people, animals, plants or things about which we may collect data. It is the entire group we are interested in, which we wish to describe or draw conclusions about. In this data specification population is referring to human beings, i.e. (1) people such as individuals or (2) people living together in households for instance. Information on people where and how they live or stay are crucial for nearly all other themes of the INSPIRE Directive, in particular to Human Health and Safety, Buildings and Natural Risk Zones.

The theme may thematically be divided into several components. The directive text points at broad groups of sub-themes:

- Human population by individual characteristics (sex, age, marital status, nationality).
- Human population activity levels (education, profession)
- Human beings living together in groups for different reasons (households, institutions such as retirement homes)

A wide range of statistics on population are available

Concerning population information, this may include total population, age, gender, nationality or place of birth in terms of stocks, or mortality, life expectancy and migration in terms of flows. Of interest could be the population with reference to a time period and activities. Use cases point to the need for a distinction between night time and daytime population for a given regional unit such as a city or a sub-city. It might be difficult is to determine how many people are in a specific place (settlement, apartment or house) and to which groups they belong. According to people's activities, they can be at work or enjoy leisure time activities (such as watching a movie in the cinema) or they can spend their time in transport facilities. For people at work, it can be shown to which economic sector in terms of NACE they belong in a specific place, region or part of it or where people of a given socio

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 2

demographic group live for instance in towns or at the border. All such information could be useful in the context of other thematic themes such as agriculture or the usage of natural resources.

No spatial features

The theme Population distribution/demography contains attributes mainly related to administrative units or statistical units. This means that this theme has no direct spatial features such as many other INSPIRE Annex III themes such as buildings, health care institutions etc., it needs to be linked to these features by the use of for example NUTS-codes or grids.

Aggregated data and confidentiality issues

The definition in the Directive 2007/2/EC specifies all kinds of features relevant to demography. It includes the term "aggregated" which means that non-aggregated data about population is excluded because of confidentiality. The data provision on a level lower than municipalities (LAU2) depends on the national situation regarding availability and privacy. There is a correlation between size of statistical units and differentiation of the attribute values. The smaller the size of the statistical units, the more undifferentiated the attributes can be to meet the data privacy rules. The data model covers all the possibilities but legal restrictions may prohibit the provision.

Open issue 1: The current specification only provides a generic model for disseminating statistical data spatial distributions. That means that the specification doesn't specify which aspects of human characteristics will be described in the datasets. Nevertheless it must be stressed that **the support for the dissemination of at least population distribution counts using a harmonized definition of spatial units (i.e. using grids with a standardized grid cell size) is an important issue and should be a requirement.** At the time this specification was laid down, the place where to put such a requirement is still an open question.

2.3 Normative References

[Directive 2007/2/EC] 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)

REGULATION (EC) No 177/2008 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 20 February 2008 establishing a common framework for business registers for statistical purposes and repealing Council Regulation (EEC) No 2186/93

REGULATION (EC) No 1893/2006 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 20 December 2006 establishing the statistical classification of economic activities NACE Revision 2 and amending Council Regulation (EEC) No 3037/90 as well as certain EC Regulations on specific statistical domains

REGULATION (EC) No 763/2008 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 9 July 2008 on population and housing censuses

COMMISSION REGULATION (EC) No 1201/2009 of 30 November 2009 implementing Regulation (EC) No 763/2008 as regards the technical specifications of the topics and of their breakdowns

COMMISSION REGULATION (EU) No 519/2010 of 16 June 2010 adopting the programme of the statistical data and of the metadata for population and housing censuses provided for by Regulation (EC) No 763/2008

REGULATION (EU) No 1151/2010 of 08 December 2010 implementing Regulation (EC) No 763/2008, as regards the modalities and structure of the quality reports and the technical format for data transmission

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 3

REGULATION (EC) No 1177/2003 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 16 June 2003 concerning Community statistics on income and living conditions (EU-SILC)

REGULATION (EC) No 295/2008 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 11 March 2008 concerning structural business statistics

COUNCIL REGULATION (EC) No 577/98 of 9 March 1998 COUNCIL REGULATION (EC) No 577/98 of 9 March 1998 on the organisation of a labour force sample survey in the Community

[ISO 19107] EN ISO 19107:2005, Geographic Information – Spatial Schema

[ISO 19108-c] ISO 19108:2002/Cor 1:2006, Geographic Information – Temporal Schema, Technical Corrigendum 1

[ISO 19113] EN ISO 19113:2005, Geographic Information – Quality principles

[ISO 19115] EN ISO 19115:2005, Geographic information – Metadata (ISO 19115:2003)

[ISO 19138] ISO/TS 19138:2006, Geographic Information – Data quality measures

[Regulation 1205/2008/EC] Regulation 1205/2008/EC implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata

2.4 Terms and definitions

General terms and definitions helpful for understanding the INSPIRE data specification documents are defined in the INSPIRE Glossary¹³.

2.5 Symbols and abbreviations

Eurostat Statistical Office of the European Communities
 NACE Statistical classification of economic activities in the European Community
 NUTS Nomenclature of Territorial Units for Statistics
 UML Unified Modelling Language

2.6 Notation of requirements and recommendations

To make it easier to identify the mandatory requirements and the recommendations for spatial data sets in the text, they are highlighted and numbered.

IR Requirement X Requirements that are reflected in the Implementing Rule on interoperability of spatial data sets and services are shown using this style.

DS Requirement X Requirements that are not reflected in the Implementing Rule on interoperability of spatial data sets and services are shown using this style.

¹³ The INSPIRE Glossary is available from <http://inspire-registry.jrc.ec.europa.eu/registers/GLOSSARY>

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 4

Recommendation 1 Recommendations are shown using this style.

2.7 Conformance

DS Requirement 1 Any dataset claiming conformance with this INSPIRE data specification shall pass the requirements described in the abstract test suite presented in Annex A.

3 Specification scopes

This data specification does not distinguish different specification scopes, but just considers one general scope.

NOTE For more information on specification scopes, see [ISO 19131:2007], clause 8 and Annex D.

4 Identification information

NOTE Since the content of this chapter was redundant with the overview description (section 2) and executive summary, it has been decided that this chapter will be removed in v3.0.

5 Data content and structure

IR Requirement 1 Spatial data sets related to the theme *Population Distribution - Demography* shall be provided using the spatial object types and data types specified in the application **schema(s)** in this section.

IR Requirement 2 Each spatial object shall comply with all constraints specified for its spatial object type or data types used in values of its properties, respectively.

Recommendation 1 The reason for a void value should be provided where possible using a listed value from the VoidValueReason code list to indicate the reason for the missing value.

NOTE The application schema specifies requirements on the properties of each spatial object including its multiplicity, domain of valid values, constraints, etc. All properties have to be reported, if the relevant information is part of the data set. Most properties may be reported as “void”, if the data set does not include relevant information. See the Generic Conceptual Model [INSPIRE DS-D2.5] for more details.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 5

5.1 Basic notions

This section explains some of the basic notions used in the INSPIRE application schemas. These explanations are based on the GCM [DS-D2.5].

5.1.1 Stereotypes

In the application schemas in this sections several stereotypes are used that have been defined as part of a UML profile for use in INSPIRE [INSPIRE DS-D2.5]. These are explained in Table 1 below.

Table 1 – Stereotypes (adapted from [INSPIRE DS-D2.5])

Stereotype	Model element	Description
applicationSchema	Package	An INSPIRE application schema according to ISO 19109 and the Generic Conceptual Model.
featureType	Class	A spatial object type.
type	Class	A conceptual, abstract type that is not a spatial object type.
dataType	Class	A structured data type without identity.
union	Class	A structured data type without identity where exactly one of the properties of the type is present in any instance.
enumeration	Class	A fixed list of valid identifiers of named literal values. Attributes of an enumerated type may only take values from this list.
codeList	Class	A flexible enumeration that uses string values for expressing a list of potential values.
placeholder	Class	A placeholder class (see definition in section 5.1.2).
voidable	Attribute, association role	A voidable attribute or association role (see definition in section 5.1.3).
lifeCycleInfo	Attribute, association role	If in an application schema a property is considered to be part of the life-cycle information of a spatial object type, the property shall receive this stereotype.
version	Association role	If in an application schema an association role ends at a spatial object type, this stereotype denotes that the value of the property is meant to be a specific version of the spatial object, not the spatial object in general.

5.1.2 Placeholder and candidate types

Some of the INSPIRE Annex I data specifications (which were developed previously to the current Annex II+III data specifications) refer to types that thematically belong and were expected to be fully specified in Annex II or III spatial data themes. Two kinds of such types were distinguished:

- *Placeholder types* were created as placeholders for types (typically spatial object types) that were to be specified as part of a future spatial data theme, but which was already used as a value type of an attribute or association role in this data specification.

Placeholder types received the stereotype «placeholder» and were placed in the application schema package of the future spatial data theme where they thematically belong. For each placeholder, a definition was specified based on the requirements of the Annex I theme. The Annex II+III TWGs were required to take into account these definitions in the specification work of the Annex II or III theme.

If necessary, the attributes or association roles in the Annex I data specification(s) that have a placeholder as a value type shall be updated if necessary.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 6

- *Candidate types* were types (typically spatial object types) for which already a preliminary specification was given in the Annex I data specification. Candidate types did not receive a specific stereotype and were placed in the application schema package of the future spatial data theme where they thematically belong. For each candidate type, a definition and attributes and association roles were specified based on the requirements of the Annex I theme. The Annex II+III TWGs were required to take into account these specifications in the specification work of the Annex II or III theme.

If the type could not be incorporated in the Annex II or III data specification according to its preliminary specification, it should be moved into the application schema of the Annex I theme where it had first been specified. In this case, the attributes or association roles in the Annex I data specification(s) that have the type as a value type shall be updated if necessary.

Open issue 2: For all Annex II+III themes for which placeholders and candidate types were specified in an Annex I data specification, it should be clearly indicated in the data specification, how the placeholder and candidate types were taken into account. If the proposed solution would require any changes to an Annex I data specification (and the corresponding section in the IR for interoperability of spatial data sets and services), this should also be clearly indicated.

A thorough investigation of the implications of the proposed changes of candidate types (in particular related to requirements of Annex I maintenance) will have to be performed for v3.0 of the data specifications.

5.1.3 Voidable characteristics

If a characteristic of a spatial object is not present in the spatial data set, but may be present or applicable in the real world, the property shall receive this stereotype.

If and only if a property receives this stereotype, the value of *void* may be used as a value of the property. A *void* value shall imply that no corresponding value is contained in the spatial data set maintained by the data provider or no corresponding value can be derived from existing values at reasonable costs, even though the characteristic may be present or applicable in the real world.

It is possible to qualify a value of void in the data with a reason using the *VoidValueReason* type. The *VoidValueReason* type is a code list, which includes the following pre-defined values:

- *Unpopulated*: The characteristic is not part of the dataset maintained by the data provider. However, the characteristic may exist in the real world. For example when the “elevation of the water body above the sea level” has not been included in a dataset containing lake spatial objects, then the reason for a void value of this property would be ‘Unpopulated’. The characteristic receives this value for all objects in the spatial data set.
- *Unknown*: The correct value for the specific spatial object is not known to, and not computable by the data provider. However, a correct value may exist. For example when the “elevation of the water body above the sea level” of a *certain lake* has not been measured, then the reason for a void value of this property would be ‘Unknown’. This value is applied on an object-by-object basis in a spatial data set.

NOTE It is expected that additional reasons will be identified in the future, in particular to support reasons / special values in coverage ranges.

The «voidable» stereotype does not give any information on whether or not a characteristic exists in the real world. This is expressed using the multiplicity:

- If a characteristic may or may not exist in the real world, its minimum cardinality shall be defined as 0. For example, an if an Address may or may not have a house number, the multiplicity of the corresponding property shall be 0..1.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 7

- If at least one value for a certain characteristic exists in the real world, the minimum cardinality shall be defined as 1. For example, if an Administrative Unit always has at least one name, the multiplicity of the corresponding property shall be 1..*.

In both cases, the «voidable» stereotype can be applied. A value (the real value or void) only needs to be made available for properties that have a minimum cardinality of 1.

5.1.4 Code lists and Enumerations

5.1.4.1. Style

All code lists and enumerations use the following modelling style:

- No initial value, but only the attribute name part, is used.
- The attribute name conforms to the rules for attributes names, i.e. is a lowerCamelCase name. Exceptions are words that consist of all uppercase letters (acronyms).

5.1.4.2. Governance of code lists

Two types of code lists are defined in INSPIRE. These two types are distinguished using the tagged value “extendableByMS” in the UML data model:

- *Code lists that **may not** be extended by Member States.* For these code lists, the tagged value is set to “false”. They shall be managed centrally in the INSPIRE code list register, and only values from that register may be used in instance data.
- *Code lists that **may** be extended by Member States.* For these code lists, the tagged value is set to “true”.

5.2 Application schema Population Distribution - Demography

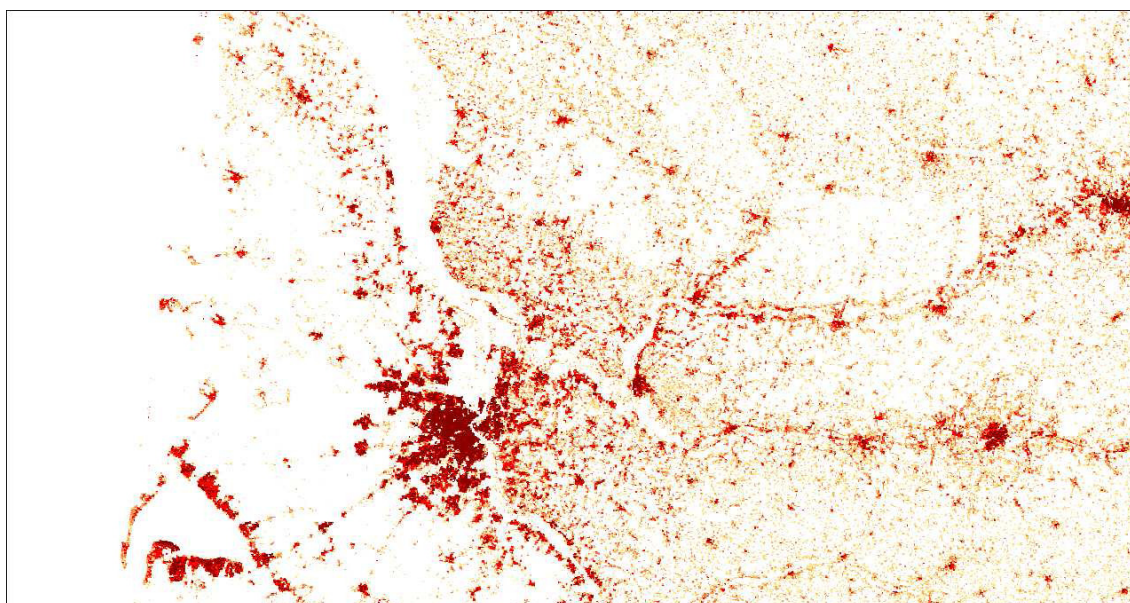
5.2.1 Description

5.2.1.1. Narrative description

A statistical data distribution describes the structure of a "dissemination area" based on a set of measures. The notion of "*population distribution*" can be best understood with the following example. The picture represents the way the population is distributed within the surroundings of Bordeaux, here defined as a geographical rectangle. The surroundings of Bordeaux are our "**area of Dissemination**". This area is pictured with a set of red or white pixels which are actually small 200m wide grid cells.

For each cell population counts were produced: the "**value**" is just a number of people. The pixels are white when population is zero and darken when the population count gets higher. But the whole set of pixels - the whole set of grid cells - provide a complete picture of the area.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 8



The aim of the proposed data model is to provide capabilities of exchanging the data that is behind such pictures. In this example: the whole set of figures that are associated to each pixel that are necessary to build the global picture that will represent the area of dissemination. The model is not restricted just to one form of area of dissemination, it is not restricted to grid cells as elementary components of the area of dissemination and it is not restricted to simple population counts. The model is a generic one to allow almost any statistical data to be exchanged.

One part of the disseminated data is a kind of metadata that is not covered by the ISO 19115/9 metadata standards and that should be added as a header in the dataset.

An example:

Attribute	Value
Domain	"demography"
Measure	"populationAtResidencePlace"
Classifications	"sex"
measurementUnit	"person"
areaOfDissemination	"surroundings of Bordeaux"

The **domain** describes the domain of statistical knowledge the value belongs to and the **measure** contains the short term of the phenomenon which is counted. The **classifications** describe the type of classification used to calculate or aggregate the values. The **area of Dissemination** describes the area for which the statistical data are available.

Temporal dimension of population distribution is considered in the three aspects: the actual period of measurement and the theoretical reference period of the measurement and the period the data remain valid.

The second part contains the elementary data values and the description by dimensions. This part applies as often as values are available in the area disseminated / requested.

Attribute	Value
Value	<number>
classificationItem	"male"
geoReference	<Statistical Unit>

NOTE: This example is simplified and does not represent all attributes of the model.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 9

The Population distribution – demography data model is defined in a very generic way, and nearly any statistical value describing human population can be considered a part of it. Use cases show the need for statistical variables as *population at residence place*, *population at work place* and *population density*. These statistics need to be provided on a more detailed territorial level than what is possible to access already through the Eurostat databases, preferable on statistical units like grids or census districts.

The requirements (which will be prepared for version 3 of this data specification) must reflect the role of this theme in the Inspire context. The National Statistical Institutes should not be forced to provide Inspire services for all statistics covering human population, as building two parallel systems (regional statistics already provided by Eurostat and regional statistics provided by the member states through Inspire) will both be costly and time consuming.

It is recommended to at least provide the most common attributes in combination with the most common statistical units. As an orientation providers should publish the variables as much as possible like it is written in data requirements listed in EU Regulations, such as the case of the 2011 Population and Housing Census in Europe (EC/763/2008).

Open issue 3: Reviewers and testers are asked to declare their possibilities / expectations according the possible provision of variables and the minimum level of aggregation, to give an orientation for the TWG (see 5.2.1.2 Codelists)

5.2.1.2. UML Overview

An overview of the Population distribution - demography package and referenced packages is provided in Figure 1 below. The Population Distribution - Demography package refers to the Base Types package for Identifier data type. In addition the StatisticalUnit feature type of Annex III theme Statistical units is referenced to support the indirect positioning of statistical data with respect to the location on the Earth.

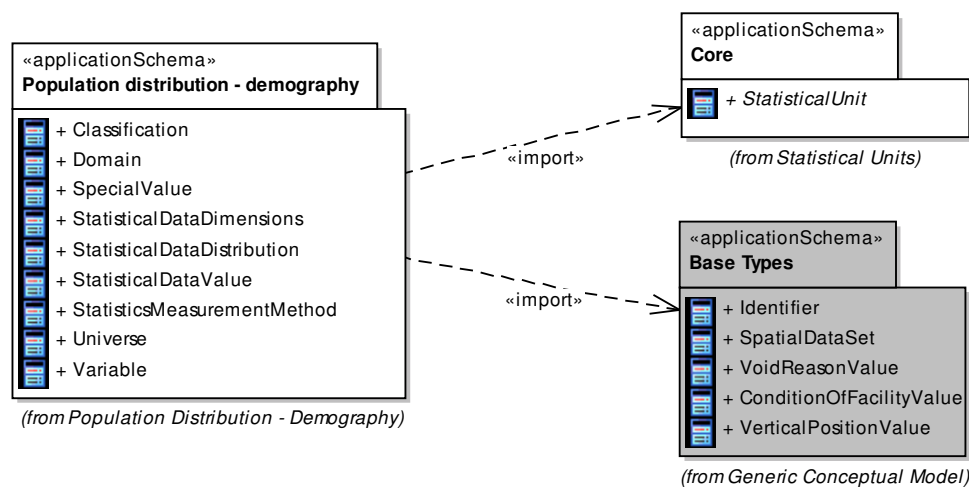


Figure 1: Package Overview

The application schema for Population distribution - demography is shown in Figure 2 and described below.

The schema is structured with three types of classes: “StatisticalDataDistribution”, “StatisticalDataValue” and “StatisticalDataDimensions” and five code lists.

Open issue 4: Please note that the code lists **domain**, **variable** and **classification** need further considerations why feedback during the review is essential! The TWG propose that the code lists

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 10

should contain the minimum required data which the user can expect but that they should be extendable by data providers.

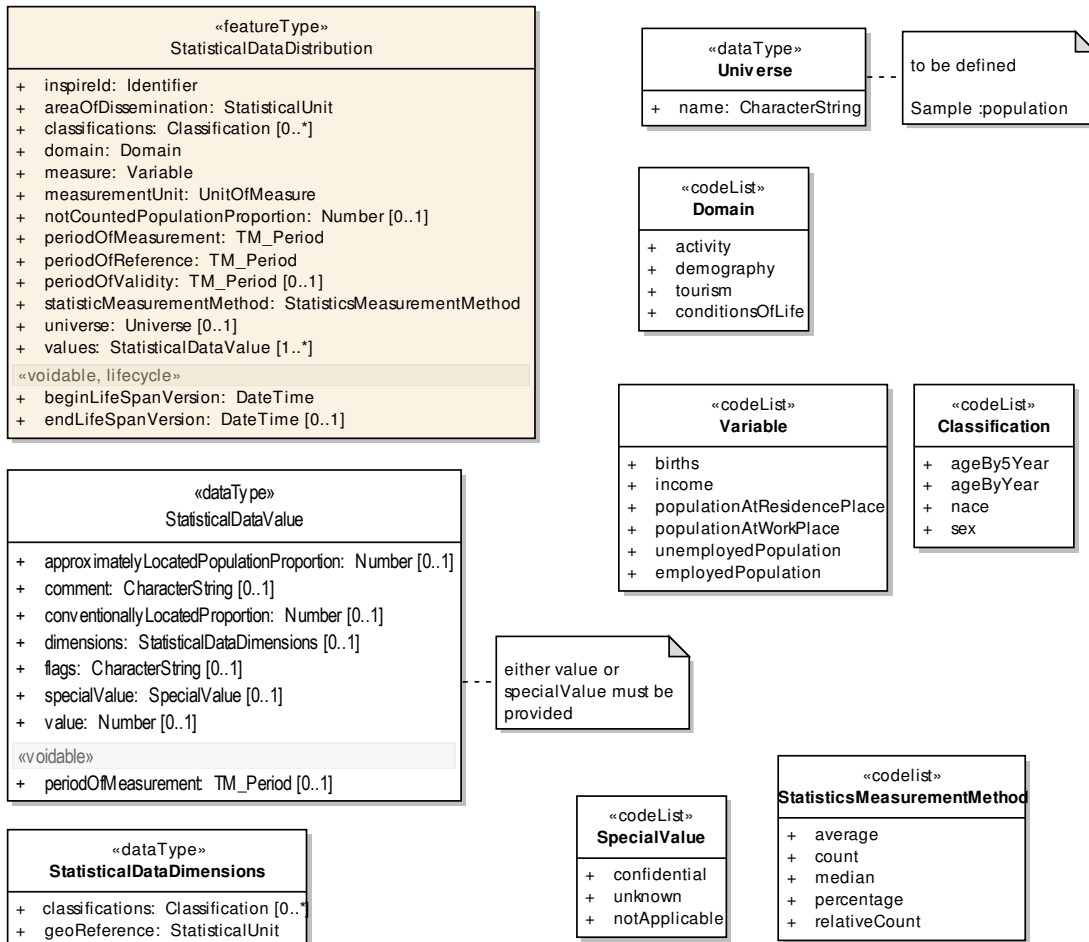


Figure 2: class diagram Population Distribution - Demography

The StatisticalDataDistribution is defined as a feature type for practical reasons to be able to generate the according GML. This type mainly acts as a header to keep general information about the distribution and contains common properties of each component of the distribution.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 11

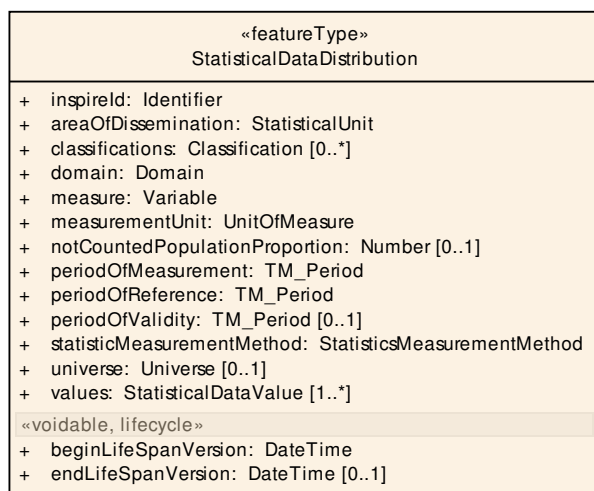


Figure 3: feature type StatisticalDataDistribution

InspireId is probably not necessary, but included to be decided later.

areaOfDissemination: This attribute is a particular instance of a statistical unit as they can be defined using the "*statistical units*" guidelines. Practically, the areas of dissemination will be fixed areas tightly related to the statistical production system. If the area of dissemination is described as a set of grid cells, every cell will have the same size. If it is described as a set of NUTS, all the NUTS will have to belong to the same level. The reason is that it is very unlikely that a statistical system is organized to make data production on a mix of different geographical levels. It could indeed produce data for different geographical levels but as distinct levels.

The optional **classification** attribute can be used to represent a set of distributions with respect to the way of splitting the statistical values into different groups using one or several classifications of the individuals according to their characteristics.

The **domain** attribute provides information about the domain of statistical knowledge.

The attribute **measure** contains the value of the code list variable, which describes in short- term the phenomena for which data provided.

The attributes **measurementUnit** and **statisticsMeasurementMethod** support description of the kind of measurements provided by the distribution. The possible values are taken from the corresponding code lists **UnitOfMeasure** and **statisticsMeasurementMethod**.

REMARK: For a harmonised and better reading, the naming of code lists will be changed according to the recommendations, provided in the summary of the domain value discussion.

Since the time dimension is an important component of statistical information, the model includes three attributes to address this issue:

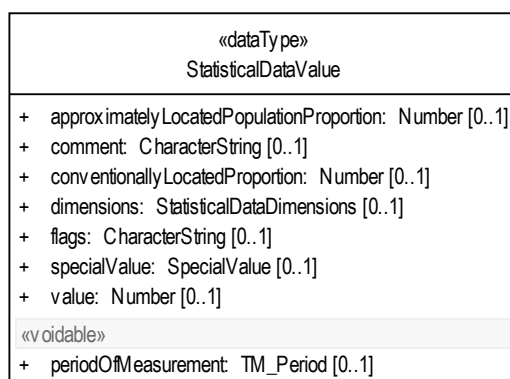
- the period when the data used for statistics was collected in **periodOfMeasurement**,
- the period the statistical information refers to in **periodOfReference**,
- the period for which the statistical information remains relevant in **periodOfValidity**.

According the common approach in Inspire to describe the quality of the content of a dataset, some data quality considerations appear through the **populationNotCountedPorportion** (will be renamed **notCountedProportion**) attribute that reports the relative amount of population known to be in the area of interest but not delivered in any of its spatial elements.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 12

The **universe** attribute provides information if the statistics cannot be assigned to domain and measure. This is the case if statistics are produced for a certain use case on user request and must be documented here.

The data type **StatisticalDataValue** supports the description of the individual values of the distributions: This description contains the reference to understand to which spatial element and to which classification items it refers to. The elementary datum can be seen as something that can be located in an n-dimensional space comprising the classical two dimensional one but also the dimensions relative to the classifications used. For example, a simple population count on each municipality or a one country will only need a reference to the two dimensional space viewed through the municipalities list. But the same population divided by sex will need an additional third dimension.



In the centre of this item is the **value** attribute, which contains the number statistically calculated.

In case the value cannot be provided (e.g. privacy, error, confidentiality) the reason for value absence can be expressed using the **specialValue** attribute. It is required that either a value or a special value is provided.

Values can be accompanied with **flags** and footnotes (**comment** attribute) giving any kind of information useful to describe particular situations.

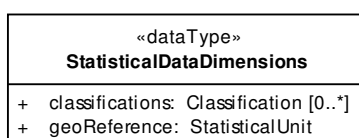
REMARK: The attribute “flags” is a candidate for a code list.

Two additional attributes related to data quality gives hints about the accuracy of the measure: **conventionallyLocatedPopulationPorportion** and **approximatelyLocatedPopulationProportion**. They are discussed in the quality section.

Additionally, for each StatisticalDataValue a time stamp (**periodOfMeasurement**) can be provided, relatively to its actual date of measurement to allow exceptions to what was specified in the general information.

According to the SDMX view on dimensions, all dimensions of the value (classification and geographical) are summarized in a unique attribute **dimensions** and can be found in the data type StatisticalDataDimensions.

Note: The cardinality of the attribute dimensions is 1, because at least the geographical reference has to be provided.



INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 13

In case an additional splitting / grouping of the statistical values are needed according to one or several classifications, the attribute **classifications** will be used.

The geographic reference (**geoReference**) contains the Inspire Identifier of the Statistical Unit the value is related to. It is a specific instance of the StatisticalUnit feature type.

Code lists

The model contains 6 code lists. Domain and Variable are used to identify the statistical object, the attribute Classifications subdivides the values or Special Values and measurement unit together with the statistical measurement method characterise the value.

The code lists **Domain**, **Variable** and **Classifications** (explicitly the type of classification) are used to define the data which are provided. These code lists are used to separate the mandatory from the optional values.

According the generally discussed use of domain values, the mandatory values should be listed in the basic code list and the optional ones in an extension under the control of the provider / member state. This also should include the flexibility to react to further requirements coming from new legislations. Due to the fact that this concept is not supported by corresponding registers, testers are asked to build their own code lists for testing.

The code lists are defined in the feature catalogue below (5.2.2.3). The code list values are examples and are not repeated here. The TWG will check if the measurement oriented codes can be aligned with the general concept of "Observation & Measurement", when it's available.

NOTE: The code list "Classifications" describes the kind / type of classification used and is only applicable to the attribute "classification" in the feature type StatisticalDataDistribution (should be renamed to "ClassificationType". The value of the attribute "classification" in the data type StatisticalDataDimensions describes the characteristic of the value and can be named "ClassificationValue".

Example: For the classification type "ageBy5Year" the classification values are "<5", "5 to <10", "10 to <15" etc.

Whether the classification values should be collected in a code list is debatable.

Open issue 5: Because it is controversial within the TWG which data should be mandatory to provide, the TWG is asking the reviewers and testers what should finally be included as required (=mandatory) data, on which regional level and in which classification types / classification values. These statements should be given in the comment- sheet and justified.

5.2.1.3. Consistency between spatial data sets

There is no specific recommendation regarding consistency between spatial data sets.

5.2.1.4. Identifier management

The feature type StatisticalDataDistribution has an identifier attribute. Following the general Inspire rules each object has an ID, but exceptions are possible for themes of Annex III.

In the case of including the ID, this identifier shall be maintained by the national or regional authority. The identifier shall consist of two parts: the namespace and a local id. The namespace is to uniquely identify a national registry wherein the identifier is registered, whereas id is to unique identifying an object within this namespace. The pragmatic approach to making it internationally unique is to add a prefix of the Member State identifier along with a theme specific identifier for the namespace. Since INSPIRE objectives refer to data exchange only, the maintenance and management of unique INSPIRE identifiers is out of INSPIRE scope, and is in the responsibility of the Member States.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 14

If no ID is included some other data about the provider and base has to be provided. This should be the providing institution and if applicable the legislation. (See: AM specs)

Open issue 6: It is an open question for the TWG to include such an ID or provider information.

5.2.2 Feature catalogue

Table 3 - Feature catalogue metadata

Feature catalogue name	INSPIRE feature catalogue Population distribution - demography
Scope	Population distribution - demography
Version number	2.0.1
Version date	2011-07-13
Definition source	INSPIRE data specification Population distribution - demography

Table 4 - Types defined in the feature catalogue

Type	Package	Stereotypes	Section
Classification	Population distribution - demography	«codeList»	5.2.2.3.1
Domain	Population distribution - demography	«codeList»	5.2.2.3.2
SpecialValue	Population distribution - demography	«codeList»	5.2.2.3.3
StatisticalDataDimensions	Population distribution - demography	«dataType»	5.2.2.2.1
StatisticalDataDistribution	Population distribution - demography	«featureType»	5.2.2.1.1
StatisticalDataValue	Population distribution - demography	«dataType»	5.2.2.2.2
Universe	Population distribution - demography	«dataType»	5.2.2.2.3
Variable	Population distribution - demography	«codeList»	5.2.2.3.4

5.2.2.1. Spatial object types

5.2.2.1.1. *StatisticalDataDistribution*

StatisticalDataDistribution	
Name:	Statistical Data Distribution
Definition:	Entire set of measures describing how a phenomenon is spread within some part of the 2D world.
Status:	Proposed
Stereotypes:	«featureType»
URI:	null
Attribute: areaOfDissemination	
Value type:	StatisticalUnit
Definition:	The part of the 2D world the StatisticalDataDistribution describes.
Description:	This attribute links to some StatisticalUnit as described in the “Statistical Units” guidelines.
Multiplicity:	1
Attribute: beginLifeSpanVersion	
Value type:	DateTime

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 15

StatisticalDataDistribution

Definition: Date and time at which this version of the spatial object was inserted or changed in the spatial data set.
 Multiplicity: 1
 Stereotypes: «voidable,lifecycle»

Attribute: classifications

Value type: Classification
 Definition: Additional classifications used to split a total value of the described phenomenon. In this way the StatisticalDistribution object will provide actually several distributions, one for each item of the used classification.
 Multiplicity: 0..*

Attribute: domain

Value type: Domain
 Definition: The part of statistical knowledge the data refers to.
 Description: EXAMPLES: Demography, activity, etc.
 Multiplicity: 1

Attribute: endLifeSpanVersion

Value type: DateTime
 Definition: Date and time at which this version of the spatial object was superseded or retired in the spatial data set.
 Multiplicity: 0..1
 Stereotypes: «voidable,lifecycle»

Attribute: inspireId

Value type: Identifier
 Definition: External object identifier of the spatial object.
 Description: An external object identifier is a unique object identifier published by the responsible body, which may be used by external applications to reference the spatial object. The identifier is an identifier of the spatial object, not an identifier of the real-world phenomenon.
 Multiplicity: 1

Attribute: measure

Value type: Variable
 Definition: What the distribution measures.
 Description: EXAMPLE: Population by age and sex, households by size, etc.
 Multiplicity: 1

Attribute: measurementUnit

Value type: UnitOfMeasure
 Definition: The unit of the measurement.
 Multiplicity: 1

Attribute: notCountedPopulationProportion

Value type: Number
 Definition: The proportion of population of the area of interest that is not counted in any of its spatial components.
 Description: EXAMPLE: This situation can occur when the total population of the area of interest is collected by some process (e.g. administrative register) but population in the spatial components of the area of interest by another process (e.g. estimation).

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 16

StatisticalDataDistribution	
Multiplicity:	0..1
Attribute: periodOfMeasurement	
Value type:	TM_Period
Definition:	The time period when the data was collected.
Multiplicity:	1
Attribute: periodOfReference	
Value type:	TM_Period
Definition:	The time period to which the data is supposed to give a picture of the area of interest.
Multiplicity:	1
Attribute: periodOfValidity	
Value type:	TM_Period
Definition:	The time period during which the data is usable. Outside this period the data should not be used and must be replaced by a fresher one.
Multiplicity:	0..1
Attribute: statisticMeasurementMethod	
Value type:	StatisticsMeasurementMethod
Definition:	The description of the statistic measurement method.
Multiplicity:	1
Attribute: universe	
Value type:	Universe
Definition:	When distribution is related to a subset of the population and not the population in its whole, the literal description of the way this subset was defined. EXAMPLE : A dataset can give the population distribution of the unemployed persons under 25. The variable will be "unemployedPopulation", the measure a "count " and the measurementUnit "person". The only way to express that the distribution doesn't cover the whole unemployed population would be to describe that situation in the Universe attribute.
Multiplicity:	0..1
Attribute: values	
Value type:	StatisticalDataValue
Definition:	The pieces of datum of the distribution.
Multiplicity:	1..*

5.2.2.2. Data types

5.2.2.2.1. *StatisticalDataDimensions*

StatisticalDataDimensions	
Name:	Statistical Data Dimensions
Definition:	The identification of what the piece of datum refers to in terms of geographic location or individual characteristics.
Status:	Proposed
Stereotypes:	«dataType»
URI:	null
Attribute: classifications	
Value type:	Classification
Definition:	A set of individual characteristics.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 17

StatisticalDataDimensions

Description: The actual value are codelists items taken from the classifications codelists that are listed in the 'classifications' attribute of the StatisticalData object.

EXAMPLE: A StatisticalDataDistribution is delivered as a dataset containing separately the spatial distribution of the male persons and of the female persons. The 'classifications' attribute of the StatisticalDataDistribution object will be 'Sex', while for each elementary spatial component data will be provided alternatively with the attribute classificationItem' having the value 'male' and the value 'female'.

Multiplicity: 0..*

Attribute: geoReference

Value type: StatisticalUnit

Definition: The location the piece of datum refers to.

Multiplicity: 1

5.2.2.2.2. StatisticalDataValue

StatisticalDataValue

Name: Statistical Data Value

Definition: The pieces of datum of the distribution.

Status: Proposed

Stereotypes: «dataType»

URI: null

Attribute: approximatelyLocatedPopulationProportion

Value type: Number

Definition: The proportion of population count that doesn't follow the common rule for location. "Population" has here a general meaning : it can be persons if persons are counted, dwellings if the StatisticalDataDistribution is about dwellings...
EXAMPLE : A statistical distribution of the people is supposed to have been produced after the people were located in the cadastral parcel they live in. For some incorrectness in the registers used, the actual location of the cadastral parcel was not found for some of them and an estimated location was computed. The number of people in this case must be reported in the attribute.

Multiplicity: 0..1

Attribute: comment

Value type: CharacterString

Definition: Some free style comment about the piece of datum.

Multiplicity: 0..1

Attribute: conventionallyLocatedProportion

Value type: Number

Definition: The proportion of population counted in the piece of datum but that cannot actually be physically located anywhere within the area of interest. EXAMPLE: This proportion will include homeless people who were given the address of the organisation that distributes social benefits to them.

Multiplicity: 0..1

Attribute: dimensions

Value type: StatisticalDataDimensions

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 18

StatisticalDataValue	
Definition:	The part of the world the piece of datum refers to. Dimensions contains a description of the geographic location (2D dimension) together with possible additional dimensions when population counts are produced simultaneously for different individual characteristics. EXAMPLE: A simple population count for each part of the area of Dissemination will need only identification of these parts, while population count by sex will require an additional dimension (sex) to make a difference between the male and the female distributions.
Multiplicity:	0..1
Attribute: flags	
Value type:	CharacterString
Definition:	A set of one-character encoded comments about the piece of datum.
Multiplicity:	0..1
Attribute: periodOfMeasurement	
Value type:	TM_Period
Definition:	The period of data collection for the piece of datum, when this one differs from the one of the other pieces of datum of the distribution.
Multiplicity:	0..1
Stereotypes:	«voidable»
Attribute: specialValue	
Value type:	SpecialValue
Definition:	Some conventional string when value for the piece of datum cannot be provided: Missing value, value hidden because of confidentiality...
Description:	Either value of specialValue must be provided.
Multiplicity:	0..1
Attribute: value	
Value type:	Number
Definition:	The value for the piece of datum.
Description:	Either value of specialValue must be provided.
Multiplicity:	0..1

5.2.2.2.3. *Universe*

Universe	
Name:	Universe
Definition:	The universe.
Status:	Proposed
Stereotypes:	«dataType»
URI:	null
Attribute: name	
Value type:	CharacterString
Definition:	The name of the universe.
Multiplicity:	1

5.2.2.3. **Code lists**

5.2.2.3.1. *Classification*

Classification	
Name:	Classification
Definition:	The classification used.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 19

Classification

Description: This code list should be extended, the items listed here are only for example A classification is actually a codelist. Such codelists are not described here, they are supposed to be external.

Status: Proposed

Stereotypes: «codeList»

Governance: May be extended by Member States.

URI:

Value: ageBy5Year

Definition: Age by 5 years periods.

Value: ageByYear

Definition: Age by year.

Value: nace

Definition: The economic activity classification.

Value: sex

Definition: Male or female.

5.2.2.3.2. Domain

Domain

Name: Domain

Definition: The domain of statistical knowledge.

Description: This code list should be extended, the items listed here are only for example.

Status: Proposed

Stereotypes: «codeList»

Governance: May be extended by Member States.

URI:

Value: activity

Definition: Everything that deals with economic activity of the human population. This domain includes the wages distributed by the enterprises.

Value: conditionsOfLife

Definition: Everything that deals with condition of life : this domain includes the financial resources of the population : income from activity and social benefits.

Value: demography

Definition: Pure population counts, births and deaths.

Value: tourism

Definition: Everything that deals with tourism.

5.2.2.3.3. SpecialValue

SpecialValue

Name: Special

Definition: A conventional replacement value when the value cannot be provided.

Description: The replacement value also tells why the value cannot be provided. This code list should be extended; the items listed here are only for example.

Status: Proposed

Stereotypes: «codeList»

Governance: May be extended by Member States.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 20

SpecialValue	
URI:	
Value: confidential	
Definition:	Providing the actual value would break some confidentiality rule.
Value: notApplicable	
Definition:	Providing a value would not make any sense.
Value: unknown	
Definition:	The actual value could have been measured but was not.

5.2.2.3.4. Variable

Variable	
Name:	Variable
Definition:	The statistic.
Description:	This code list should be extended, the items listed here are only for example
Status:	Proposed
Stereotypes:	«codeList»
Governance:	May be extended by Member States.
URI:	
Value: births	
Definition:	The number of births.
Value: employedPopulation	
Definition:	The employed population.
Value: income	
Definition:	The income.
Value: populationAtResidencePlace	
Definition:	The number of people measured in the place where they sleep.
Value: populationAtWorkPlace	
Definition:	The number of people measured in the place they where they work.
Value: unemployedPopulation	
Definition:	The unemployed population.

5.2.2.4. Imported types (informative)

This section lists definitions for feature types, data types and enumerations and code lists that are defined in other application schemas. The section is purely informative and should help the reader understand the feature catalogue presented in the previous sections. For the normative documentation of these types, see the given references.

5.2.2.4.1. CharacterString

CharacterString	
Package:	INSPIRE Consolidated UML Model::Foundation Schemas::ISO TC211::ISO 19103:2005 Schema Language::Basic Types::Primitive::Text [Include reference to the document that includes the package, e.g. INSPIRE data specification, ISO standard or the GCM]

5.2.2.4.2. DateTime

DateTime	
-----------------	--

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 21

DateTime

Package: INSPIRE Consolidated UML Model::Foundation Schemas::ISO TC211::ISO 19103:2005 Schema Language::Basic Types::Primitive::Date and Time [Include reference to the document that includes the package, e.g. INSPIRE data specification, ISO standard or the GCM]

5.2.2.4.3. *Identifier*

Identifier

Package: INSPIRE Consolidated UML Model::Generic Conceptual Model::Base Types [Include reference to the document that includes the package, e.g. INSPIRE data specification, ISO standard or the GCM]

Definition: External unique object identifier published by the responsible body, which may be used by external applications to reference the spatial object.

Description: NOTE1 External object identifiers are distinct from thematic object identifiers.

NOTE 2 The voidable version identifier attribute is not part of the unique identifier of a spatial object and may be used to distinguish two versions of the same spatial object.

NOTE 3 The unique identifier will not change during the life-time of a spatial object.

5.2.2.4.4. *Number*

Number (abstract)

Package: INSPIRE Consolidated UML Model::Foundation Schemas::ISO TC211::ISO 19103:2005 Schema Language::Basic Types::Primitive::Numerics [Include reference to the document that includes the package, e.g. INSPIRE data specification, ISO standard or the GCM]

5.2.2.4.5. *StatisticalUnit*

StatisticalUnit (abstract)

Package: INSPIRE Consolidated UML Model::Themes::Annex III::Statistical Units::Core [Include reference to the document that includes the package, e.g. INSPIRE data specification, ISO standard or the GCM]

Definition: Unit for dissemination or use of statistical information.

Description: SOURCE [INSPIRE Directive:2007].

EXAMPLE grid cell, point, line, polygon.

NOTE Spatial features of any INSPIRE application schema can be considered as a statistical unit, because all can be used as spatial reference. This class is provided to represent features that are used only to disseminate statistical information and that are not included in another INSPIRE application schema.

5.2.2.4.6. *TM_Period*

TM_Period

Package: INSPIRE Consolidated UML Model::Foundation Schemas::ISO TC211::ISO 19108:2006 Temporal Schema::Temporal Objects [Include reference to the document that includes the package, e.g. INSPIRE data specification, ISO standard or the GCM]

5.2.2.4.7. *UnitOfMeasure*

UnitOfMeasure (abstract)

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 22

UnitOfMeasure (abstract)

Package: INSPIRE Consolidated UML Model::Foundation Schemas::ISO TC211::ISO 19103:2005 Schema Language::Basic Types::Derived::Units of Measure [Include reference to the document that includes the package, e.g. INSPIRE data specification, ISO standard or the GCM]

6 Reference systems

6.1 Coordinate reference systems

6.1.1 Datum

IR Requirement 3 For the coordinate reference systems used for making available the INSPIRE spatial data sets, the datum shall be the datum of the European Terrestrial Reference System 1989 (ETRS89) in areas within its geographical scope, and the datum of the International Terrestrial Reference System (ITRS) or other geodetic coordinate reference systems compliant with ITRS in areas that are outside the geographical scope of ETRS89. Compliant with the ITRS means that the system definition is based on the definition of the ITRS and there is a well established and described relationship between both systems, according to EN ISO 19111.

6.1.2 Coordinate reference systems

IR Requirement 4 INSPIRE spatial data sets shall be made available using one of the three-dimensional, two-dimensional or compound coordinate reference systems specified in the list below.

Other coordinate reference systems than those listed below may only be used for regions outside of continental Europe. The geodetic codes and parameters for these coordinate reference systems shall be documented, and an identifier shall be created, according to EN ISO 19111 and ISO 19127.

1. Three-dimensional Coordinate Reference Systems
 - Three-dimensional Cartesian coordinates
 - Three-dimensional geodetic coordinates (latitude, longitude and ellipsoidal height), using the parameters of the GRS80 ellipsoid
2. Two-dimensional Coordinate Reference Systems
 - Two-dimensional geodetic coordinates, using the parameters of the GRS80 ellipsoid
 - Plane coordinates using the Lambert Azimuthal Equal Area projection and the parameters of the GRS80 ellipsoid
 - Plane coordinates using the Lambert Conformal Conic projection and the parameters of the GRS80 ellipsoid
 - Plane coordinates using the Transverse Mercator projection and the parameters of the GRS80 ellipsoid

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 23

3. Compound Coordinate Reference Systems

- For the horizontal component of the compound coordinate reference system, one of the two-dimensional coordinate reference systems specified above shall be used
- For the vertical component on land, the European Vertical Reference System (EVRS) shall be used to express gravity-related heights within its geographical scope
- Other vertical reference systems related to the Earth gravity field shall be used to express gravity-related heights in areas that are outside the geographical scope of EVRS. The geodetic codes and parameters for these vertical reference systems shall be documented and an identifier shall be created, according to EN ISO 19111 and ISO 19127
- For the vertical component measuring the depth of the sea floor, where there is an appreciable tidal range, the Lowest Astronomical Tide shall be used as reference surface. In marine areas without an appreciable tidal range, in open oceans and effectively in waters that are deeper than 200 m, the depth of the sea floor shall be referenced to the Mean Sea Level
- For the vertical component measuring depths above the sea floor in the free ocean, barometric pressure shall be used
- For the vertical component in the free atmosphere, barometric pressure, converted to height using ISO 2533:1975 International Standard Atmosphere shall be used

6.1.3 Display

IR Requirement 5 For the display of the INSPIRE spatial data sets with the View Service specified in D003152/02 Draft Commission Regulation implementing Directive 2007/2/EC of the European Parliament and of the Council as regards Network Services, at least the two dimensional geodetic coordinate system shall be made available.

6.1.4 Identifiers for coordinate reference systems

IR Requirement 6 For referring to the non-compound coordinate reference systems listed in this Section, the identifiers listed below shall be used.

For referring to a compound coordinate reference system, an identifier composed of the identifier of the horizontal component, followed by a slash (/), followed by the identifier of the vertical component, shall be used.

- ETRS89-XYZ for Cartesian coordinates in ETRS89
- ETRS89-GRS80h for three-dimensional geodetic coordinates in ETRS89 on the GRS80 ellipsoid
- ETRS89-GRS80 for two-dimensional geodetic coordinates in ETRS89 on the GRS80
- EVRS for height in EVRS
- LAT for depth of the sea floor, where there is an appreciable tidal range
- MSL for depth of the sea floor, in marine areas without an appreciable tidal range, in open oceans and effectively in waters that are deeper than 200m
- ISA for pressure coordinate in the free atmosphere
- PFO for Pressure coordinate in the free ocean
- ETRS89-LAEA for ETRS89 coordinates projected into plane coordinates by the Lambert Azimuthal Equal Area projection
- ETRS89-LCC for ETRS89 coordinates projected into plane coordinates by the Lambert Conformal Conic projection
- ETRS89-TMzn for ETRS89 coordinates projected into plane coordinates by the Transverse Mercator projection

6.2 Temporal reference system

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 24

IR Requirement 7 The Gregorian Calendar shall be used for as a reference system for date values, and the Universal Time Coordinated (UTC) or the local time including the time zone as an offset from UTC shall be used as a reference system for time values.

6.3 Theme-specific requirements and recommendations on reference systems

There are no theme-specific requirements or recommendations on reference systems.

7 Data quality

This chapter includes a description of data quality elements and sub-elements as well as the associated data quality measures (section 7.1). The selected data quality measures should be used to evaluate quality of data sets for a specific data quality element / sub-element. The evaluation can be performed at the level of spatial object, spatial object type, dataset or dataset series.

The results of the evaluation are then reported at the spatial object type or dataset level in metadata utilising the same data quality elements and measures (see chapter 8).

NOTE The selection of appropriate data quality measures represents the first step towards the harmonisation of documenting data quality.

In addition, for some of the data quality elements described in section 7.1, minimum data quality requirements or recommendations may be defined. These are described in the section 1.2.

Recommendation 1 If data quality information is required at spatial object level then it should be modelled in the data model as an attribute of a relevant spatial object type.

7.1 Data quality elements and measures

Recommendation 2 To evaluate and report the data quality of data sets related to the spatial data theme **Population Distribution - Demography**, the elements and measures listed in Table 2 should be used.

Table 2 – Data quality elements for evaluating and reporting the data quality of data sets related to the spatial data theme Population Distribution - Demography

Section	Data quality element and sub-element
7.1.1	Completeness – Commission
7.1.2	Completeness – Omission

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 25

Error! Reference source not found.	Completeness - Population not counted
Error! Reference source not found.	Rough location - Conventionally located population
7.1.3	Rough location - Approximately located population

7.1.1 Completeness – Commission

Commission should be documented using the rate of excess items.

Name	Rate of excess items
Alternative name	–
Data quality element	Completeness
Data quality sub-element	Commission
Data quality basic measure	Error rate
Definition	Number of excess items in the dataset in relation to the number of items that should have been present.
Description	–
Evaluation scope	data set
Reporting scope	data set
Parameter	–
Data quality value type	Real, percentage, ratio (example: 0,0189 ; 98,11% ; 11:582)
Data quality value structure	–
Source reference	–
Example	–
Measure identifier	3 (ISO 19138)

7.1.2 Completeness – Omission

Omission should be documented using the rate of missing items.

Name	Rate of missing items
Alternative name	–
Data quality element	Completeness
Data quality sub-element	Omission
Data quality basic measure	Error rate
Definition	Number of missing items in the dataset in relation to the number of items that should have been present.
Description	–
Evaluation scope	data set
Reporting scope	data set
Parameter	–
Data quality value type	Real, percentage, ratio (example: 0,0189 ; 98,11% ; 11:582)
Data quality value structure	–
Source reference	–
Example	–
Measure identifier	7 (ISO 19138)

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 26

7.1.3 Completeness – Population not counted

Omission of some population in the *population distribution* should be documented using the following indicator.

Name	Population not counted
Alternative name	-
Data quality element	Completeness
Data quality sub-element	Population not counted
Data quality basic measure	Error rate
Definition	Proportion of individuals (= humans) that are present in the <i>area of dissemination</i> but not counted in any <i>spatial element</i>
Description	-
Evaluation scope	data set
Reporting scope	data set
Parameter	-
Data quality value type	Percentage
Data quality value structure	-
Source reference	-
Example	<p>When individual location is chosen as the centre of the cadastral parcel containing their residence, the indicator will report the proportion of the individuals for which the exact computation was not done and no estimate of the location was attempted (the actual cadastral parcel was not known and additional information was not available).</p> <p>When population distribution is related to population work place based, this indicator can be significantly non zero if the statistics attempt to reach the real work place concept, i.e. the geographic place the individual spends his work day time: Every people that have a moving working place (e.g. taxi drivers) should be counted as not located and reported in this indicator.</p>
Measure identifier	None

7.1.4 Rough location – Conventionally located population

Population counts that are recorded for each spatial element occurring in a population distribution could result from an exact measurement but in most of the cases it will result from a mixture of exact measurements and conventional choices and estimation processes. For example, with census based data, the location of homeless people is mainly conventional: they contribute to the population count of the data collection spatial unit they are found to be in at the time of the data collection. With register based data such people are generally attached to some conventional administrative addresses, e.g. the address of the charity organisation that is in charge for their health insurance.

Dealing with spatial elements that are defined by their geometric boundaries, particularly with grid cells or point locations, makes the question of recording population in the right spatial element more acute because counting can only occur after a precise geometric location was obtained for each individual and this is mainly matter of convention. Among these conventions are: the centre of the property unit, the centre of the building, the entrance of the building, etc... Additionally, in some cases such precise location will not be available and the final count will come from some estimation process. Again, the final value for a population count related to a spatial element can be depicted as a combination of a partial value obtained in conformance with the general rules for establishing the population count and another partial value obtained by softening the general rules : conventional choices or rough or modelled locations. Reporting the proportion related to the later figure is important because it directly

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 27

inform the user of the relevance of spatial or temporal comparisons. For this reason, the information has to be reported for each spatial element of a population distribution.

Name	Conventionally located population
Alternative name	-
Data quality element	Rough location
Data quality sub-element	Conventionally located population
Data quality basic measure	Level of lack of accuracy
Definition	Proportion of the individuals (= human persons) that could not be located in any specific <i>spatial element</i> and was counted in the spatial element.
Description	-
Evaluation scope	data set
Reporting scope	data set
Parameter	-
Data quality value type	Percentage
Data quality value structure	-
Source reference	-
Example	Homeless people
Measure identifier	None.

7.1.5 Rough location – Approximately located population

Name	Approximately located population
Alternative name	-
Data quality element	Rough location
Data quality sub-element	Approximately located population
Data quality basic measure	Level of lack of accuracy
Definition	Proportion of the individuals (= human persons) that could not be located using the basic rule used for determining the location of an individual, and thus which may be actually located in some other neighboring <i>spatial element</i> than the one they are counted into.
Description	-
Evaluation scope	data set
Reporting scope	data set
Parameter	-
Data quality value type	Percentage
Data quality value structure	-
Source reference	-
Example	When individual location is chosen as the centre of the cadastral parcel containing their residence, the indicator will report the proportion of the individuals for which the exact computation was not done (the actual cadastral parcel was not known and some other geographic elements were used).
Measure identifier	None.

7.2 Minimum data quality requirements and recommendations

No minimum data quality requirements are defined.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 28

8 Dataset-level metadata

Metadata can be reported for each individual spatial object (spatial object-level metadata) or once for a complete dataset or dataset series (dataset-level metadata). Spatial object-level metadata is fully described in the application schema (section 5). If data quality elements are used at spatial object level, the documentation shall refer to the appropriate definition in section 7. This section only specifies dataset-level metadata elements.

For some dataset-level metadata elements, in particular on data quality and maintenance, a more specific scope can be specified. This allows the definition of metadata at sub-dataset level, e.g. separately for each spatial object type. When using ISO 19115/19139 to encode the metadata, the following rules should be followed:

- The scope element (of type DQ_Scope) of the DQ_DataQuality subtype should be used to encode the scope.
- Only the following values should be used for the level element of DQ_Scope: Series, Dataset, featureType.
- If the level is featureType the levelDescription/MDScopeDescription/features element (of type Set< GF_FeatureType>) shall be used to list the feature type names.

NOTE The value featureType is used to denote spatial object type.

Mandatory or conditional metadata elements are specified in Section 8.1. Optional metadata elements are specified in Section 8. The tables describing the metadata elements contain the following information:

- The first column provides a reference to a more detailed description.
- The second column specifies the name of the metadata element.
- The third column specifies the multiplicity.
- The fourth column specifies the condition, under which the given element becomes mandatory (only for Table 3 and Table 4).

8.1 Common metadata elements

IR Requirement 8 The metadata describing a spatial data set or a spatial data set series related to the theme **Population Distribution - Demography** shall comprise the metadata elements required by Regulation 1205/2008/EC (implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata) for spatial datasets and spatial dataset series (Table 3) as well as the metadata elements specified in Table 4.

Table 3 – Metadata for spatial datasets and spatial dataset series specified in Regulation 1205/2008/EC (implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata)

Metadata Regulation Section	Metadata element	Multiplicity	Condition
1.1	Resource title	1	
1.2	Resource abstract	1	
1.3	Resource type	1	

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 29

1.4	Resource locator	0..*	Mandatory if a URL is available to obtain more information on the resource, and/or access related services.
1.5	Unique resource identifier	1..*	
1.7	Resource language	0..*	Mandatory if the resource includes textual information.
2.1	Topic category	1..*	
3	Keyword	1..*	
4.1	Geographic bounding box	1..*	
5	Temporal reference	1..*	
6.1	Lineage	1	
6.2	Spatial resolution	0..*	Mandatory for data sets and data set series if an equivalent scale or a resolution distance can be specified.
7	Conformity	1..*	
8.1	Conditions for access and use	1..*	
8.2	Limitations on public access	1..*	
9	Responsible organisation	1..*	
10.1	Metadata point of contact	1..*	
10.2	Metadata date	1	
10.3	Metadata language	1	

Table 4 – Mandatory and conditional common metadata elements

INSPIRE Data Specification Population Distribution - Demography Section	Metadata element	Multiplicity	Condition
8.1.1	Coordinate Reference System	1	
8.1.2	Temporal Reference System	0..*	Mandatory, if the spatial data set or one of its feature types contains temporal information that does not refer to the Gregorian Calendar or the Coordinated Universal Time.
8.1.3	Encoding	1..*	

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 30

8.1.4	Character Encoding	0..*	Mandatory, if an encoding is used that is not based on UTF-8.
8.1.5	Data Quality – Logical Consistency – Topological Consistency	0..*	Mandatory, if the data set includes types from the Generic Network Model and does not assure centreline topology (connectivity of centrelines) for the network.

8.1.1 Coordinate Reference System

Metadata element name	Coordinate Reference System
Definition	Description of the coordinate reference system used in the dataset.
ISO 19115 number and name	13. referenceSystemInfo
ISO/TS 19139 path	referenceSystemInfo
INSPIRE obligation / condition	mandatory
INSPIRE multiplicity	1
Data type (and ISO 19115 no.)	189. MD_CRS
Domain	<p>Either the referenceSystemIdentifier (RS_Identifier) or the projection (RS_Identifier), ellipsoid (RS_Identifier) and datum (RS_Identifier) properties shall be provided.</p> <p>NOTE More specific instructions, in particular on pre-defined values for filling the referenceSystemIdentifier attribute should be agreed among Member States during the implementation phase to support interoperability.</p>
Implementing instructions	
Example	<pre>referenceSystemIdentifier: code: ETRS_89 codeSpace: INSPIRE RS registry</pre>
Example XML encoding	<pre><gmd:referenceSystemInfo> <gmd:MD_ReferenceSystem> <gmd:referenceSystemIdentifier> <gmd:RS_Identifier> <gmd:code> <gco:CharacterString>ETRS89 </gco:CharacterString> </gmd:code> </gmd:codeSpace> <gco:CharacterString>INSPIRE RS registry</gco:CharacterString> </gmd:codeSpace> </gmd:RS_Identifier> </gmd:referenceSystemIdentifier> </gmd:MD_ReferenceSystem> </gmd:referenceSystemInfo></pre>
Comments	

8.1.2 Temporal Reference System

Metadata element name	Temporal Reference System
Definition	Description of the temporal reference systems used in the dataset.
ISO 19115 number and name	13. referenceSystemInfo
ISO/TS 19139 path	referenceSystemInfo

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 31

INSPIRE obligation / condition	Mandatory, if the spatial data set or one of its feature types contains temporal information that does not refer to the Gregorian Calendar or the Coordinated Universal Time.
INSPIRE multiplicity	0..*
Data type (and ISO 19115 no.)	186. MD_ReferenceSystem
Domain	No specific type is defined in ISO 19115 for temporal reference systems. Thus, the generic MD_ReferenceSystem element and its reference SystemIdentifier (RS_Identifier) property shall be provided. NOTE More specific instructions, in particular on pre-defined values for filling the referenceSystemIdentifier attribute should be agreed among Member States during the implementation phase to support interoperability.
Implementing instructions	
Example	referenceSystemIdentifier: code: GregorianCalendar codeSpace: INSPIRE RS registry
Example XML encoding	<pre> <gmd:referenceSystemInfo> <gmd:MD_ReferenceSystem> <gmd:referenceSystemIdentifier> <gmd:RS_Identifier> <gmd:code> <gco:CharacterString>GregorianCalendar</gco:CharacterString> </gmd:code> </gmd:code> <gmd:codeSpace> <gco:CharacterString>INSPIRE RS registry</gco:CharacterString> </gmd:codeSpace> </gmd:RS_Identifier> </gmd:referenceSystemIdentifier> </gmd:MD_ReferenceSystem> </gmd:referenceSystemInfo> </pre>
Comments	

8.1.3 Encoding

Metadata element name	Encoding
Definition	Description of the computer language construct that specifies the representation of data objects in a record, file, message, storage device or transmission channel
ISO 19115 number and name	271. distributionFormat
ISO/TS 19139 path	distributionInfo/MD_Distribution/distributionFormat
INSPIRE obligation / condition	mandatory
INSPIRE multiplicity	1
Data type (and ISO 19115 no.)	284. MD_Format
Domain	See B.2.10.4. The property values (name, version, specification) specified in section 9 shall be used to document the default and alternative encodings.
Implementing instructions	

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 32

Example	name: Population Distribution - Demography GML application schema version: version 2.0.1 , GML, version 3.2.1 specification: D2.8.III.10 Data Specification on Population Distribution - Demography – Draft Guidelines
Example XML encoding	<pre> <gmd:MD_Format> <gmd:name> <gco:CharacterString> Population Distribution - Demography GML application schema </gco:CharacterString> </gmd:name> <gmd:version> <gco:CharacterString>2.0.1, GML, version 3.2.1</gco:CharacterString> </gmd:version> <gmd:specification> <gco:CharacterString>D2.8.III.10 Data Specification on Population Distribution - Demography – Draft Guidelines</gco:CharacterString> </gmd:specification> </gmd:MD_Format> </pre>
Comments	

8.1.4 Character Encoding

Metadata element name	Character Encoding
Definition	The character encoding used in the data set.
ISO 19115 number and name	
ISO/TS 19139 path	
INSPIRE obligation / condition	Mandatory, if an encoding is used that is not based on UTF-8.
INSPIRE multiplicity	0..*
Data type (and ISO 19115 no.)	
Domain	
Implementing instructions	
Example	-
Example XML encoding	<pre> <gmd:characterSet> <gmd:MD_CharacterSetCode codeListValue="8859part2" codeList="http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/Codelist/ML_gmxCodellists.xml#CharacterSetCode">8859-2</gmd:MD_CharacterSetCode> </gmd:characterSet> </pre>
Comments	

8.1.5 Data Quality – Logical Consistency – Topological Consistency

Metadata element name	Data Quality – Logical Consistency – Topological Consistency
Definition	Correctness of the explicitly encoded topological characteristics of the dataset as described by the scope
ISO 19115 number and name	18. dataQualityInfo
ISO/TS 19139 path	dataQualityInfo

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 33

INSPIRE obligation / condition	Mandatory, if the data set includes types from the Generic Network Model and does not assure centreline topology (connectivity of centrelines) for the network.
INSPIRE multiplicity	0..*
Data type (and ISO 19115 no.)	115. DQ_TopologicalConsistency
Domain	Lines 100-107 from ISO 19115
Implementing instructions	This metadata should be filled, at least, with these elements: - valueUnit: UnitOfMeasure - value: Record
Example	
Example XML encoding	
Comments	See clauses on topological consistency in section 7 for detailed information. This metadata element is mandatory if connectivity is not assured for network centrelines in the dataset. In this case the <i>Connectivity tolerance</i> parameter – as described in section 7 – must be provided in order to ensure automatic and unambiguous creation of centreline topology in post-process.

8.2 Metadata elements for reporting data quality

Recommendation 3 For reporting the results of the data quality evaluation quantitatively, the data quality elements and measures defined in chapter 7 should be used.

The scope for reporting may be different from the scope for evaluating data quality (see section 7). If data quality is reported at the data set or spatial object type level, the results are usually derived or aggregated.

Metadata element name	See chapter 7
Definition	See chapter 7
ISO 19115 number and name	80. report
ISO/TS 19139 path	dataQualityInfo*/report
INSPIRE obligation / condition	optional
INSPIRE multiplicity	0..*
Data type (and ISO 19115 no.)	Corresponding DQ_xxx element from ISO 19115, e.g. 109. DQ_CompletenessCommission
Domain	Lines 100-107 from ISO 19115 100. nameOfMeasure : CharacterString [0..*] 101. measureIdentification : MD_Identifier [0..1] 102. measureDescription : CharacterString [0..1] 103. evaluationMethodType : DQ_EvaluationMethodTypeCode [0..1] 104. evaluationMethodDescription : CharacterString [0..1] 105. evaluationProcedure : CI_Citation [0..1] 106. dateTime : DateTime [0..*] 107. result : DQ_Result [1..2]

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 34

Implementing instructions	<p>Recommendation 4 For each DQ result included in the metadata, at least the following properties should be provided:</p> <p>100. nameOfMeasure NOTE This should be the name as defined in Chapter 7.</p> <p>103. evaluationMethodType</p> <p>104. evaluationMethodDescription NOTE If the reported data quality results are derived or aggregated (i.e. the scope levels for evaluation and reporting are different), the derivation or aggregation should also be specified using this property.</p> <p>106. dateTime NOTE This should be data or range of dates on which the data quality measure was applied.</p> <p>107. result NOTE This should be of type DQ_QuantitativeResult</p>
Example	
Example XML encoding	
Comments	See Chapter 7 for detailed information on the individual data quality elements and measures to be used.

Open issue 7: In the ongoing revision of ISO 19115 and development of new ISO 19157 standard (Geographic Information – Data quality), a new element is introduced (DQ_DescriptiveResult). This element enables to describe and report qualitative results of the data quality evaluation and could be used instead of DQ_QuantitativeResult. Once the new (version of the) standards are approved, these guidelines will be revisited and be updated if necessary.

Open issue 8: For reporting compliance with minimum data quality requirements and recommendations specified in section 7, the INSPIRE conformity metadata element should be used.

However, since this issue is part of the larger discussion on the Abstract Test Suite and the definition of conformance classes for the data specification, detailed instructions on how to provide metadata on compliance with minimum data quality requirements and recommendations will only be provided for v3.0.

8.3 Theme-specific metadata elements

No mandatory theme-specific metadata elements are defined for this theme.

Recommendation 5 The metadata describing a spatial data set or a spatial data set series related to the theme **Population Distribution - Demography** should comprise the theme-specific metadata elements specified in Table 5.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 35

Table 5 – Optional theme-specific metadata elements for the theme Population Distribution - Demography

INSPIRE Data Specification Population Distribution - Demography Section	Metadata element	Multiplicity
8.3.1	Maintenance Information	0..1
8.3.2	Data Quality – Completeness – Commission	0..*
8.3.3	Data Quality – Completeness – Omission	0..*

8.3.1 Maintenance Information

Metadata element name	Maintenance information
Definition	Information about the scope and frequency of updating
ISO 19115 number and name	30. resourceMaintenance
ISO/TS 19139 path	identificationInfo/MD_Identification/resourceMaintenance
INSPIRE obligation / condition	optional
INSPIRE multiplicity	0..1
Data type (and ISO 19115 no.)	142. MD_MaintenanceInformation
Domain	This is a complex type (lines 143-148 from ISO 19115). At least the following elements should be used (the multiplicity according to ISO 19115 is shown in parentheses): <ul style="list-style-type: none"> – maintenanceAndUpdateFrequency [1]: frequency with which changes and additions are made to the resource after the initial resource is completed / domain value: MD_MaintenanceFrequencyCode: – updateScope [0..*]: scope of data to which maintenance is applied / domain value: MD_ScopeCode – maintenanceNote [0..*]: information regarding specific requirements for maintaining the resource / domain value: free text
Implementing instructions	
Example	
Example XML encoding	
Comments	

8.3.2 Data Quality – Completeness – Commission

Metadata element name	Commission
Definition	Data absent from the dataset, as described by the scope.
ISO 19115 number and name	18. dataQualityInfo
ISO/TS 19139 path	dataQualityInfo
INSPIRE obligation / condition	optional
INSPIRE multiplicity	0..*
Data type (and ISO 19115 no.)	110. DQ_CompletenessOmission
Domain	Lines 100-107 from ISO 19115.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 36

8.3.3 Data Quality – Completeness – Omission

Metadata element name	Omission
Definition	Data absent from the dataset, as described by the scope.
ISO 19115 number and name	18. dataQualityInfo
ISO/TS 19139 path	dataQualityInfo
INSPIRE obligation / condition	optional
INSPIRE multiplicity	0..*
Data type (and ISO 19115 no.)	110. DQ_CompletenessOmission
Domain	Lines 100-107 from ISO 19115.

8.4 Guidelines on using metadata elements defined in Regulation 1205/2008/EC

8.4.1 Conformity

The *Conformity* metadata element defined in Regulation 1205/2008/EC allows to report the conformance with the Implementing Rule for interoperability of spatial data sets and services or another specification. The degree of conformity of the dataset can be *Conformant* (if the dataset is fully conformant with the cited specification), *Not Conformant* (if the dataset does not conform to the cited specification) or *Not evaluated* (if the conformance has not been evaluated).

Recommendation 6 The Conformity metadata element should be used to report conceptual consistency with this INSPIRE data specification. The value of Conformant should be used for the Degree element only if the dataset passes all the requirements described in the abstract test suite presented in Annex A. The Specification element should be given as follows:

- title: "INSPIRE Data Specification on <Theme Name> – Draft Guidelines"
- date:
 - dateType: publication
 - date: 2011-07-13

Open issue 9: Conformance testing is still an open issue under discussion.

Instructions on conformance testing and a common abstract test suite (including detailed instructions on how to test specific requirements) will be added at a later stage.

This may also lead to an update of the recommendations on how to fill the conformity metadata element.

8.4.2 Lineage

Recommendation 7 Following the ISO 19113 Quality principles, if a data provider has a procedure for quality validation of their spatial data sets then the data quality elements listed in the Chapters 7 and 8 should be used. If not, the *Lineage* metadata element (defined in Regulation 1205/2008/EC) should be used to describe the overall quality of a spatial data set.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 37

According to Regulation 1205/2008/EC, lineage “is a statement on process history and/or overall quality of the spatial data set. Where appropriate it may include a statement whether the data set has been validated or quality assured, whether it is the official version (if multiple versions exist), and whether it has legal validity. The value domain of this metadata element is free text”.

The Metadata Technical Guidelines based on EN ISO 19115 and EN ISO 19119 specify that the statement sub-element of LI_Lineage (EN ISO 19115) should be used to implement the lineage metadata element.

Recommendation 8 To describe the transformation steps and related source data, it is recommended to use the following sub-elements of LI_Lineage:

- For the description of the transformation process of the local to the common INSPIRE data structures, the LI_ProcessStep sub-element should be used.
- For the description of the source data the LI_Source sub-element should be used.

NOTE 1 This recommendation is based on the conclusions of the INSPIRE Data Quality Working Group to avoid overloading of the overall lineage statement element with information on the transformation steps and related source data.

NOTE 2 In order to improve the interoperability, domain templates and instructions for filling these free text elements (descriptions) may be specified in an Annex of this data specification.

Open issue 10: The suggested use of the LI_Lineage sub-elements needs to be discussed as part of the maintenance of the INSPIRE metadata Technical Guidelines.

8.4.3 Temporal reference

According to Regulation 1205/2008/EC, at least one of the following temporal reference metadata elements shall be provided: temporal extent, date of publication, date of last revision, date of creation. If feasible, the date of the last revision of a spatial data set should be reported using the *Date of last revision* metadata element.

9 Delivery

9.1 Delivery medium

DS Requirement 2 Data conformant to this INSPIRE data specification shall be made available through an INSPIRE network service.

DS Requirement 3 All information that is required by a calling application to be able to retrieve the data through the used network service shall be made available in accordance with the requirements defined in the Implementing Rules on Network Services.

EXAMPLE 1 Through the Get Spatial Objects function, a download service can either download a pre-defined data set or pre-defined part of a data set (non-direct access download service), or give direct access to the spatial objects contained in the data set, and download selections of spatial objects

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 38

based upon a query (direct access download service). To execute such a request, some of the following information might be required:

- the list of spatial object types and/or predefined data sets that are offered by the download service (to be provided through the Get Download Service Metadata operation),
- and the query capabilities section advertising the types of predicates that may be used to form a query expression (to be provided through the Get Download Service Metadata operation, where applicable),
- a description of spatial object types offered by a download service instance (to be provided through the Describe Spatial Object Types operation).

EXAMPLE 2 Through the Transform function, a transformation service carries out data content transformations from native data forms to the INSPIRE-compliant form and vice versa. If this operation is directly called by an application to transform source data (e.g. obtained through a download service) that is not yet conformant with this data specification, the following parameters are required:

Input data (mandatory). The data set to be transformed.

- Source model (mandatory, if cannot be determined from the input data). The model in which the input data is provided.
- Target model (mandatory). The model in which the results are expected.
- Model mapping (mandatory, unless a default exists). Detailed description of how the transformation is to be carried out.

9.2 Encodings

9.2.1 Default Encoding(s)

DS Requirement 4 Data conformant to the application schema(s) defined in section 5.2 shall be encoded using the encoding(s) specified in this section.

9.2.1.1 Default encoding for application schema **Population Distribution - Demography**

Name: Population Distribution - Demography GML Application Schema

Version: version 2.0.1, GML, version 3.2.1

Specification: D2.8.III.10 Data Specification on **Population Distribution - Demography** – Draft Guidelines

Character set: UTF-8

The GML Application Schema is distributed in a zip-file separately from the data specification document.

9.2.2 Alternative Encoding(s)

There is no specific requirement about the encoding of the data. But the target of the whole INSPIRE specifications is to allow GIS-like application to use the data. As it was written before population distribution objects are not spatial objects and will primarily come from production processes that are not very close to the world of GIS. can be used through a translation. Therefore it is strongly expected that the data will be first available using standards that are today in use in the world of statistics, such as SDMX. In its XML form SDMX dataset s structure is not very different from the GML one, so a general XSL tool will be provided to make the bridge between the two formats

10 Data Capture

There is no specific guidance required with respect to data capture.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 39

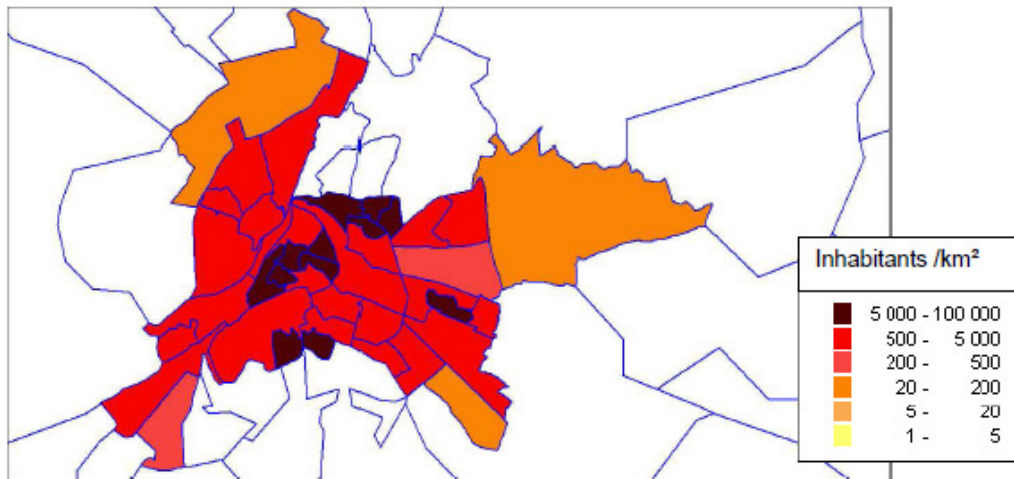
11 Portrayal

The unique object defined in the “Population Distributions – Demography” guidelines, *StatisticlaDataDistribution* is not a spatial object so there is no direct portrayal of it. But as distributions of statistical data are intended to give a description of how some phenomena is spread within an area of dissemination, viewing the contents of a distribution as an additional layer is a very common task. Two different techniques can be used:

- Changing the aspect of the elementary spatial elements of the distribution.
- Adding symbols

This section doesn't contain any portrayal requirement but is a remainder of some common sense considerations.

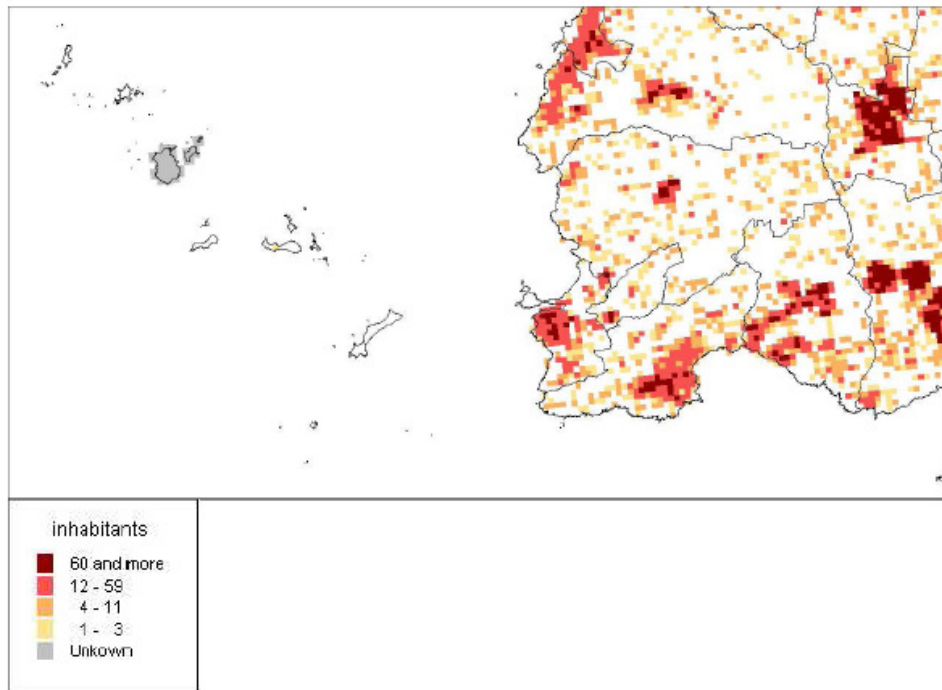
Changing the aspect of the elementary spatial elements



EXAMPLE 1: Viewing population density distribution for the city of Poitiers. Elementary spatial elements are neighborhoods. Their shapes are outlined in blue and they are filled with different colors, as darker as the density is high. Neighborhoods outside the city are left blank. Some rules must be followed when using this kind of viewing:

- Always provide the correspondence between the color and the range of values they represent
- Don't use more than seven colors
- Bare in mind that colours can be associated with feelings. like the extra attention that the colour "red" can give or the calming effect of green
- Bare in mind that the size of the elementary spatial elements will influence the reading of the picture, giving more importance to the larger elements. So, with conventional statistical units, this kind of portrayal must not be used with absolute quantitative measures like population counts. This restriction doesn't apply to grids.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 40



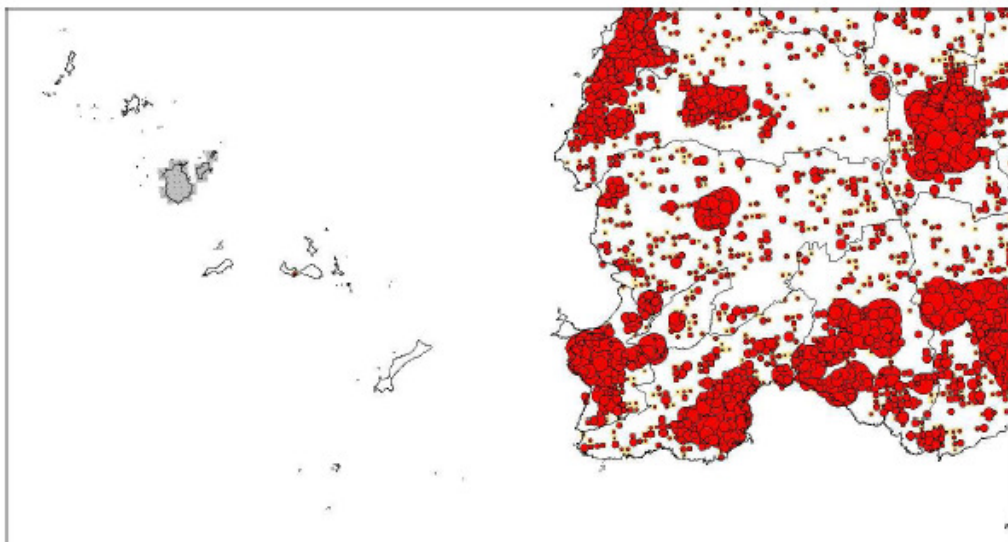
EXAMPLE 2: Viewing population distribution in Brittany, using 200m wide grid cells (partial view). Dark outlines (municipalities) are drawn only for improving the readability of the map and are not related to the population distribution object.

Since cells have the same size everywhere, the grid cells show actual population counts. But the threshold values for the colors are chosen to meet standard density thresholds in use to define the degree of urbanization: 1500 inhabitants per km² and 300 inhabitants per km² (respectively dark red and red). Colored grid cells are more useful than proportional symbols because they better show the extent of urbanized areas.

In this example, there are grid cells whose population is strictly zero and grid cells where actual population is not known. It is important to make a distinction between the two cases: first ones are left blanks, second ones are painted in grey.

Adding symbols

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 41



EXAMPLE 3: Viewing population distribution using symbols whose size is proportional to the population. The population is shown with variable size discs, each one giving the population of one elementary spatial element (they are 200m wide grid cells as in example 2). In this case the population distribution is an additional layer of a map containing something else: here a map of the boundaries of the municipalities. This kind of presentation is useful when different information have to be represented separately:

- Size and color can be chosen to represent different things: for example size for population and color for the intensity of young people in it.
- Also there could be different sets of symbols: there could be symbols for population at residence place and symbols for population at work place.
- One could also use one layer with colored vector statistical units and one layer with symbols. This is also a good solution to depict absolute quantitative data when using conventional statistical units. Since this kind of presentation doesn't directly refer to the actual shape of the statistical units the reader is much less influenced by the differences in area of the underlying statistical units. One should notice that using symbols doesn't remove the question of the presentation of unknown data. In the example it was represented as cells painted in grey like in the previous example.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 42

Bibliography

[DS-D2.3] INSPIRE DS-D2.3, Definition of Annex Themes and Scope, v3.0,
http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.3_Definition_of_Annex_Themes_and_scope_v3.0.pdf

[DS-D2.5] INSPIRE DS-D2.5, Generic Conceptual Model, v3.1,
http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.5_v3.1.pdf

[DS-D2.6] INSPIRE DS-D2.6, Methodology for the development of data specifications, v3.0,
http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.6_v3.0.pdf

[DS-D2.7] INSPIRE DS-D2.7, Guidelines for the encoding of spatial data, v3.0,
http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.7_v3.0.pdf

[ISO 19101] EN ISO 19101:2005 Geographic information – Reference model (ISO 19101:2002)

[ISO 19103] ISO/TS 19103:2005, Geographic information – Conceptual schema language

[ISO 19107] EN ISO 19107:2005, Geographic information – Spatial schema (ISO 19107:2003)

[ISO 19115] EN ISO 19115:2005, Geographic information – Metadata (ISO 19115:2003)

[ISO 19118] EN ISO 19118:2006, Geographic information – Encoding (ISO 19118:2005)

[ISO 19139] ISO/TS 19139:2007, Geographic information – Metadata – XML schema implementation

Eurostat Regional Yearbook 2009

(http://epp.eurostat.ec.europa.eu/portal/page/portal/publications/regional_yearbook)

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 43

Annex B (normative)

Abstract Test Suite

Any dataset conforming to this INSPIRE data specification shall meet all requirements specified in this document.

Open issue 11: Conformance testing is still an open issue under discussion.

Instructions on conformance testing and a common abstract test suite (including detailed instructions on how to test specific requirements) will be added at a later stage.

INSPIRE	Reference: D2.8.III.10_v2.0.1		
TWG-PD	Data Specification on <i>Population Distribution - Demography</i>	2011-07-13	Page 44

Annex A **(informative)** **Use cases**

No use cases have been formalised for Population Distribution theme. The basic use case is to answer to the questions:

- Where are human populations located?
- How a place is populated?

Many other INSPIRE themes include formalised use cases on population distribution.