



**UN-GGIM
EUROPE**

UNITED NATIONS
COMMITTEE OF EXPERTS ON
GLOBAL GEOSPATIAL
INFORMATION MANAGEMENT



Core Spatial Data Theme 'Address' Recommendation for Content

Working Group A - Deliverable of Task 1.b

Version 1.0 - 2017-11-10

Version History

Version number	Date	Modified by	Comments
1.0	2017.11.10	WG A	Consolidated draft, for review by geographical and statistical community

Content

- 1 Executive Summary 4
- 2 Foreword 5
 - 2.1 Document purpose and structure 5
 - 2.2 Core data context 6
 - 2.3 Document objectives and principles 7
 - 2.4 Abbreviations 8
 - 2.5 Glossary 8
 - 2.6 Reference documents 8
- 3 Overview 9
 - 3.1 General scope 9
 - 3.2 Use cases 9
- 4 Data content 10
 - 4.1 Feature types and attributes 10
 - 4.2 Levels of detail 11
 - 4.3 Geographical extent 11
 - 4.4 Data capture 11
 - 4.5 Quality 12
- 5 Other recommendations 13
 - 5.1 Coordinate Reference System (CRS) 13
 - 5.2 Metadata 13
 - 5.3 Delivery 13
- 6 Considerations for future 14
- 7 Annex A: Relationship with INSPIRE 16
 - 7.1 Data model 16
- 8 Annex B: Methodology 18

1 Executive Summary

In September 2015 the countries of the United Nations adopted the 2030 Agenda for Sustainable Development; a set of goals to end poverty, protect the planet, and ensure prosperity for all as part of a new sustainable development agenda. Each goal has specific targets to be achieved over the next 15 years. The 17 Sustainable Development Goals (SDGs) of the 2030 Agenda are supported by 169 targets and 230 indicators.

Geospatial data supports the measuring, achieving and monitoring of many of the goals and targets set by the 2030 Agenda. The 2030 Agenda demands 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 the 2030 Agenda depends on senior administrators owning and leading the geospatial efforts in their respective countries.

In Europe, building on the INSPIRE Directive redirecting the focus on a cohesive spatial data infrastructure without gaps in content and discrepancies in quality, stakeholders are working on geospatial standardization and increasing richness of data through Core Data Recommendations for Content that correspond to the first phase of WGA work program. Core data is primarily meant for fulfilling the common user requirements related to SDGs in Member States and European institutions.

Achievement of the SDGs requires knowledge of the location of people, buildings and events as well as efficient government administration. Address data underpins government administration at all levels, and is crucial for ensuring access to utility services. It also enables effective communication with citizens: informing them of policies applying to them, notifying them of events or incidents affecting them and supporting the carrying out of social surveys. In many analytical uses address data plays a crucial role in linking the records of otherwise-unrelated datasets. A geocoded address dataset enables spatial analysis which is of benefit not only to government actors but also to the private sector for purposes such as retail location planning or logistics. Thus addresses have significant economic value.

Address is one of the candidate themes for point-based statistics recommended by UN-GGIM: Europe WG B, and is required as source data for censuses.

Core Spatial Data Theme “Addresses” Recommendation for Content includes a two-step approach for its adoption by UN member states: a) Core recommendations, such as full geographic coverage of the territory, a specified minimum set of attributes for each address, a definition of ‘building’ as the core addressable unit, and quality requirements about completeness or correct spelling of address components; and b) good practices, for example allocating valid addresses to vacant plots and buildings under construction and recommending the creation of “true addresses” with street names and house numbers in areas where address locator is missing.

2 Foreword

2.1 Document purpose and structure

2.1.1 Purpose

This document provides the main characteristics of core data for theme Address 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 Address.

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’](http://un-ggim-europe.org/content/wg-a-core-data) 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 the 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 follows:

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

CRS	Coordinate Reference System
ELF	European Location Framework
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

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
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.

2.6 Reference documents

INSPIRE Data Specification on Addresses – Technical Guidelines 3.1:

http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AD_v3.1.pdf

3 Overview

3.1 General scope

Definition: Location of properties based on address identifiers, usually by road name, house number, postal code [INSPIRE Directive 2007/2/EC].

Description: In core data the basic unit of addressing is a building; a permanent construction, intended or used for the shelter of people, having at least one entrance from publicly-accessible space. This could be a house (whether or not attached to its neighbours), office, factory, retail unit, leisure centre, bar or place of worship [WG A definition].

The general scope of core data theme Addresses is narrower than that of the corresponding INSPIRE theme, because core data requirements apply only to basic units of addressing as described above. The other key difference is that core data, unlike INSPIRE data, is not constrained to what already exists.

NOTE 1: In practice governments and local jurisdictions may include a wide range of other entities within the scope of their address data; examples being moorings, agricultural buildings, electricity substations, car parks, street furniture and so on. These addressable objects, though of interest, are not considered as core because they have minor importance for the SDGs.

NOTE 2: Addresses for flats and apartments are not required as core data, as they are of lesser interest for SDG purposes but they may be included where available

More detailed comparison with INSPIRE is available in the technical explanations document.

3.2 Use cases

The most obvious use of addresses is the physical delivery of mail, but the power of address data in the digital age lies in its geocoding capability. A great deal of information is linked to addresses, and a geocoded address database allows such information to be linked to a physical location. Thus address data can add significant economic value, for example in marketing and logistics.

Address data underpins government administration at all levels, and good administration is a prerequisite for the achievement of the SDGs. It supports the provision of services and also enables effective communication with citizens: informing them of policies applying to them, notifying them of events or incidents affecting them and supporting the carrying out of social surveys.

The analysis carried out by WG-A use case focuses on the role of addresses in locating people; locating buildings; locating services; and locating events. Addresses can also act as a proxy for the density of population or buildings.

Theme 'Address' is one of the candidate themes for point-based statistics recommended by UN GGIM: Europe WG B, and is required as source data for censuses.

A systems view of the use cases reduces the wide variety of applications of address data to three broad categories:

- **Search** for addresses
- **Navigation** to address (enabled by good links; see Transport Network theme)
- **Analysis** of addresses and of other data using address as a link.

4 Data content

4.1 Feature types and attributes

Core Recommendation 1

Core data should comprise feature type Address with at least the following attributes: one two-dimensional geographic position, one locator (e.g. number or name) if available, and such other address components as are in current use. A unique and persistent identifier is also required.

NOTE 1: The terminology here is taken from the INSPIRE data specification. A **geographic position** locates the address spatially. A **locator** is a building number and/or name that distinguishes an address from its neighbours. An address is associated with a variable number of **address components**, of which there are four subclasses: administrative unit name, address area name, thoroughfare name (e.g. street, road) and postal descriptor (postal code). Address components may be used in any combination.

NOTE 2: The INSPIRE specification requires that an address should have at least one locator and one geographic position. However, it is recognised that in some countries this requirement is not currently met; for instance, in remote villages there may be no street names and no house numbers so the address is given only at village level (address area name). As a consequence, several buildings may share the same address. However, this state-of-play is raising many issues, e.g. for emergency rescue, for utility management and even for everyday life.

Good Practice 1

All basic units of addressing should be provided with a unique address enabling their unambiguous location, i.e. an address with a locator and geographic position.

NOTE: This will generally imply that authorities should create, for all core addressable objects, detailed addresses (with street names and building numbers or names) in areas where they are missing.

Good Practice 2

Geographic position should be further refined by including the geometry specification attribute, which describes the type of spatial object used to derive the position. Examples from the INSPIRE code list are: building, entrance, parcel, postal delivery point, postal descriptor and administrative unit. Wherever possible, building or entrance should be used, for reasons of precision.

4.1.1 Temporal aspects

Core Recommendation 2

Current, valid addresses are considered as core data.

Good Practice 3

Obsolete addresses should be retained in the data and their current status indicated using the INSPIRE mechanism of life-cycle attributes and versioning. Pre-allocated or provisional addresses, where available, should be managed in the same way.

4.2 Levels of detail

Core theme Addresses corresponds to Master Level 0 data.

NOTE: Levels of detail are defined by a scale range in the glossary.

4.3 Geographical extent

Core Recommendation 3

Core address data should cover the full geographic extent of the territory.

NOTE: Countries are encouraged to provide full geographical coverage in order to permit maximum use of address data to support the SDGs across the whole territory.

Good Practice 4

Core address data should also encompass isolated buildings and buildings in remote areas

4.4 Data capture

Core Recommendation 4

The basic unit of addressing is a building; a permanent construction, intended or used for the shelter of people, having at least one entrance from publicly-accessible space. This could be a house (whether or not attached to its neighbours), office, factory, retail unit, leisure centre, bar or place of worship.

NOTE: Structures ancillary to human habitation are excluded from core address data; into this category would fall garden summer houses, workshops, toilet blocks and the like.

Many buildings are composed of building units, having their own entrance from shared space ; this can include apartments and individual shops within an enclosed shopping centre or mall.

There are various ways to deal with addresses in case of multi-units buildings:

- a) only the building has a geographic address;
- b) the building and each of the building units has a geographic address;
- c) only the building units have geographic addresses.

Good Practice 5

For a multi-unit building, if there is no address at building level, the addresses of building units should be considered as core data.

NOTE 1: Good practice 5 addresses case c).

NOTE 2: In case b), where building units have been allocated unique addresses, these addresses may be added to core data sets. More generally, a core data set shall contain the minimum data in content and in quality defined by the core recommendations but may include more or better information.

Good Practice 6

Vacant building plots and sites under construction should be allocated addresses where possible.

4.5 Quality

4.5.1 Completeness

INSPIRE does not require national administrations to fill gaps in data availability, but the achievement of the SDGs depends on having adequate data for all parts of the territory.

Core Recommendation 5

Compliant addresses should be available for all buildings meeting the criteria of the §3.1 description, on the basis of continuous update.

NOTE: It is recognised that 100% completeness is not achievable, but the aim should be to maintain completeness of 95% or more of core addressable units.

4.5.2 Location accuracy

Core Recommendation 6

Co-ordinates of an address should be accurate to within 5 metres of the true position of the building centroid or entrance, where possible.

NOTE: Relatively high location accuracy is desirable to ensure that individual addresses are readily distinguished within the immediate locality. Particular care is required to locate the address on the correct side of the street.

4.5.3 Thematic accuracy

Good Practice 7

Care should be taken to ensure that address components are recorded consistently, especially in respect of the spelling of address components.

NOTE: Geocoding gives good results only if there is common semantics (i.e. mainly same spelling of street names and address area names) between the file to be geocoded and the geographic address dataset(s). To achieve this it is best to have a single register of addresses, ideally at national level, that may be used as reference data by all public bodies and other stakeholders to capture or to check the addresses of their features of interest (e.g. business register, person register, taxpayer register, customer register, etc.)

5 Other recommendations

5.1 Coordinate Reference System (CRS)

5.1.1 Case of 2D data

Good Practice 8

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

Only two-dimensional positions are required as core data; the third dimension may be represented if required as a locator, e.g. Floor 3.

5.2 Metadata

Good Practice 9

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.

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 10

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.

Good Practice 11

It may be helpful to provide the address semantics in the form of a simple text string as an additional attribute. Free text addresses are simple, readily accessible and human-readable within the local context.

6 Considerations for future

6.1.1 Historical addresses

There may be a requirement to maintain addresses that have been superseded. Typically these may be required to check identity documents or to geocode business registers; they may also have been used to define regulated areas in legal documents.

Capturing addresses from the past raises many issues: is it achievable, and how?; how far back in the past?; are the benefits worth the costs? It may be of interest in future to share experiences and to encourage research activities on these topics.

6.1.2 Linked data

In order to maximise the usefulness of core address data, it may be advisable to publish it as linked data. However, as this technology is still relatively new, more experience and more feedback on costs and benefits of such practice would be useful to support a potential future recommendation.

6.1.3 A common unique and persistent identifier

This deliverable recommends a unique and persistent identifier for each address feature, i.e. an identifier unique in the geographic Address database.

The GEOSTAT 2 project (from the statistical community) is more demanding as it recommends a unique identifier that would be common to all address stakeholders, or at least to the National Statistical Office and to the geographic Address data producer. This common unique and persistent identifier would enable more reliable geocoding than the simple semantic matching. The principle would be to have a unique identifier of the real-world object.

Though attractive, this may be quite difficult to achieve in practice: from the organisational point of view (how to persuade the various address stakeholders to work together?); from the technical point of view (how to construct such common identifiers? which common life-cycle rules?); and from the financial point of view (what are the costs? what are the benefits?)

Research activities and knowledge exchange should be encouraged on this topic to get better understanding of the current state of the art, and the advantages and difficulties of such a proposal.

6.1.4 Addresses in complex infrastructures

The “recommendation for content” for core data theme Address reflects the current address system, mainly based on street names and house numbers. Core data is expected to be efficient to fulfil the use case requirements for “simple” buildings such as single house or blocks of flats. However, the current address system may be not enough to locate a feature of interest in complex built infrastructures, such as airport, railway station, shopping centres.

An OGC Working Group is dealing with Mobile Location Service, trying to define relevant address or index system for localisation in such complex infrastructures.

The conclusions of this OGC working group may have to be considered to define the core data of theme Address in a few years.

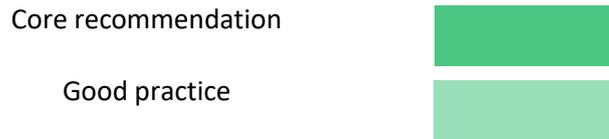
7 Annex A: Relationship with INSPIRE

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.

For theme Address the UML model is based in the INSPIRE conceptual model, without extension.

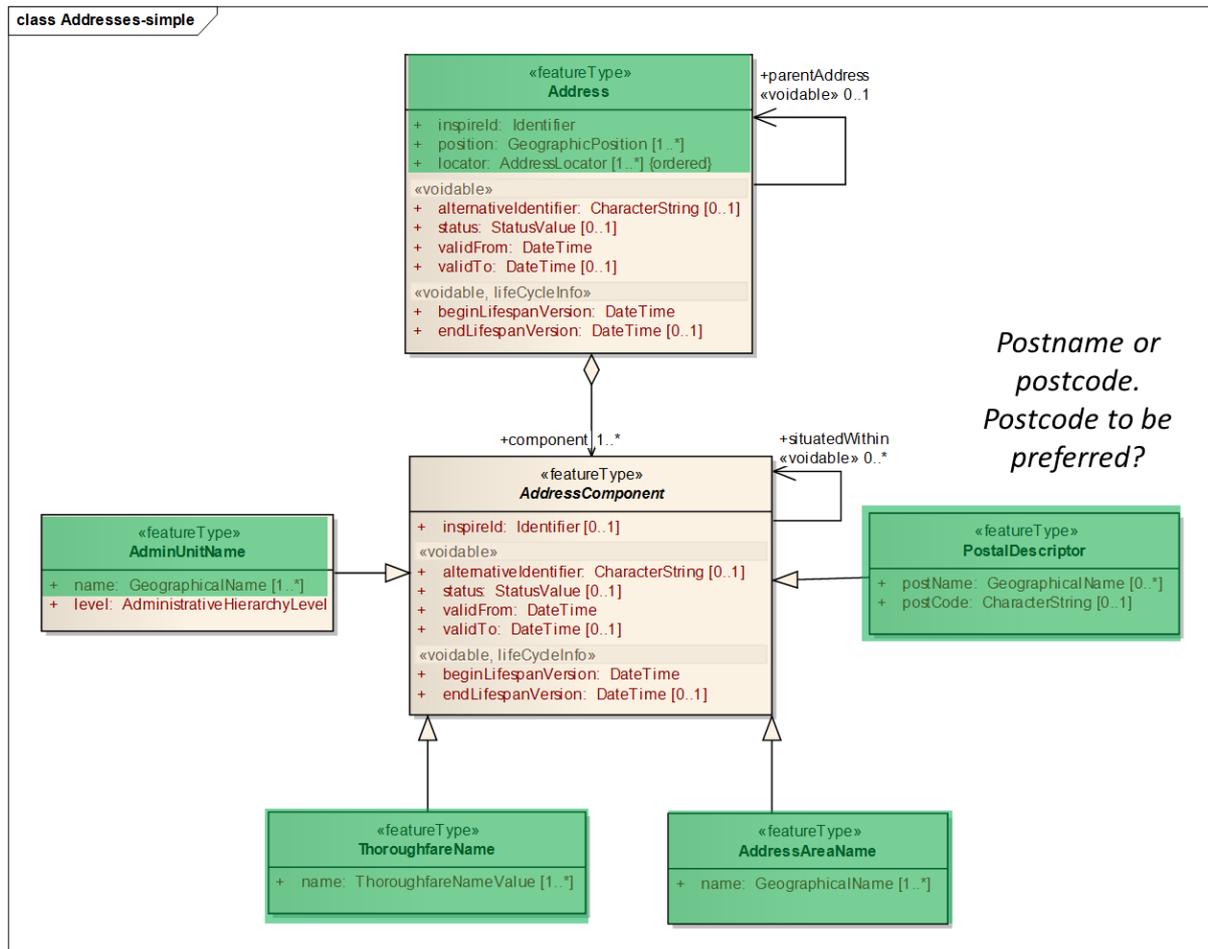
The recommendations for content are represented by highlighted the selected attributes in the following way:



7.1.1 Comparison between Core Data and INSPIRE content

Core Recommendation 1

Core data should comprise feature type Address with at least the following attributes: one two dimensional geographic position, one locator (e.g. number or name) if available, and such other address components as are in current use. A unique and persistent identifier is also required.



NOTE 1: Each of the four subclasses of address component is voidable, as not all are relevant to all addresses, but sufficient components are needed to provide unambiguous identification at country level.

NOTE 2: in INSPIRE, there is the association between parent and child addresses, the parent address being the main address and the child address being a sub-address.

In summary, in case of multi-unit buildings, core addresses are the parent addresses whereas sub-addresses are optional.

- If parent address at building level, the building parent address is core and the children addresses at building unit level are optional
- If addresses only at building unit level, these addresses are the main addresses and should be included as core.

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.

The work began with an analysis of the INSPIRE data specifications and a comparison between the INSPIRE context (focus on environmental policy and cross-border interoperability) and the core data context (focus on the Sustainable Development Goals).

This work was supplemented by examination of ISO 19160 Part 1 and the ISA Core Vocabulary for addresses, as well as study of the work of the Universal Postal Union.

Questions were added to a survey of the members of EuroGeographics' Cadastre and Land Registry Knowledge Exchange Network, to ascertain the current state of play in the countries of Europe.

Draft proposals were shared with WG A members at workshops in June and November 2016, leading to further development and refinement of the proposals.