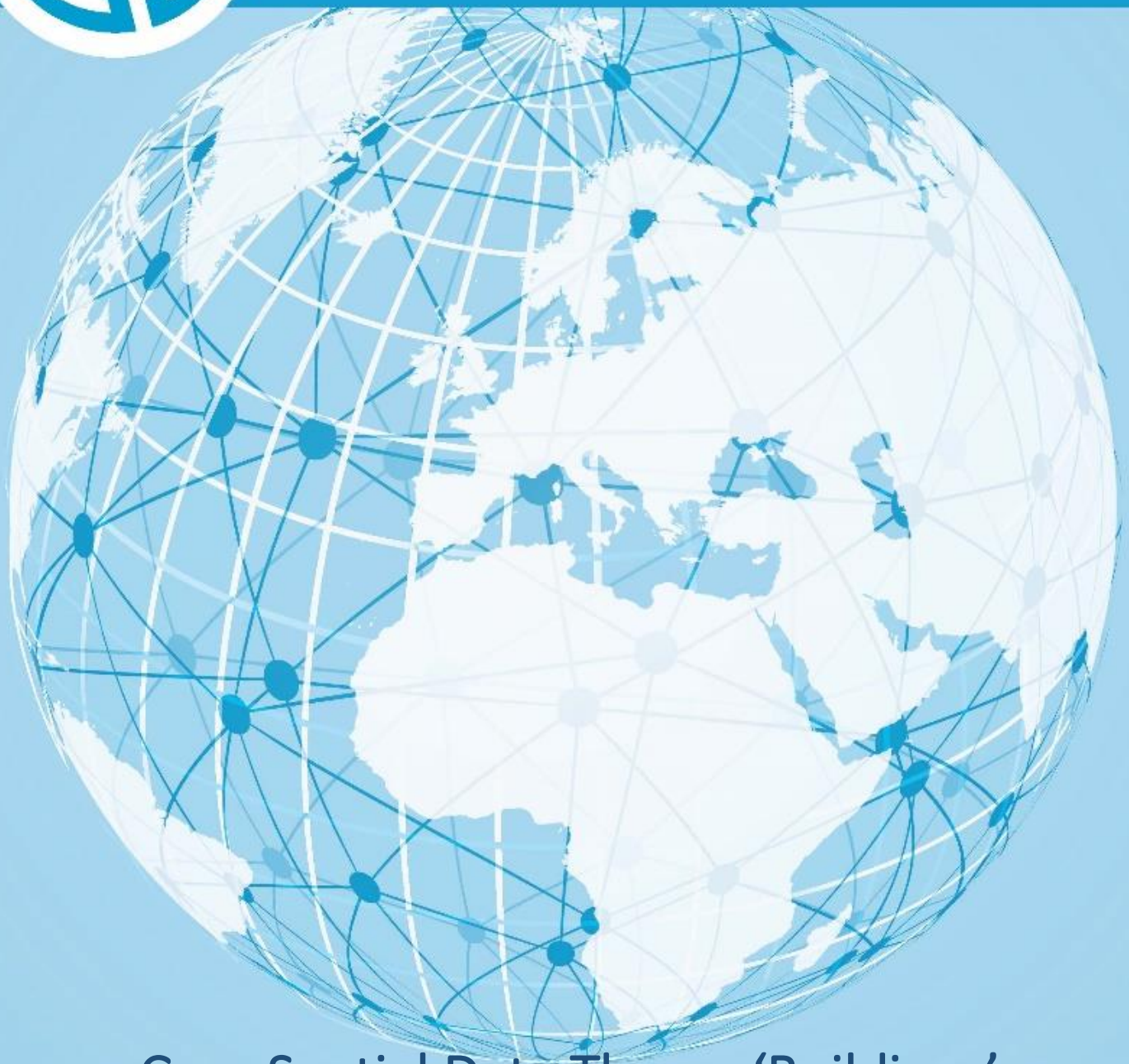
UN-GGIM EUROPE

UNITED NATIONS COMMITEE OF EXPERTS ON GLOBAL GEOSPATIAL INFORMATION MANAGEMENT



Core Spatial Data Theme 'Buildings' Recommendation for Content

Working Group A - Deliverable of Task 1.b

Version 1.1 - 2022-08-17

Version History

Version number	Date	Modified by	Comments
1.0	2018-06-01	WG A	Consolidated draft, for review by geographical and statistical community
1.1	2022.08.17	WG A	Final version, taking into account the comments from the geographical and statistical community

Content

1 Executive Summary

As the United Nations (U.N.) Millennium Development Goals (2000) era came to a conclusion with the end of the year, the U.N. announced the 2030 Agenda for Sustainable Development in September 2015, an ambitious, integrated, indivisible and transformational global agenda with 17 Sustainable Development Goals, 169 associated targets and 230 indicators promising to achieve sustainable development in its three dimensions – economic, social and environmental – in a balanced way. Geospatial data supports measuring, achieving and monitoring several if not all goals and targets set by the 2030 Agenda. The 2030 Agenda mentions the need for new data acquisition and integration approaches to improve the availability, quality, timeliness and disaggregation of data. Goal 17 explicitly emphasizes the need for developing capacities and partnerships. In this context the success of Agenda 2030 depends on senior administrators owning and leading the geospatial efforts in their respective countries.

Building on INSPIRE Directive and pertinent documentation and redirecting the focus on a cohesive Spatial Data Infrastructure without gaps in content and discrepancies in quality, stakeholders in Europe are working on geospatial standardization and increasing richness of data through Core Data Recommendation for Content that corresponds to the first phase of the WG A work program. Core Data is primarily meant for fulfilling the common user requirements related to SDGs in Member States and European Institutions.

Theme Buildings is key information for most SDG because buildings are places where people live, work and more generally spend most of their time; therefore, building data is necessary to locate accurately where people are and so to make relevant decisions in many domains. In addition, buildings are topographic objects whose physical characteristics are of strong interest for many SDG related use cases (vulnerability to risk, energy performance …).

The core recommendations include a 2D geometric representation at large scale (accuracy better than 2 metres) and a few set of key attributes: in addition to provide a unique and persistent identifier, it is strongly recommended to capture the height and number of floors above ground, the building classification (current use and nature, i.e. physical aspect), the date of construction and the number of dwellings.

The theme Buildings includes the buildings themselves but also the other constructions of interest for the SDG, such as elevated constructions that are valuable landmarks and obstacles to aerial navigation or the constructions aiming to mitigate risks or pollution.

As second priority, it is advised to capture additional semantic attributes, such as the physical characteristics of a building (material), its underground description (height or number of floors below ground), its energy characteristics and its address. It is also advised to capture a 3D geometric representation of buildings, with description of roof shape, i.e. at Level 2 of CityGML.

Last but not least, if these various pieces of information are scattered between various data producers, these producers should coordinate in order to provide core data in a user-friendly way, e.g. through a unique access point or through a unique reference data set.

2 Foreword

2.1 Document purpose and structure

2.1.1 Purpose

This document provides the main characteristics of core data for theme Buildings with focus on the recommendation for content. This document aims to help decision makers (from governments, data producers, national coordination bodies, etc.) to define their policy regarding the improvement of existing data and production of new geospatial data. It addresses digital data.

This document has annexes containing more detailed explanations targeting the technical people who will be in charge of implementing or adapting core data recommendations (e.g. for production purpose, as source of other standards, etc.).

2.1.2 Structure

The executive summary synthesizes the main conclusions of the Working Group A (WG A) process and results to develop the recommendation for content. It is meant mainly for high level decision makers.

The foreword reminds the general context of core data, the first step achieved by WG A (i.e. selecting core data themes), and it explains the general principles set by WG A to develop the recommendations for content of core data specifications for all selected themes.

The 'recommendation for content' document itself includes four chapters:

- Overview: it provides the general scope of the theme and describes the main use cases addressed;
- Data content: it provides the main characteristics of the recommended content, such as the list of core features and attributes (for vector data), as well as data capture and quality rules;
- Other recommendations: e.g. Coordinate Reference System, Metadata, Delivery;
- Considerations for future: this chapter addresses some key trends or significant user requirements that cannot be considered as core today but that might be considered in future.

The 'recommendation for content' document is meant for medium level decision makers. It is written in natural and not too technical language.

The technical explanations included in annexes describe the relationship between the recommendation for content and the corresponding INSPIRE specification, and contain any other appropriate information useful for this theme.

2.2 Core data context

2.2.1 Rationale for core data

The following background of harmonised pan-European data was identified.[1]

Authoritative geospatial data are used to support both the implementation of public policies and the development of downstream services. Moreover, geospatial data are required to be homogenous to enable the implementation of public policies in a coherent and coordinated way among countries and at regional or global level. Likewise, significant opportunities exist if services developed by industry can be exploited without requiring country specific adaptation.

The INSPIRE Directive has set up the legal and technical framework for harmonisation of the existing data related to the themes in annexes I, II and III. INSPIRE specifications provide common data models that ensure a first step towards interoperability, however ensuring homogeneous content is outside their scope, as they contain no indication about levels of detail, very few recommendations about quality, and as most features and attributes are "voidable", i.e. to be supplied if available or derivable at reasonable cost.

This background led the UN-GGIM: Europe Regional Committee to setup in 2014 the Working Group A on Core Data to deal with core data content and quality, production issues, funding and data availability.

Recommendations for content of core data will complement INSPIRE data specifications by defining the priorities on the core content that is encouraged to be made available in Europe in order to fulfil the main user requirements that are common to many countries, with focus on the SDG related ones.

Core data availability may be ensured either through upgrading of existing data when feasible or through production of new data when necessary.

2.2.2 Core data scope

In its first phase, WG A selected core data themes according to the following criteria: core data is the geospatial data that is the most useful, either directly or indirectly, to analyse, to achieve and to monitor the Sustainable Development Goals.

Among the 34 INSPIRE data themes, 14 have been considered as core including theme Buildings.

More information about the selection process and results may be found in document "Core Data Scope - Working Group A - First Deliverable of Task 1.a - Version 1.2" on http://un-ggim-europe.org/content/wg-a-core-data

[1] Extract from the Report by the Preparatory Committee on the establishment of the UN-GGIM: Europe Regional Committee, European Commission Ref. Ares(2014)1491140 - 09/05/2014.

2.3 Document objectives and principles

2.3.1 Encouraging content availability

This deliverable provides recommendations for national governments and data producers, aiming to help them to define their priorities for enriching existing data or producing new data. This deliverable is meant mainly for data producers; however it defines the recommended result and target but not the production process.

2.3.2 Complementing INSPIRE

Core data specifications are built upon INSPIRE data specifications. On one hand, they often simplify INSPIRE by selecting core feature types and attributes and by restricting or clarifying the scope; On the other hand, they enrich INSPIRE by recommending specific levels of detail, quality rules and sometimes data model extensions. Besides, the INSPIRE common terminology is thoroughly used for naming core features and attributes.

Regarding the levels of detail, the ELF (European Location Framework) project terminology has been used. The ELF levels of detail are the following: Global, Regional, Master level 2, Master level 1, Master level 0. These terms are defined in the glossary.

Regarding delivery, core data may be supplied according to several ways. It is expected that, very often, the core data recommendations will be used to enrich and upgrade existing products. In this case, core data will be available through these improved products. Core data may also be delivered through INSPIRE conditions (specifications and services).

2.3.3 Status of core data recommendations

This document contains recommendations that are not legally binding. However, some recommendations are more important than others. This order is indicated as follow:

Core Recommendation X

It is first priority recommendation, considered as both necessary and achievable in principle. Ideally, it should encourage involved stakeholders to launch short-term actions (typically within a couple of years).

Core recommendations are usually addressing only technical aspects and are meant for the organisations in charge of producing this theme. The set of core recommendations defines the basic expectations on core data.

Good Practice X

It is second priority recommendation; if adopted, it will provide significant added value to core data; it indicates a relevant trend to be adopted as much as possible. It encourages involved stakeholders to take these recommendations into account in long term, if not possible in short term.

NOTE: some of these good practices may be quite easy to achieve and are already effective in some countries whereas some others may be more difficult to achieve. This is typically the case when these good practice recommendations involve other stakeholders in addition to the organisations in charge of producing this theme, and when they address not only technical aspects but also legal or organisational ones.

A “core data set” should contain the minimum data defined by the core recommendations (and ideally also by the good practices) of this deliverable but may of course contain more and/or better information.

2.4 Abbreviations

BU	Buildings
CRS	Coordinate Reference System
ELF	European Location Framework
GML	Geographic Markup Language
LOD	Level of Detail
LC	Land Cover
LU	Land Use
PD	Population Distribution
SDG	Sustainable Development Goal
UN-GGIM	United Nations initiative on Global Geospatial Information Management
WG A	(UN-GGIM: Europe) Working Group on Core data

2.5 Glossary

Master level 0	Level of detail defined by ELF: data to be used generally at scales larger than 1: 5 000; typically, data at cadastral map level, for local level actions.
Master level 1	Level of detail defined by ELF: data to be used generally at scales between 1: 5 000 and 1: 25 000; data for local level actions.
Master level 2	Level of detail defined by ELF: data to be used generally at scales between 1: 25 000 and 1: 100 000); data for regional (sub-national) actions.
Regional	Level of detail defined by ELF: data to be used generally at scales between 1: 100 000 and 1: 500 000; data for national or regional (European or cross-border) actions.
Global	Level of detail defined by ELF: data to be used generally at scales between 1: 500 000 and 1: 1 000 000, i.e. mainly at international level

2.6 Reference documents

INSPIRE Data Specification on Buildings – Technical Guidelines 3.1:
http://inspire.ec.europa.eu/id/document/tg/bu

Eurostat classification of Buildings:
http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Classification_of_types_of_construction_(CC)

3 Overview

3.1 General scope

Definition: Considered as under scope of the theme *Buildings* are constructions above and/or underground, intended or used for the shelter of humans, animals, things, the production of economic goods or the delivery of services that refer to any structure permanently constructed or erected on its site [from INSPIRE Data Specifications on Buildings].

NOTE 1: The scope of Core data theme Buildings is globally the same as the one of INSPIRE theme Buildings. More detailed comparison with INSPIRE is available in the annex A of this document.

NOTE 2: The definition of theme Buildings is quite generic and so includes a wide range of constructions. However, not all those constructions should be considered as core data. More detailed guidelines are provided in clause 4.4 about data capture.

3.2 Use cases

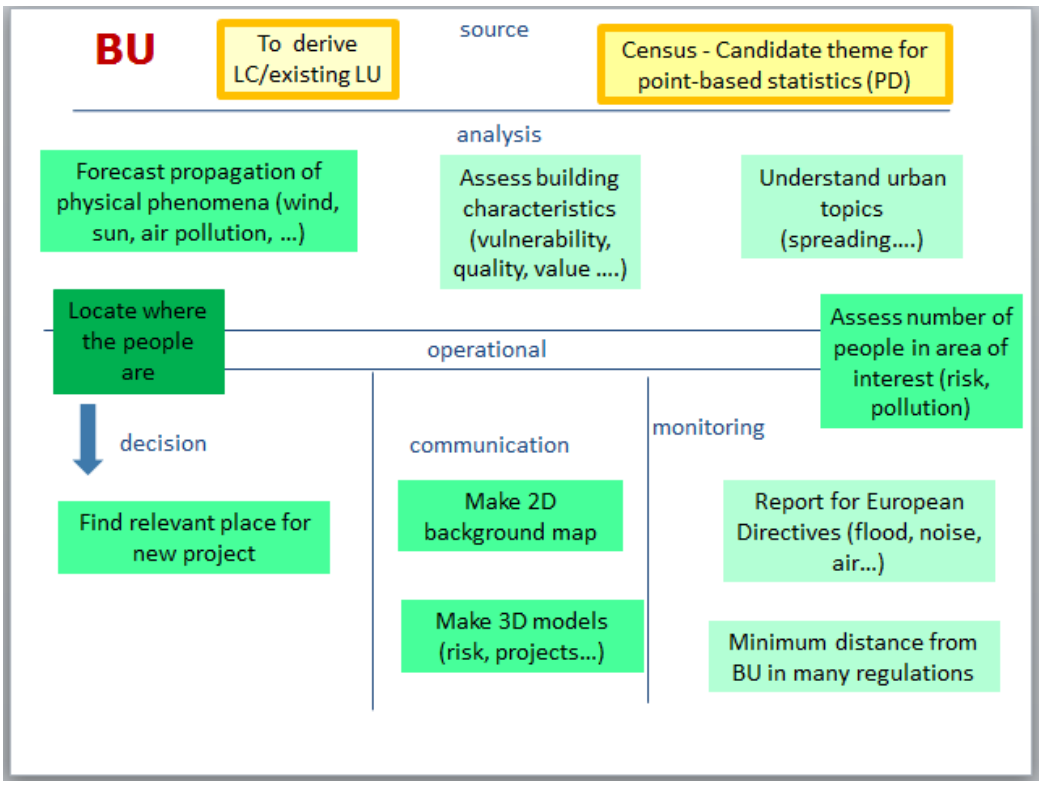


Figure 1: map of use cases for theme Buildings

Buildings are, above all, the place where people live, work and spend much of their time. Knowledge about location of buildings may be useful indirectly, as source or ancillary data to derive land use or land cover information enabling various studies at medium scales, e.g. to better understand and forecast city spreading, to assess the requirements for basic services (water, energy, schools etc.), to assess the human pressure on environment etc. Location of buildings is of course quite necessary to carry out the same kind of studies but at most detailed scales. In addition, there are many regulations aiming to protect people that prohibit some activities within a given distance of the

buildings where they live. Building data (with height and use) enable transfer of population data from administrative or statistical units to any area of interest. This is key information to assess the number of persons submitted to risk or pollution and to report it for European Directives (noise, flood, air etc.) or to compute the SDG indicators. This is also necessary to choose relevant location of new equipment, in order to ensure that it serves the citizens as much as possible (school, hospital, public transport stop etc.) or that it disturbs them as little as possible (waste landfill, water treatment plant etc.).

Building data is also required by the statistical community in order to conduct the census surveys and may be part of basic infrastructure for point-based statistics; building data is already in this position in many statistical organisations. It is expected to become a component of the Global Statistical Geospatial Framework. Building data may give location information also for people, businesses or other statistical basic units.

In addition, buildings are 3D topographic objects and, as such, may influence the propagation of physical phenomena. Typically, air pollution will not spread in the same way in a street lined with high buildings as in a street lined with small houses. Building data is also quite necessary to make visibility or inter-visibility maps, to understand the urban heat island phenomenon (and so to try to avoid it when designing new building projects), to forecast how noise will propagate etc. At the same time, the physical phenomena may also impact the building. The most obvious use case is the assessment of the vulnerability of buildings to various kinds of risks (earthquake, fire, flood etc.), according to the physical characteristics of the building. But it is also of significant interest to assess the ability of the building to SDG related improvements, e.g. can the roof host solar panels? Should the building be isolated from noise or from heat losses? Etc. In addition, buildings are valuable economic assets and for some of them, part of historic patrimony; that should be taken into account in risk management.

Regarding communication purposes, buildings are part of most of large scale maps and of 3D models.

4 Data content

4.1 Features types and attributes

Core Recommendation 1
Core data should include the feature type Building with following attributes:
- Geometry (as surface or multi-surface)
- Unique and persistent identifier
- Height above ground
- Number of floors above ground
- Current use
- Building nature (for specific buildings)
- Date of construction
- Number of dwellings

NOTE 1: Geometry has to be supplied as surface (or multi-surface); this implies that core data on Buildings will be 2D or 2.5D data.

NOTE 2: In the INSPIRE data model, the above attributes may be carried either by the Building or by the BuildingParts composing the Building. This alternative solution is considered as quite conform to the "Recommendation for Content", as far as the above attributes are provided. In other words, a data producer may use (or not) the concept of BuildingPart.

NOTE 3: The attribute "building nature" is related to the physical aspect of buildings and the recommendation to capture it applies only for specific, noticeable buildings (e.g. church, castle, greenhouse, stadium).

Core Recommendation 2
Core data should include the feature type OtherConstruction with following attributes:
- Geometry
- Unique and persistent identifier
- Nature
- Height above ground (at least for elevated constructions)

NOTE 1: The definition and description of feature types "Building" and "OtherConstruction" is provided in the INSPIRE data specifications. In summary, the main difference is that a building has to be "enclosed" whereas this condition does not apply to other construction.

NOTE 2: elevated constructions are obstacles to aerial navigation; this is why the "height above ground" is significant information.

Good Practice 1
In addition to core attributes, it is recommended to provide also for feature type Building, the following information:
- Physical description of building (material of structure, of roof, of façade, roof shape)
- Description of underground (height and/or number of floors below ground)
- Energy characteristics (energy performance, heating source and system)
- Linkage mechanisms (address, external reference to other information systems or to documents)

NOTE 1: there are two possible ways to provide the additional information of interest (physical description, underground description, energy characteristics), either directly in the same data set as the core data or by using the linkage mechanisms to another data set. The first solution is the most user-friendly and is easily achievable if all the necessary building data is managed by a single data producer. However, in some countries, data on Building is scattered between various data producers; in these cases, the second solution, though less user-friendly, might be more achievable and constitute an acceptable alternative.

NOTE 2: documents of interest include building permits, and especially for significant buildings, evacuation plans.

Good Practice 2
It is recommended to capture also buildings as 3D data, using at least the LOD2 of CityGML.

NOTE 1: The LOD2 of CityGML consists in the description of buildings as solids, with representation of walls, roof and possibly ground surface.

NOTE 2: 3D data at CityGML LOD2, i.e. with roof shapes, are necessary for some SDG related use cases (e.g. assessing solar potential of a building) or provide quite better results than CityGML LOD1, i.e. buildings as “shoe box” that might be derived from 2D data and the “height above ground” attribute for many of other SDG related use cases (e.g. physical modelling, 3D models for decision making when infrastructure project).

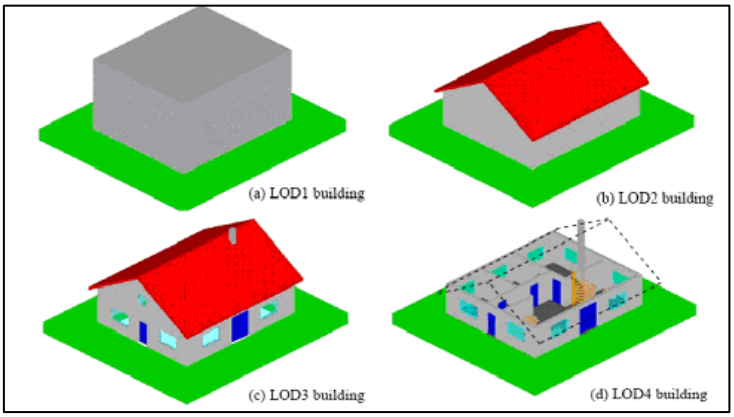


Figure 2: the 4 LOD (levels of detail) of CityGML

NOTE 3: Until now, 3D data has been mainly captured in and by big cities. However, as the user requirements are not limited to main cities, it is advised to envisage national coverage of 3D data for buildings.

4.1.1 Temporal aspects

Core Recommendation 3
Current, valid features are considered as core data.

NOTE 1: features of the past (ruin, demolished) or features of the future (projected, under construction) are not considered as core. However, in future, it would be a good practice to create the first version of a spatial data object of the building in the step of a construction process where the building is still in the status of projected.

NOTE 2: a data producer may of course also capture features of the past or features of the future as additional data; in these cases, it is advised to document the attribute “condition of construction” in order to make distinction between past, current and future features.

However, once features have been captured, it is recommended to keep them in the data base, even after their end in the real world.

Good Practice 3
It is recommended to manage the history of features, using the mechanism provided by the INSPIRE data specifications: versioning and life-cycle attributes.

NOTE 1: the mechanisms provided by the INSPIRE data specifications, namely versioning and temporal life-cycle attributes, enable the management of evolutions in the database and the delivery of change-only updates or of data at a given date in the past.

4.2 Levels of detail

Core Recommendation 4
Core data on Buildings should be captured at large scale (master level 1 or master level 0).

NOTE 1: Building data is often captured by mapping agencies (at master level 1 – around 10K) and/or by cadastral agencies or local authorities at master level 0 (better than 5K). The above recommendation aims to ensure national coverage of building data at a scale corresponding to master level 1 or at better scale (master level 0).

4.3 Geographical extent

Core Recommendation 5
Core data on Buildings should be captured on whole national territory.

NOTE 1: this recommendation implies that building data should be captured on land (where most buildings are located) but also on water (e.g. it may be the case of oil research platforms, wind turbines).

4.4 Data capture

4.4.1 Selection of core buildings and other constructions

The definition of theme Buildings is quite generic and so includes a wide range of constructions. However, only the buildings and constructions that "shall be in INSPIRE" or that "should be in INSPIRE" are considered as core data.

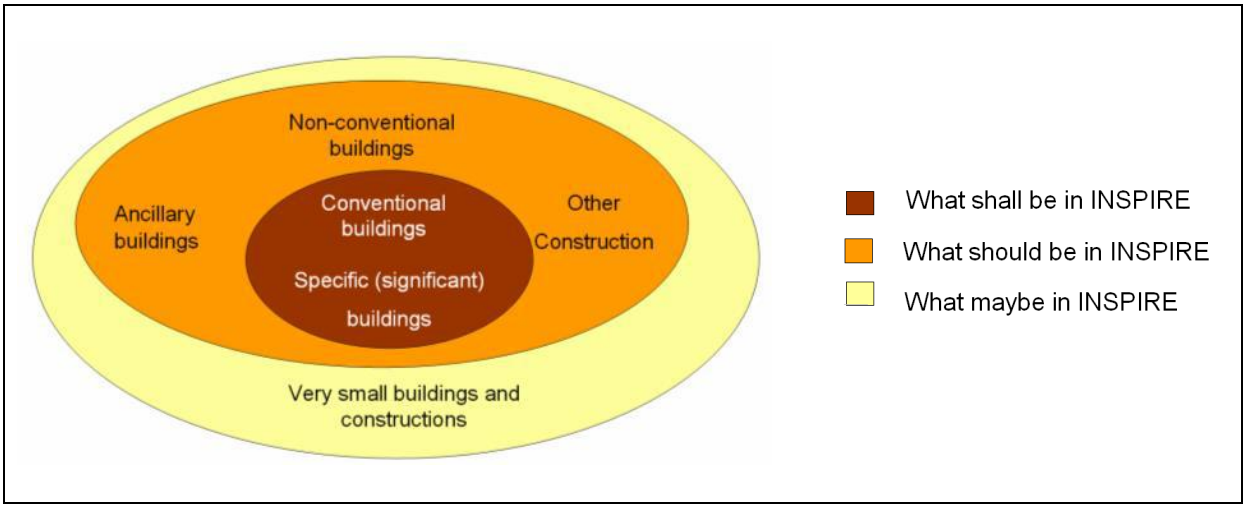


Figure 3: the modular scope of INSPIRE

The various types of buildings and constructions mentioned in the above figure are defined in clause "2.2.2.2 – modular scope" of INSPIRE Data Specifications of Buildings and described with more details in clause "9.5 - Scope of theme Buildings".

The other constructions of interest are the elevated constructions (obstacles to aerial navigation, landmarks), the constructions aiming to mitigate risk or pollution, bridges and tunnels.

Core Recommendation 6
All the buildings and constructions that shall or should be in INSPIRE are considered as core data and should be captured.

NOTE 1: as a consequence, the very small buildings and construction that "may be in INSPIRE" are not considered as core data. Of course, they may nevertheless be captured if there is national or local interest to do so.

NOTE 2: in addition to the buildings and constructions listed in INSPIRE theme Buildings, it is advised to capture the most significant man-made objects of INSPIRE theme Hydrography (e.g. embankments, dams, aqueducts, shoreline constructions).

4.4.2 Segmentation of buildings

There may be different views on Buildings: except in case of very simple architecture, it is not obvious to decide if contiguous constructions constitute a single Building, several Buildings or a Building with several BuildingParts.

More detailed explanations may be found in clause "9.6 Use of Building and BuildingPart" of INSPIRE Data Specifications of Buildings.

It is up to data producer to decide of the relevant segmentation of buildings (and possibly of building parts). As general guideline, for core data, it is advised to have reasonable segmentation:

- on the one hand, the generalized view (only one Building) should be avoided if it implies a significant lack of quality on core information;

- on the other hand, "over-segmentation" conducting to very small buildings or building parts may be costly (increasing capture efforts and data volume) with limited interest for the SDG related use cases. However, it may be required by specific national or local use cases.

4.4.3 Building geometry

The INSPIRE Buildings2D model offers the possibility of multi-representation of Buildings: a building may have several geometries that may be point, surface or multi-surface. The geometry may be 2D or 2.5D.

Core Recommendation 7
Regarding core data, the minimum content for Building geometry consists in a single geometry, as surface (or multi-surface) and as 2D data.

NOTE 1: the above recommendation is defining expected minimum content for core data; it is of course possible to capture data with more details (e.g. capture Z coordinate or more geometries).

NOTE 2: the INSPIRE data model offers several values for the horizontalGeometryReference. Regarding core data, for simple architecture buildings, the most relevant values are the footprint (most adapted to many use cases) or the roof edge (easy to capture by stereo-plotting, required by some use cases) whereas for complex architecture buildings, the other values may be more relevant (see INSPIRE Data Specifications clause 9.7 Geometric Representation).

NOTE 3: the INSPIRE Buildings2D data model offers the mechanisms to document the horizontal (and possibly vertical) geometry reference and the horizontal (and possibly vertical) estimated accuracy at

feature level. For core data, these mechanisms should be used if the information is heterogeneous within the data set; if the geometry reference and/or estimated accuracy are homogeneous, they may be captured only at dataset level.

4.4.4 Height above ground

The INSPIRE Buildings model offers the possibility of multi-representation of the attribute heightAboveGround, with documenting the low and high reference (see clause 5.3.1.1.3 of INSPIRE Data Specification).

Regarding core data, several values may be of interest for the high reference:

- Roof eave (or roof edge): in most cases, it allows to make a coarse assessment of number of floors above ground; **if the attribute "number of floors" cannot be captured, it is strongly recommended to capture the building height above ground using the roof eave** (or if not possible, the roof edge) as high reference;
- General roof: it is relevant to capture the height of the building, using for instance the half roof or 2/3 roof as high reference in order to derive 3D data as LOD1 of CityGML from the 2D geometry and the building height; this provides more realistic results than using the roof eave or roof edge; this solution is advised when there is no other true 3D data, i.e. no data in LOD2 or better of CityGML.

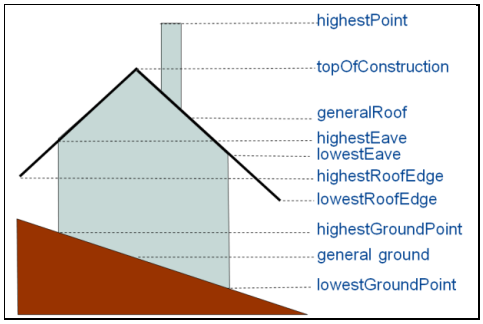


Figure 4: Examples of low and high references for height above ground (from INSPIRE)

NOTE 3: the INSPIRE Buildings data model offers the mechanisms to provide metadata information about the high and low reference and the status of the attribute "heightAboveGround". For core data, these mechanisms should be used if several values of the height are supplied or if the metadata information is heterogeneous and varies according features. In other cases, the metadata information may be captured only at dataset level.

4.4.5 Date of construction

The INSPIRE Buildings model offers the possibility to provide the attribute Date of construction as a date or as an interval (beginning, end).

Regarding core data, date of construction is considered as very valuable information:

- It enables to go into the past and to make temporal studies (e.g. to understand urban spreading in a given area);
- It gives an idea of the building "quality": an expert may guess the building physical characteristics from the construction standards or practices at a given period of time.

However, it should be recognised that capturing date of construction of existing buildings and especially the old ones may be quite challenging. Typically, it is expected that date of construction (e.g. 2015) may be provided as construction year for recent buildings and as a period for older buildings (e.g. “between 1950 and 1960” or even just “before 1900”). The general advice is to do the best with reasonable effort.

Good Practice 4
For new buildings, it is advised to register both the beginning date (building permit) and the end date (end of work) of the construction phase of buildings.

NOTE 1: The beginning date provides information on the construction standard used for the building whereas the end date provides the starting point for effective use of the building. Registering this information is of course of specific interest for the buildings whose construction phase is long lasting.

NOTE 2: An alternative solution is to ensure link with the building permit register.

4.4.6 Current use

The INSPIRE data models allows to provide several values and offers to document their percentage.

Good Practice 5
Regarding core data, it is recommended to provide at least the dominant use and, if possible, all the uses with their percentage.

In INSPIRE, the possible values of current use are defined in a hierarchical code list.

Good Practice 6
Regarding core content, it is recommended to capture the attribute “current use” using at least the highest level of the INSPIRE code list.

NOTE 1: the highest level of INSPIRE code list includes following values: residential, agriculture, industrial, commercesAndServices, ancillary.

NOTE 2: it is of course possible to capture more detailed data, e.g. using the Eurostat classification of Buildings or a national standard easily matchable with the INSPIRE code list.

4.4.7 Nature

Regarding core data, the attribute “nature” applies both to feature types Building and OtherConstruction.

Good Practice 7
Regarding core content, it is recommended to capture the attribute “nature” using as starting point the values of the INSPIRE code lists.

NOTE 1: the INSPIRE code list may be extended to provide more detailed values (e.g. more detailed description of towers as required by the the EuroControl standard for obstacles to navigation) or just new values (e.g. the most significant man-made objects of INSPIRE theme Hydrography).

NOTE 2: it may be worth to agree on a national standard taking into account the national specificities and extending the INSPIRE code lists.

4.4.8 Energy performance

At least in European Union Member States, it is mandatory to assess and provide the energy performance of a building in case of selling or renting it. This provides good opportunity to capture this data in order to make it available for SDG related use cases (mainly SDG 7 about sustainable energy).

Good Practice 8
It is recommended to capture the energy performance with its reference date at least when new building transactions (selling or renting).

NOTE 1: A more ambitious policy consisting in registering all the energy performance assessments, including those done for past transactions, would of course be better but might be more costly to achieve.

NOTE 2: The energy performance may be measured (typically in multiple-family buildings) at building unit (e.g. apartment) level. In this case, the information about energy performance should preferably be delivered either using the BuildinUnit feature type (as in INSPIRE extended schema) or, if not possible, by providing the information in a more aggregated way, by indicating the ratio (e.g. 12% of building with "D" as energy performance).

4.5 Quality

4.5.1 Completeness

Core Recommendation 8
Completeness of core data should be ensured, at least for the buildings that shall be in INSPIRE.

NOTE 1: Regarding buildings that shall be in INSPIRE, see clause 4.4.1 of this document and/or clauses 2.2.2.2 and "Data Capture - Scope of theme Buildings" of INSPIRE Data Specifications of Buildings.

NOTE 2: nevertheless, it is recognized 100 % completeness may be difficult to achieve. The aim should be to maintain completeness of 98% or more of the buildings that shall be in INSPIRE.

Good Practice 9
Completeness is recommended in the following cases:
- "Current use" for residential, agriculture, industrial and tertiary (commerce and services) buildings;
- At least one of the attributes "number of floor above ground" or "height above ground" should be captured for all buildings;
- "Nature" should be supplied for all the elevated buildings and other constructions.

NOTE: the elevated buildings and constructions are those considered as obstacles for aerial navigation by EuroControl. **For these elevated constructions, it is also strongly advised to ensure completeness of the features and of the attribute "heightAboveGround".**

4.5.2 Geometric accuracy

Core Recommendation 9
Absolute horizontal accuracy of buildings and other construction should be better than 2 meters.

In addition, the geometric accuracy should be documented, either at feature level using the attributes horizontal or vertical estimated accuracy if the accuracy is heterogeneous in the dataset or at data set level if the accuracy is homogeneous.

Good Practice 10
If buildings are captured as 2.5D or 3D data, absolute vertical accuracy should be better than 2 meters.

4.5.3 Thematic accuracy

Good Practice 11
Great care has to be taken to ensure thematic correctness of all core attributes

4.5.4 Update frequency

Core Recommendation 10
The update frequency for core data of theme Buildings should be one year or better.

NOTE: It is reminded that Building core data is 2D data (or 2.5D data).

Good Practice 12
If buildings are captured as 3D data, the update frequency of the dataset should be better than 5 years or even 3 years if possible

5 Other recommendations

5.1 Coordinate Reference System (CRS)

5.1.1 Case of 2D data

Good Practice 13
Core data should be stored and managed in a CRS based on datum ETRS89 in areas within its geographical scope, either using geographic or projected coordinates.

NOTE 1: geographical scope of ETRS-89 excludes over-sea territories, such as Canary Islands or French Guyana or Madeira Islands and Azores Islands. In these cases, it is recommended to use a CRS based on ITRS (International Terrestrial Reference System).

NOTE 2: storing and managing data in CRS based on international datum facilitates the import of measures from modern sensors, ensures that data is managed in a well-maintained geodetic framework and of course, facilitates the export of data into international CRS (e.g. those mandated by INSPIRE).

5.1.2 Case of 2.5D or 3D data

Good Practice 14
If buildings or other constructions are captured as 2.5D data or 3D data, it is recommended to use for the Z coordinate a gravity-related height, ideally given in EVRS as vertical component of the Coordinate Reference System.

NOTE 1: EVRS stands for European Vertical Reference System.

5.2 Metadata

Good Practice 15
Core data should be documented by metadata for discovery and evaluation, as stated in the INSPIRE Technical Guidelines for metadata and for interoperability.

NOTE: this is a legal obligation for the Member states belonging to the European Union. For the other countries, this is a way to make their data easily manageable by transnational users.

The "Recommendations for content" for theme Buildings is strongly based on the INSPIRE data specifications. The INSPIRE data model includes some metadata information at feature or even at attribute value level, such as the estimated accuracy, the geometry reference ….

It is reminded that for core data, data producers may capture this metadata information at feature level, using the INSPIRE mechanisms; this is the recommended solution as it is the most flexible. However, if metadata information is homogeneous, it may be enough to capture it at dataset level, e.g. by documenting it within the national specifications of theme Buildings.

5.3 Delivery

It is expected that core data will be made available through improved existing products (or new products) or as INSPIRE data, and perhaps as specific core products (delivery issues still have to be investigated by the working group).

Good Practice 16
Core data should be made available according to the INSPIRE Technical Guidelines for interoperability, for metadata and for services.

NOTE: this is a legal obligation for the Member states belonging to the European Union. For the other countries, this is a way to make their data easily manageable by transnational users.

Data related to theme "Buildings" may be scattered among several data providers. However, users should be given an easy access to core data.

Good Practice 17

Data providers of core theme “Buildings” should coordinate their efforts and offer easy access to core data, e.g. through a unique access point or by national reference data set(s).

NOTE 1: the unique access point is the relevant solution when several local data producers provide Building data according to a common standard or specification (implementing the core data recommendations).

NOTE 2: national reference data set(s) integrating data from various data producers may be necessary if the attributes are scattered among these various data producers. Ideally, there should be one unique national reference data set for Buildings 2D data and, if possible, one unique reference data set for Buildings 3D data.

6 Considerations for future

6.1 Data integration

This document recommends to offer easy access to core data (good practice 14), even if case of multiple data producers. However, this may be difficult to achieve typically where core and good practices attributes are carried by various Building data products. In this case, some data integration may be required to gather various products into a single national reference data set.

This document also recommends, as good practice, to include some linkage mechanisms, such as the building address or the external reference (i.e. the building identifier in another Information System). This good practice aims to make data integration or data linkage easier but once again, it is unsure it is enough to ensure automatic process. This is due to the fact that there is no a single view on buildings: the same real-world entity may be considered as various features according various stakeholders, i.e. data producers will likely use different segmentations of buildings.

Data integration (when required) will raise many issues, both political (data producers should have strong will to cooperate or government should have strong policy towards reference data) and technical (how to gather various attributes carried by various geometries on a single feature). Research and knowledge exchange should be encouraged to cope with these issues.

6.2 More detailed data

6.2.1 Interior of buildings

This document advises 2D data as core recommendation and simple 3D data (LOD2 of City GML) as good practice. However, there are user requirements for more detailed data, such as description of building interior. These requirements apply mainly to big buildings open to public, such as airports, railway stations, commerce centres.

Research and experience sharing should be encouraged to evaluate the existing data models such as LOD 4 of CityGML or BIM (Building Information Modelling) and the existing tools to produce and use 3D complex data, to develop new solutions, if required, to assess the costs and benefits of building interior data etc.

6.2.2 Specific extensions

This document includes a relatively small set of basic attributes that are expected to satisfy a wide range of user requirements. However, it should be recognized that in some cases, these core attributes will be used only as proxy of the information that would be ideally expected.

More detailed information has already been defined for specific use cases. For instance, the ELISE action Energy & Location Applications has designed an extension of the INSPPIRE model to record the input and output data used for assessing the energy demand and performance of buildings. This Use case is strongly related to SDG 7 (affordable and clean energy).

Another example is provided by City GML that is a modular standard having specific extensions for well-defined Use Cases, such as the Noise one that is strongly related to SDG 3 (good health and well-being).

In future, it should be investigated if these more specific data should be captured in a systematic way and how this capture should be organised. Activities such as cost-benefit analysis, research, feasibility studies and discussion with potential producers should be encouraged.

6.3 Less detailed data

Some users are interested by less detailed data, i.e. official generalised data with schematic building representations to be used at coarser levels of resolution.

The simplest way to get less detailed data is cartographic generalisation dealing only with the building geometry (geometry simplification, grouping of neighbour buildings); it results in building blocks without any semantic information as the purpose is just to get readable maps. These building blocks are used for digital or paper maps but are generally not provided in a separate layer to users.

It is also possible to group buildings according to some common characteristics in order to get more meaningful new features.

A simple case is when buildings belong to a building complex such as industrial plants, hospital complexes, university campus ... The capture and representation of the building complex outline and of its function enables a simplified view on the concerned area.

Another example is to group buildings and their surroundings into features sharing same geographical name, such as hamlet, village, neighbourhood or city.

More systematic generalisation process is to group buildings and their surroundings into areas sharing same land cover characteristics or land use characteristics. In these cases, the grouping may be adapted to user requirements and not just imposed by the real-world context. For instance, the land cover data may just take into account what is a built-up area or not but it may also take into consideration the density of the built-up area, using as possible criteria distance between buildings, percentage of building in the area, height of the buildings …. In a similar way, the land use data may be derived using a very simple classification (e.g. residential/activity) or a more detailed one (e.g. (residential/industrial/commercial /offices/ mixt …)

Some of this data is already more or less included in the UN-GGIM: Europe recommendations for content of core data. This is the case with theme Geographical Names that recommends to capture the real geometry of settlements and to ensure completeness of these named places. This is partly the case with theme Basic Services that includes building complexes corresponding to public services or to power plants but that is not considering commercial or other industrial complexes. Land Cover and Land Use are also considered as core theme but the related recommendations for content include only very basic classification. In addition, land cover data is rather expected to be captured from images rather than from building generalisation.

There is clearly place for progress and activities about user requirements refinement, state-of-play of existing practices, knowledge exchange and research should be encouraged. For instance, in a first step, it might be investigated if delivering generalised cartographic data would be helpful. More generally, it should be investigated and tested if the best solution consists in on-the-shelf products or in offering users processes to design their own generalised data product.

7 Annex A: Relationship with INSPIRE

7.1 Data model

The UML models provided in this annex are only graphical illustrations of the core recommendations and of the good practices present in this document.

The recommendations for content are represented by highlighted the selected attributes in the following way:

Core recommendation

Good practice

7.1.1 Comparison between Core Data and INSPIRE content

7.1.2 Feature types and attributes

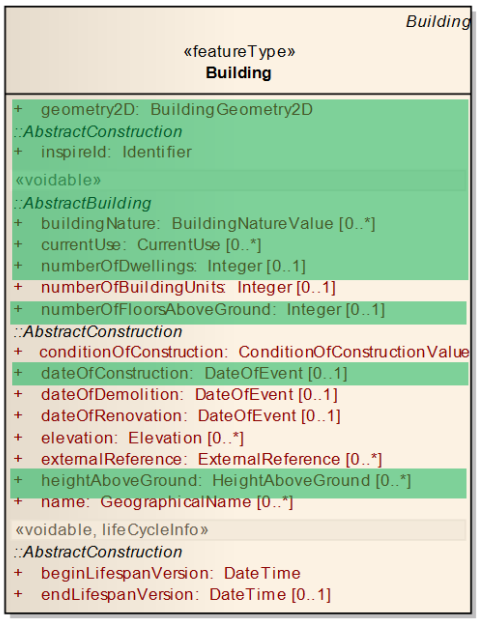
Core recommendation 1 **Core data should include the feature type Building with following attributes:** **- Geometry (as surface or multi-surface)** **- Unique and persistent identifier** **- Height above ground** **- Number of floors above ground** **- Current use** **- Nature (for specific buildings)** **- Date of construction** **- Number of dwellings**	 Extract from INSPIRE Buildings 2D

Figure 5: Comparison of core recommendation 1 with INSPIRE

Core recommendation 2

Core data should include the feature type OtherConstruction with following attributes:

- Geometry

- Unique and persistent identifier

- Nature

- Height above ground (at least for elevated constructions)

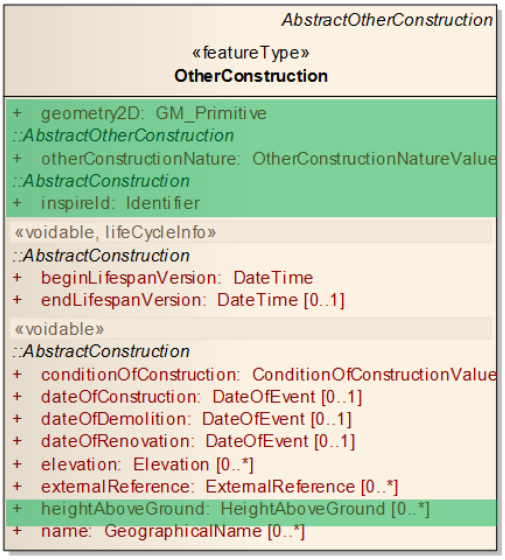


Figure 6: extract of INSPIRE Buildings Extended2D for comparison to core recommendation 2

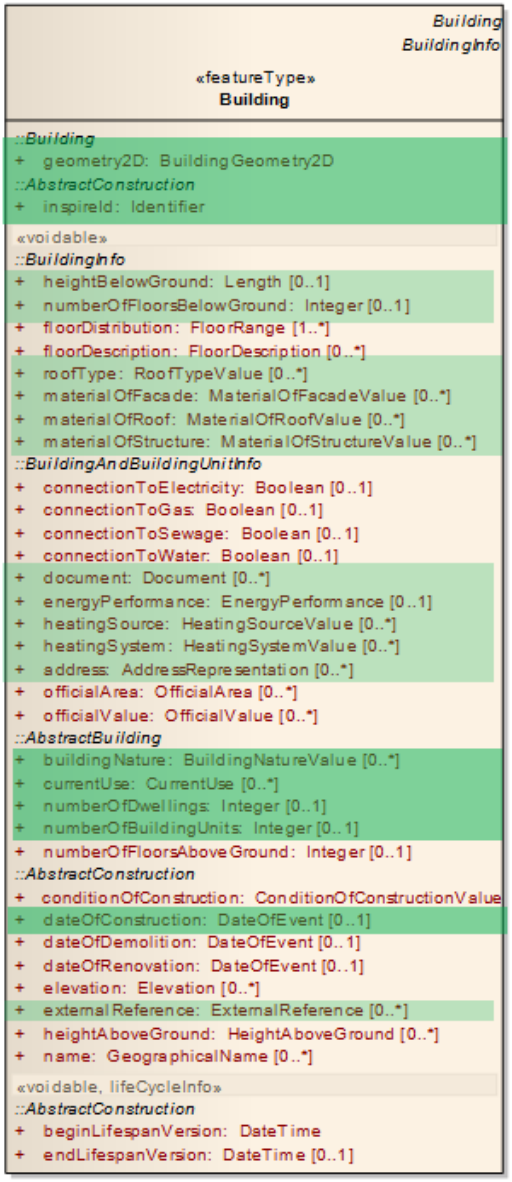
Good practice 1 In addition to core attributes, it is recommended to provide also for feature type Building, the following information: - Physical description of building (material of structure, of roof, of façade, roof shape) - Description of underground (height and/or number of floors below ground) - Energy characteristics (energy performance, heating source and system) - Linkage mechanisms (address, external reference to other information systems or to documents)	 Extract of INSPIRE Buildings Extended2D

Figure 7: Comparison of INSPIRE Buildings Extended2D with good practice 1

Good practice 2

It is recommended to capture also buildings as 3D data, using at least the LOD2 of CityGML.

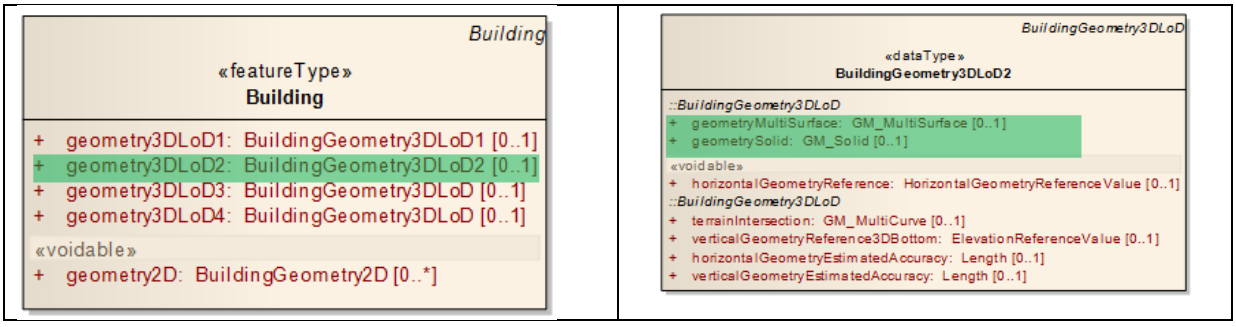


Figure 8: Extract of INSPIRE Buildings 3D

NOTE 1: The INSPIRE recommendations for 3D data geometry apply also for core data; for instance, the geometry may be defined as a Solid or as a MultiSurface, depending if the ground surface may be captured or not.

NOTE 2: the core data may be enriched by the walls, roof and ground surfaces represented in separate feature types.

NOTE 3: regarding semantic attributes, there are 2 options: either, the relevant attributes (core recommendations + possibly good practices) are carried directly the 3D building data or the 3D dataset includes for each building the “external reference” mechanism enabling to link the 3D data to the 2D core data set used as main source for the 3D data production.

7.1.3 Temporal aspects

Good practice 2 It is recommended to manage the history of features, using the mechanism provided by the INSPIRE data specifications: versioning and life-cycle attributes.

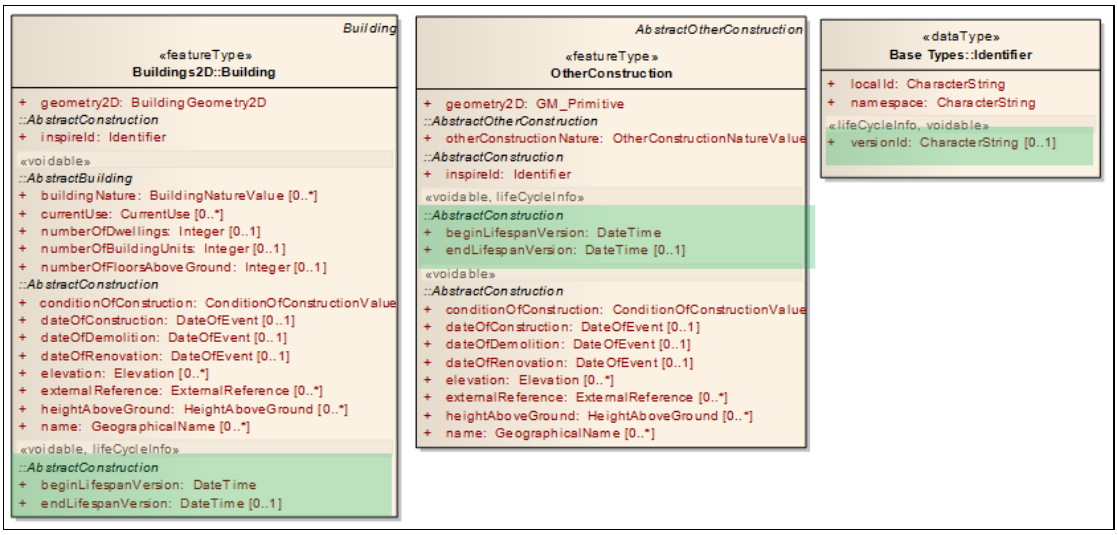


Figure 9: temporal aspects

7.1.4 Geometric representation

Core recommendation 7

Regarding core data, the minimum content for Building geometry consists in a single geometry, as surface (or multi-surface) and as 2D data.

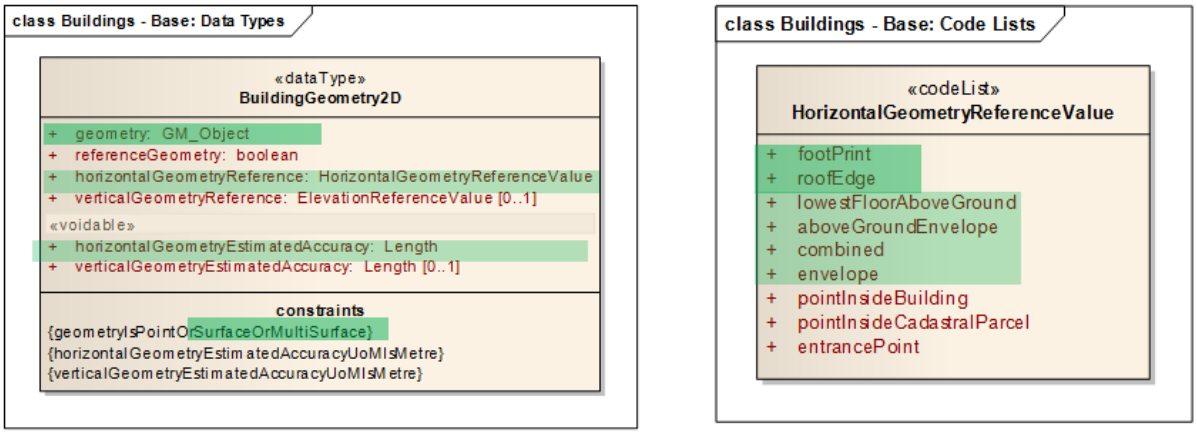


Figure 10: geometric representation of buildings

NOTE 2: the INSPIRE data model offers several values for the horizontalGeometryReference. Regarding core data, for simple architecture buildings, the most relevant values are the footprint (most adapted to many use cases) or the roof edge (easy to capture by stereo-plotting, required by some use cases) whereas for complex architecture buildings, the other values may be more relevant.

NOTE 3: if the horizontal geometry reference and/or estimated accuracy are not the same in whole data set, it is advised to have them carried at feature level, as this is expected in the INSPIRE data model.

7.1.5 Current Use

Good practice 5

Regarding core data, it is recommended to provide at least the dominant use and, if possible, all the uses with their percentage.

Good practice 6

Regarding core content, it is recommended to capture the attribute “current use” using at least the highest level of the INSPIRE code list.

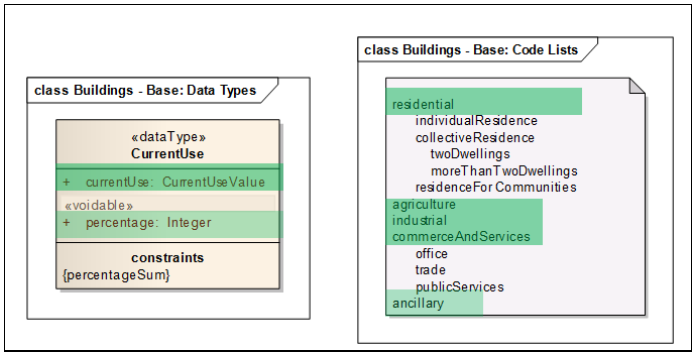


Figure 11: current use of buildings

NOTE 1: the colour coding takes into account the good practice 9 that recommends completeness for the following highest level values of INSPIRE code list: residential, agriculture, industrial, commerceAndServices. By definition, the ancillary use can’t be the dominant use of a building; it is why it is given lower priority than the other values.

7.1.6 Nature

NOTE 1: the INSPIRE code list may be extended to provide more detailed values (e.g. more detailed description of towers as required by the the EuroControl standard for obstacles to navigation) or just new values (e.g. the most significant man-made objects of INSPIRE theme Hydrography).

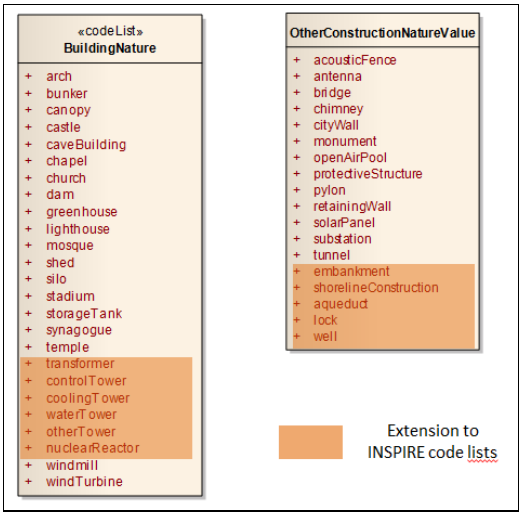


Figure 12: examples of possible extensions of INSPIRE code list for the attribute “nature”

In INSPIRE, there is some overlap between the themes Hydrography and Buildings as some hydrography related constructions are both in theme Hydrography (as man-made objects) and in theme Buildings (e.g. dams, bridges). In the core data “Recommendations for content”, the most significant hydrography related constructions are considered under theme Buildings however, in practice, the key point is to have them captured, whether in theme Buildings or in theme Hydrography.

7.1.7 EnergyPerformance

Good practice 8 recommends capturing EnergyPerformance.

NOTE 2: the energy performance may be measured (typically in multiple-family buildings) at building unit (e.g. apartment) level. In this case, the information about energy performance should preferably be delivered either using the BuildinUnit feature type (as in INSPIRE extended schema) or, if not possible, by providing the information in a more aggregated way, by indicating the ratio (e.g. 12% of building with “D” as energy performance).

Case 1: the energy performance is captured and delivered at “atomic” level, either on the Building (or Building Part) if homogeneous on whole building [it is typically the case of single-family houses] or on the Building Unit [it is typically the case of blocks of flats]. In these cases, the energy performance may be provided using the INSPIRE extended data model.

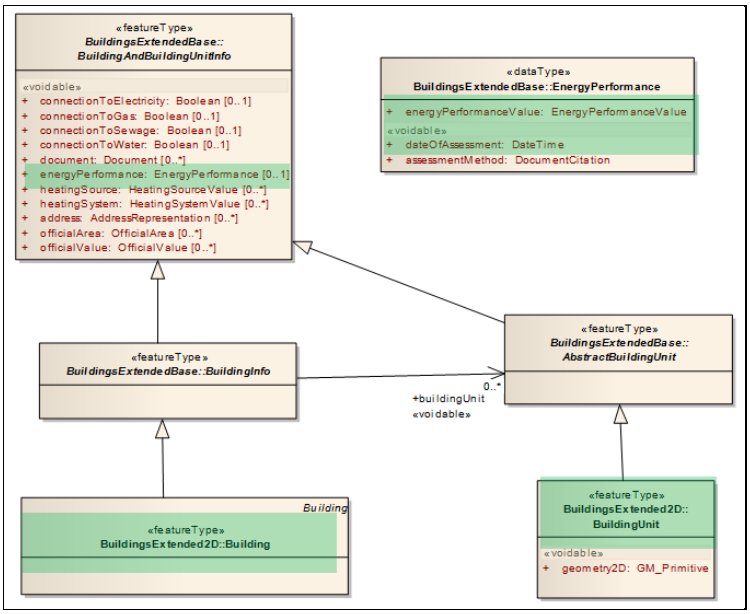


Figure 13: Energy performance is provided using the (standardised) Extended INSPIRE data model

NOTE: information may be provided as attribute of Building (if homogeneous value on whole Building) or of BuildingUnit.

Case 2: the energy performance has been captured on some of the Building Units of a Building, but the information is delivered in an aggregated way on the Building. In this case, the energy performance may be provided using the INSPIRE basic data model and making a specific extension.

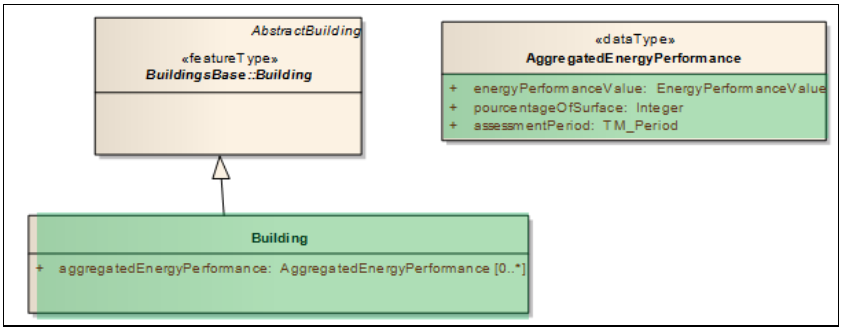


Figure 14: Energy performance is provided using an ad-hoc extension of basic INSPIRE data model

8 Annex B: Methodology

Core data specifications have been elaborated based on one hand on user requirements (with focus on the ones related to SDG) and on the other hand on INSPIRE data specifications.

In case of theme Buildings, the user requirements are also coming from INSPIRE, more accurately from the use case investigation already done by the INSPIRE Thematic Working Group. The user interviews and the use cases descriptions (INSPIRE Data specifications –annex B) have been used as starting point to select the candidate core attributes: they are those mentioned the most often.

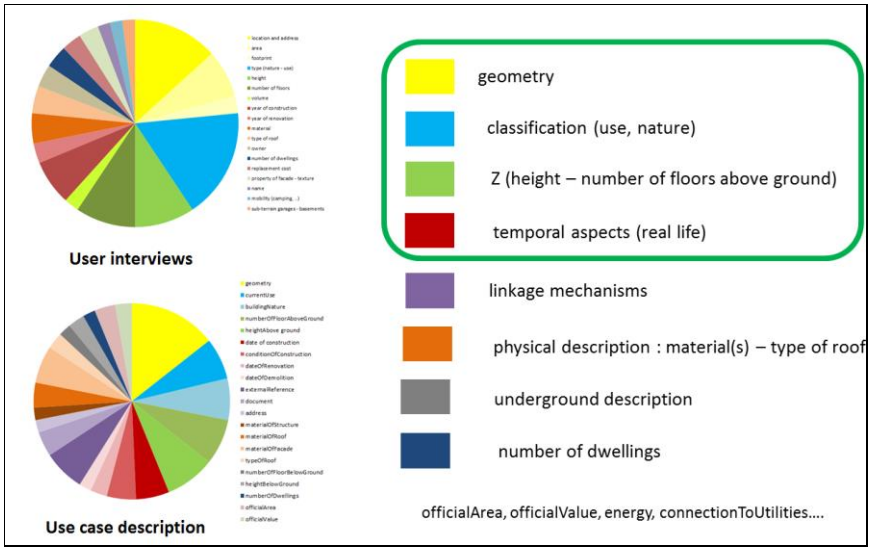


Figure 15: selection of core attributes

This initial list has been refined, through expert discussion within WGA, in order to adapt it to the context of core data that is driven by the SDG whereas INSPIRE was driven by environment policies. For instance, the “number of dwellings” has been added as core attribute because of its strong interest for the statistical community; the energy related attributes have been added as good practice attribute because of their strong interest for SDG related use cases (SDG 7 about sustainable energy).

A questionnaire to cadastral community was also sent to test some hypothesis (information about connection to networks, official identifiers on buildings) in order to assess feasibility of some candidate proposals.

A key question was about the geometric representation of buildings: INSPIRE includes both the 2D and 3D options. What about core data? Whereas it was obvious that at least 2D data should be captured as core, the decision for 3D data was far less obvious. **Production of 3D data should clearly be encouraged as it is of strong interest from many SDG related use cases** but it is still difficult to achieve at least in some countries and not first priority. This is why the capture of 3D (LOD 2 of CityGML) data has been advised but only as “good practice”.