NIVA – NEW IACS VISION IN ACTION

WP3 Harmonisation and interoperability

D3.2 Common Semantic Model

Deliverable Lead: IGN

Deliverable due date: 30/11/2022

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 842009.

Document Control Page

Document Control Page	
Title	D3.2 Common Semantic Model
Creator	Dominique Laurent (IGN)
Description	Third and last version of the Common Semantic Model, capitalising the NIVA experience.
Publisher	"NIVA - New IACS Vision in Action" Consortium
Contributors	Jesus Estrada, Anej Viskovic, Kadri Pärnpuu, Finn Moller Andreassen, Ike Van Randen, Rémi Pas
Date of delivery	30/11/2022
Type	Text
Language	EN-GB
Rights	Copyright "NIVA - New IACS Vision in Action"
Audience	☒ Public ☐ Confidential ☐ Classified
Status	☒ In Progress ☐ For Review x For Approval ☐ Approved

Revision History			
Version	Date	Modified by	Comments
0.1	20/03/2022	Laurent Dominique	Draft Table of Content
0.2	23/08/2022	Laurent Dominique	Initial draft
1.0	30/11/2022	Laurent Dominique	Input from contributors taken into account

Disclaimer

Table of Contents

List of Tables

List of Figures

1. Executive summary

General aim of the Common Semantic Model is to provide recommendations to improve IACS semantic interoperability. This deliverable has been provided according 3 versions, supplied respectively at M9 (core geospatial data), M 12 (Use Cases data) and M42 (definitive version) of the NIVA project calendar.

The target readers of this document are the stakeholders dealing with IACS data, such as Paying Agencies and their technical partners, European Commission (JRC – DG AGRI) and researchers.

This first part of this document is capitalising the NIVA experience about core geospatial data; it is also promoting IACS data sharing. To do so, the data model is composed of 3 diagrams designed according to INSPIRE themes:

- Land use that is modelling agricultural parcels and their crop types
- Main Land Cover that is modelling Land Cover (arable land, permanent crop, permanent grassland) on agricultural areas
- Secondary Land Cover for non-agricultural zones such as landscape features or non-eligible areas.

The second part of this document is proposing a meta-model called "Base types for EO monitoring". The model aims to provide a standardised way to describe the strategy and operations to be conducted when setting up the Area Monitoring System. It might be used in future by Paying Agencies to ensure transparency of their new monitoring system or by the academic world to conduct benchmark studies about EO monitoring experiences.

The third part of this document is capitalising the NIVA experience about data exchanges between IACS and FMIS. In practice, it is composed of a chapter about the exchange itself (based on the use of the eCrop standard) and by a chapter about the Farm Registry data model that is expected to be used as a receiver of FMIS data. Receiving data from FMIS is considered as one on the new technologies to be used to implement the CAP post-2022, contributing for instance to the Seamless Claim. The examples of eCrop messages and the Farm Registry model are expected to be used by Paying Agencies as starting point when designing their future IACS.

Though being globally self-understanding, this deliverable should ideally be complemented at least by the reading of the previous version of the Common Semantic Model (for the study conducted about common crop type list) and by the reading of deliverable D3.5 Recommendations for standardised connections between IACS and other applications (for the considerations about technical interoperability regarding the use of EO and FMIS data).

2. Foreword

2.1 General objectives

The Common Semantic model aims to provide recommendations for promoting IACS data semantic interoperability.

2.2 Deliverable versions

- **Versions calendar**

The Common Semantic Model is a living document that should be officially delivered on M9, M12 and M36 of the project, i.e. respectively 9 months, 12 months and 36 months after the official beginning of the project (that took place on 01 June 2019).

The current document is the third version of the Common Semantic Model that was officially scheduled on M36 and then postponed to M42 due to the 6 month extension allowed to the NIVA project.

- **Content and objectives**

Version	V1 (M9)	V2 (M12)	V3 (M42)
Content	Core geospatial data	Common crop type list EO monitoring IACS-FMIS data exchange	Core geospatial data EO monitoring IACS-FMIS data exchange Farm Registry
Objective	Provide a common view on current IACS data	Support and capitalise NIVA innovations	Provide starting point for the design of some blocks of future IACS
Target readers	UC teams	UC teams Paying Agencies	Paying Agencies European Commission

Table 1 – Overview of the various versions of NIVA Common Semantic Model

2.3 Current deliverable

- **Sources**

This last version of the Common Semantic Model is based on various activities:

- Collecting the feed-back of NIVA Use Cases on the WP3 initial recommendations mainly about core geospatial data and common crop type list.
- Capitalising the innovations brought by the NIVA project (mainly Use Case teams) ; this is typically the case of the Farm Registry data model and of the data exchanges between FMIS and IACS
- Taking into account the WP3 investigations about IACS data sharing; in other words, these investigations were conducted mainly for legal interoperability but they have some impact on data semantic.

- **Objectives**

This deliverable has two main objectives. The first and main one is to provide starting point to Paying Agencies in the design of the new IACS to be set up in order to implement the changes required by the CAP post 2022, e.g. the new monitoring system. The NIVA Common Semantic Model doesn't claim to provide an exhaustive solution for all the IACS components but only for those that have been considered during the project. The second objective is to facilitate IACS data sharing, and when possible its publication as open and interoperable data.

The maturity level varies according to the chapters of this document. The core geospatial data chapter is based on strong findings (JRC common data model, existing practices, feed-back from NIVA experience) and therefore the related chapter provides a set of well-defined recommendations that are expected to be adopted more or less as such by Paying Agencies. The other chapters are about more innovative topics, their content is also expected to be taken into account by Paying Agencies but as a starting point for building future IACS, i.e. significant amount of additional work and national adaptation will likely be required but it will be nevertheless quite more efficient than just beginning from scratch.

- **Reading recommendations**

This deliverable has been written in order to provide main or definitive findings of the NIVA project regarding IACS data semantic interoperability. As a consequence, it is supposed to be self-understanding, i.e. it can be understood without having to read previous versions of D3.2.

However, for motivated readers willing more detailed information, it may be useful to consult also some chapters of first and second version of the Common Semantic Model or other NIVA deliverables.

Chapter in D3.2 (v3)	Other NIVA deliverables	Reading advise
Core geospatial data	D3.2 (v1) provides an analysis of the JRC data model, a survey of existing practices in PA and some recommendations. The chapter about crop types of D3.2 (v2) provides a detailed analysis about the interest, difficulties and possible solutions for common crop type list for IACS.	D3.2 (v3) summarises and/or updates the previous versions Read first D3.2 (v3) If interested, consult D3.2 (v1) and/or the crop type chapter of D3.2 (v2)
Base types for EO monitoring	D3.2 (v2) proposes a draft version of a data model about EO monitoring.	D3.2 (v3) updates and replaces D3.2 (v2) Read only D3.2 (v3)
Data exchanges between IACS and FMIS	D3.2 (v2) includes a chapter about this topic based on early stage of progress and is therefore partly outdated. D3.5 includes a detailed chapter on this topic, providing explanations about FMIS, capitalising NIVA experiences and addressing whole set of interoperability issues	D3.2 (v3) is summarising and updating D3.5 with focus on semantic interoperability. Read D3.5 and D3.2 (v3)
Farm Registry	Data Model of the UC3 Farm Registry	D3.2 (v3) provides a generic and conceptual model, based on the content identified by UC3 Read first D3.2 (v3) If interested, consult UC3 data model (implementation example)

Table 2– How to read NIVA deliverables related to semantic interoperability

In addition, Core geospatial data and Farm registry are also concerned by legal interoperability; therefore, it may be useful to consult also deliverable D3.8 Profile of priority data for external applications.

- **Modelling principles**

This deliverable includes some data models. They are conceptual data models that will have to be implemented through logical models (to be decided by Paying Agencies).

These models are using the standardised modelling principles of geographic information, such as UML (Unified Modelling Language) diagrams and reuse of ISO 19000 types.

Main patterns used in UML are:

- Feature types that are classes of objects
- Feature types have various attributes that may be geometric (points, curves, surfaces), character strings, numbers (real, integer …), dates, boolean or defined as code lists or as data types. These attributes have a multiplicity expressing the number of values the attribute may take for each object instance of the feature type. By default, the multiplicity is 1 and is not indicated.
- Data types that are generally used for composed attributes
- Code lists that provide the potential values of an attribute
- Inheritance
  - Inheritance is between a child and a parent feature types
  - It is shown by a specific arrow
  - The arrow destination is the parent feature type
  - The arrow origin is the child feature type
  - The child feature type inherits from the attributes of the parent feature type
  - In summary, the child feature type is a sub-class of the parent feature type having both the generic attributes of the parent and its own specific attributes.
- Associations
  - Associations are shown by lines
  - A simple line is used for ordinary associations
  - A line with diamond is used for compositions

	
Inheritance arrow	Composition arrow

Table 3 Some modelling patterns used in UML

In practice, the NIVA model is built on the JRC data model (modeling current IACS) that is importing itself the external data models of INSPIRE and ISO.

The NIVA data model may be found on the Zenodo web site: https://zenodo.org/record/7383022#.Y5nmu1GZPIU

2.4 Glossary

Definition of the terms and acronyms used in this deliverable may be found in the Common Glossary, available on the "Deliverable" page of NIVA web site: https://www.niva4cap.eu/deliverables/

3. Core geospatial data

3.1 Introduction

First version of the NIVA Common Semantic Model focused on considerations of the core geospatial data used in current IACS, i.e. in IACS dedicated to CAP 2014-2022.

These general considerations included:

- A reminder and analysis of the related part from the JRC data model (current IACS)
- Results of a survey about semantic interoperability conducted among Paying Agencies (mainly those being NIVA partners)
- Potential alternative model, generally just a simplification of the JRC data model.

This initial version was mainly dedicated to the NIVA partners and more especially to the Use Case teams in order to facilitate design and development of their tools, by making them aware of the communalities and differences between the various national IACS.

The second version of the NIVA Common Semantic Model includes a big chapter about common crop type classification: requirements, current practices in Paying Agencies, existing standards and possible solutions for NIVA project and future IACS.

This chapter of the last version of the NIVA Common Semantic Model aims to provide feed-back from the NIVA experience about use of core geospatial data, including not only the practical experience from Use Case development but also the most theoretical experience gained from the investigation conducted by WP3 about IACS data sharing. In addition, this chapter aims to provide recommendations for core geospatial data modeling in order to facilitate both IACS data processing (for control of CAP payments in the context of CAP post-2022) and IACS data sharing towards external users (e.g. in environment or climate change domains).

3.2 Agricultural parcel

Agricultural parcel is a key feature type in IACS and even more for the NIVA project. It has been used by almost all Use Cases and it is also the information most widely required by users for IAS data sharing. As a consequence, this is the feature type for which there is the most significant feed-back from NIVA experience.

- **Initial context and NIVA proposal**

The initial NIVA proposal was not providing any definition of an agricultural parcel but just recognizing diversity of national practices and recommending Use Cases teams to document the definition and/or selection criteria chosen the for this feature type.

The proposed model itself was quite simple as shown by following figure.

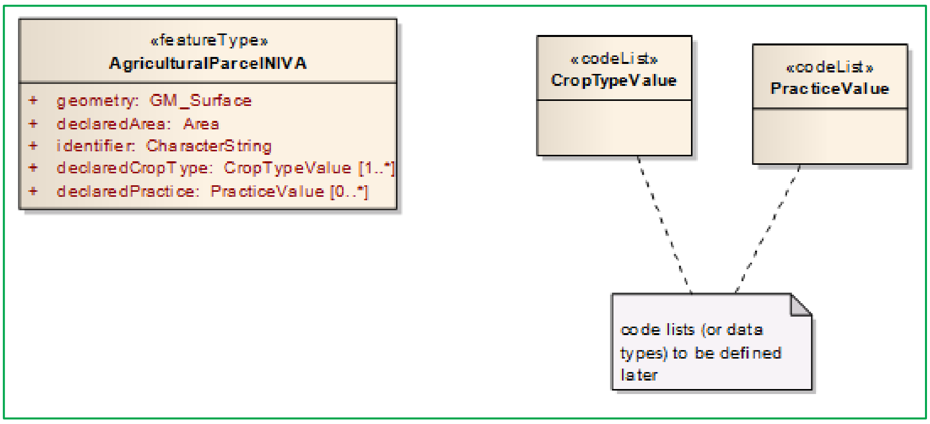


Figure 1 Initial proposal for modelling of Agricultural Parcels

- **What we have learned from NIVA experiences and investigation**

Definition of Agricultural Parcels

The JRC data model proposes a relatively generic definition of agricultural parcels "continuous area of land, declared by one farmer, which does not cover more than one single crop group ...." but makes distinction between Agricultural Parcels and Ecological Focus Area (EFA) that are modelled as two feature types.

This distinction does not seem widely implemented by Paying Agencies who tend to model a single feature type Agricultural Parcel, the EFA associated practices generally being modelled through the national crop type list or through specific attributes.

This distinction does not seem to be relevant for external users who are mainly interested in getting any parcel with associated crops without having interest in the associated payment scheme. In addition, most of these external users are not aware of EFA and would not understand their meaning.

The other drawback of the JRC definition is the reference to the farmer declaration whereas in the new CAP, the necessity of farmer declaration might or even should be removed.

Recommendation n° 1

For the practical implementation of new IACS, it is recommended to adopt a generic definition of the feature type Agricultural Parcel. Distinction between Agricultural Parcels and Ecological Focus Areas may be done through specific attributes.

Agricultural Parcel geometry

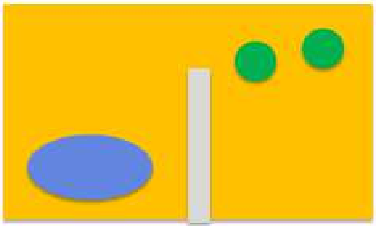
		
Real world : in the agricultural parcel, there is a lane (ineligible element), a pond and 2 trees (landscape features)	Option 1: Paying Agency captures simple geometry of the parcel; landscape features and ineligible elements are captured in separate layers. Eligible surface is later derived by computation for payments.	Option 2: Paying Agency natively excludes landscape features and ineligible elements from the parcel geometry.

Figure 2 Capturing geometry of Agricultural Parcels

With current IACS, the two options of the above illustration are valuable, with advantages and drawbacks and the NIVA investigation have shown that both are used.

With the new monitoring system, the second option looks quite better:

- The geometry to be submitted to an EO process should involve only the cultivated part; it would not be relevant to take into account the pixels corresponding to ineligible elements or to landscape features
- In FMIS, crop parcels are representing only the cultivated part so it will be easier and more meaningful to match them with the IACS agricultural parcels captured according option2.

Recommendation n° 2

The geometry of Agricultural Parcel should be natively captured excluding landscape features and ineligible elements. If this is not the case, an additional layer should be derived by computation in order to provide the geometry of "cultivated part" only.

Agricultural Parcel identifier

Agricultural Parcels are the main priority for IACS data sharing as their publication is both widely required by external users and mandated by some European regulations. However, protection of personal data has also to be ensured due to GDPR and to national laws and more generally for ethics reasons (not harming farmers).

Recommendation n°3

To facilitate IACS data sharing, there is interest for an anonymous identifier (without any personal information) that may be made publicly available for everyone whereas the link to the farm or to the farmer is managed through an external key, used only by the Paying Agency and the farmer (or farmer representative).

For the common data model, the UC3 team investigated the idea of a potential common format for parcel identifiers in Europe and proposed "NUTS code + sequential number on 8 characters". This proposal was considered as too demanding and was strongly rejected by the Paying Agencies answering the survey about semantic interoperability.

However, the UC1b experience has shown the interest of an identifier integrating some administrative coding. Some French data about agricultural parcels had to be extracted at department level (NUTS3) for the computation of agro-environmental indicators. In a first step, the extraction was done (without great care) by using current parcel identifier that is composed of a farmer identifier + parcel number. As a result, there was some selection errors as the extraction provided the parcels belonging to a farm located on a given department instead of providing the parcels located on this department. In addition the French law forbids providing the link between the parcel and the farm; so the native parcel identifiers had to be removed. In a second step, the IACS data was better prepared, by removing current identifier and replacing it by the department code and a meaningless incremental number, making the extraction easier and enabling to provide parcel identifier for data publication.

Recommendation n° 4

A parcel identifier including code(s) of national administrative units may be of interest to facilitate data extraction and should be considered by Paying Agencies for the new IACS.

NOTE: It is expected that administrative units will be the area of interest of many external users of IACS data. The extraction of required data is always possible through a spatial query but the presence of administrative code(s) in the identifier will enable the data extraction to be processed by a simpler semantic query.

From the investigations about technical interoperability and from the discussions with the Open IACS project, it also appears that providing identifiers as URI (Uniform Resource Identifier) as recommended par the Linked Open Data principles would facilitate reuse of IACS data on the Web.

Though this has not been identified as a key priority by the user survey conducted by WP3, the design of the new IACS to be set up may be a good opportunity for Paying Agencies to envisage and possibly adopt this practice.

Recommendation n° 5

When designing their new IACS, Paying Agencies are invited to consider the provision of agricultural parcel identifiers as URI, in order to enable their use as Linked Data.

In databases, persistency is often a characteristic expected from identifiers. A general and minimum rule is that if there is no change in a given feature characteristics, its identifier should remain the same. In case of small changes, the notion of versioning is often used:

- The feature is considered to remain the same (and to keep its identifier)
- But it is becoming a new version

This is the solution recommended by the INSPIRE Directive together with life-cycle information, through the attributes beginLifespanVersion and endLifespanVersion that document the start and end date of a given version of a feature in the database.

Of course, this should be completed by well-documented life-cycle rules, explaining for instance what are the small changes leading to a new version or the big changes leading to new feature(s).

The WP3 survey about semantic interoperability has not investigated the practices about persistency of identifiers on agricultural parcels but general feeling is that, as agricultural parcels have to be declared on a yearly basis, most of Paying Agencies don't really care about persistent identifiers on this feature type.

However, during the NIVA project, some requirements have been identified: persistent (and shared) identifiers would facilitate data exchange between IACS and FMIS; they might help also to aggregate in a meaningful way the results of the nitrate indicator (UC1b) currently computed at pixel level on a period over 2 agricultural campaigns. More generally, persistent identifiers could facilitate the Paying Agency control of parcels under a several years payment scheme (permanent grassland, crop rotation) and the farmer declaration in GSAA.

Recommendation n° 6

When designing their new IACS, Paying Agencies should consider the adoption of persistent identifiers, the documentation of their life-cycle rules and the provision of temporal attributes according to the INSPIRE mechanisms.

NOTE 1: The above recommendation means that the issue has to be raised and investigated, by identifying more accurately the expected benefits and the implementation solution or difficulties (feasibility).

NOTE 2: This recommendation is a priori relevant for other IACS feature types (at least those of LPIS and GSAA) but the priority is on agricultural parcel that is the key feature type and the most widely required by external users.

Crop type

The detailed analysis conducted for previous version of this deliverable about potential common crop type list was proposing following solution:

- For data exchange, a simple and standardized botanical classification, such as LUCAS
- For native adoption by Paying Agencies, the combination of the detailed botanical classification of EPPO and of the product oriented classification GPC.

National existing classifications used by Paying Agencies are often mixing various concepts, such as species, product, agricultural practices … The main recommendation was clean modeling, i.e. an attribute for each concept.

During the NIVA project, in practice, only two Use Cases have required a common crop type list, namely UC1b about agro-environmental indicators and UC3 about Farm Registry.

For UC1b, the common crop type list has been used mainly to document the matching to be done from existing data (standard list) to the crop type list required for computation of the indicator. The matching is explained in the user guide and has been done in practice by the testing countries. In addition, this common crop type has been used for the publication of the indicator results for Carbon Tier 1. The Use Case has chosen the LUCAS classification with following results:

- The crop type lists required for computation of carbon and biodiversity indicator are simpler, more generic than LUCAS and so, the matching has been possible without any trouble
- Though being relatively simple (small number of classes), the crop type list required for computation of nitrate indicator requires seasonality information that is absent from LUCAS. The UC1b team has proposed an enrichment of the LUCAS code list by adding a more detailed level carrying the seasonality information (winter/spring or summer crops)
- The lack of seasonality information in LUCAS is limiting the analysis of the results for carbon indicator; typically, it is not possible to compare if carbon indicator is better for winter or summer crops.

For UC3, the common crop type has been adopted in the Farm Registry data model and some testing has been done, by importing current data from existing registries (Andalusia, Estonia) to the common database. In a first step, LUCAS was tested but was considered not detailed enough, entailing loss of information during the data import. In second step, the very detailed EPPO classification has been preferred.

Use Case - Component	Crop type list	Comment
UC1b – carbon indicator	LUCAS	LUCAS is enough for indicator computation but lack of seasonality information limits the result analysis.
UC1b – nitrate lixiviation indicator	LUCAS	Lack of seasonality information limits is an issue for indicator computation
UC1b – biodiversity	LUCAS	LUCAS is enough for indicator computation
UC3 – Farm Registry	LUCAS EPPO	With LUCAS, detailed information is lost EPPO enables to transfer detailed information

Table 4: Feed-back about use of standardised crop type lists

The investigations about IACS data sharing have shown the interest of interoperable data in Europe, regarding agricultural parcels and crop types. The most user-friendly solution would be publication of IACS data using a common and relatively simple botanical classification. LUCAS was considered as best candidate but the NIVA experience has shown some limits. As a consequence, the choice of a recommended common crop type list remains an open issue.

However, as these standards are based on same principles of botanical classification, there are lots of similarities and matching between them is generally relatively easy. For instance, the Open IACS project has established matching between various crop type standards, namely between LUCAS, EPPO, LCCS (from FAO), FADN, AGRIPROD and the common catalogues of varieties of agricultural plants and varieties of vegetables.

Recommendation n° 7

When designing new IACS, Paying Agencies should devote a specific attribute about crop species and fill it using a botanical classification. Other concepts linked to the crop description should be modeled using other attributes.

NOTE 1: The recommendation entails a less compact model than current practice (generally a single list mixing various concepts) but aims to facilitate IACS data sharing, by facilitating provision of interoperable data.

NOTE 2: To get harmonized data at European level, the national botanical classifications will have to be matched to a common classification. This might be done by users or by Paying Agencies (e.g. in future, if DG AGRI mandates or recommends a given common classification).

Declared area

The NIVA UC5b has developed the tools to implement the Seamless Claim, i.e. a dataflow not requiring any longer farmer declaration. The testing assessment has shown that one of the benefits was to ensure more payments to farmers: there is no longer need to declare cultivated area, so no error on this topic and so, no penalty to be applied.

The Seamless Claim is quite a disruptive process and so it will not be operational on short term in most Paying Agencies. However, this attribute should become useless in longer term.

Declared practice

The initial proposal was offering a simplified view on this topic, with a single attribute whose possible values are included a national code list.

The experience of the NIVA project has shown that the agricultural practices have a strong interest in the context of the CAP post-2022; therefore a single attribute is of course far from able to provide the required information. Potential solution is provided in chapter 6 of this document about the Farm Registry data model.

- **Proposed model**

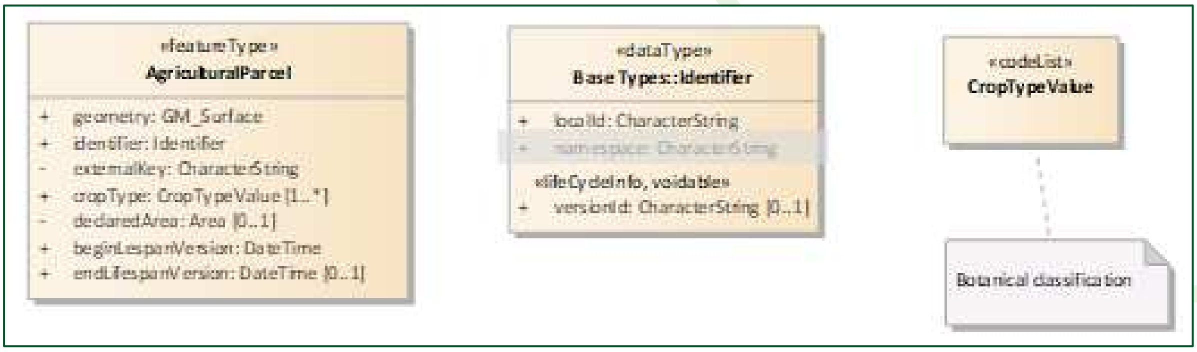


Figure 3 NIVA proposed model of Agricultural Parcels

The feature type AgriculturalParcel has following attributes:

- Geometry: to be used for CAP payments and for data publication
- Identifier: to be used mainly for data publication, this identifier should not include any personal data and should be persistent; the illustration above is using the INSPIRE data type
    - localId is the identifier to be captured by Paying Agencies
    - the namespace is dedicated to the publication of INSPIRE data; it aims to provide context of data and it is generally including country code, producer acronym, product name … (e.g. FR.ASP.RPG); there is no need to capture it natively as it may be produced on-the-fly when delivering INSPIRE data
    - versionId is to be used if Paying Agencies decide to manage versioning of agricultural parcels

- externalKey is a code enabling to ensure the link to the farm and/or to the farmer and so dedicated to enable CAP payments.
- cropType according to a botanical classification; it is to be used for CAP payments and for data publication
- declaredArea: : to be used for CAP payments; it might be removed once the Seamless Claim system is operational
- beginSpanlifeVersion and endLifespanVersion are the life-cycle attributes aiming to manage more continuous update of agricultural parcels (with the versionId); they may be used both for CAP payments and for data publication.

NOTE: more detailed information about INSPIRE identifiers and temporal attributes may be found in the INSPIRE Generic Conceptual Model: [INSPIRE Generic Conceptual Model | INSPIRE (europa.eu)]

3.3 Reference parcels and agricultural areas

- **Initial context and NIVA proposal**

NIVA was recognizing the various options offered to Member States regarding Reference Parcels, as illustrated by following figure.

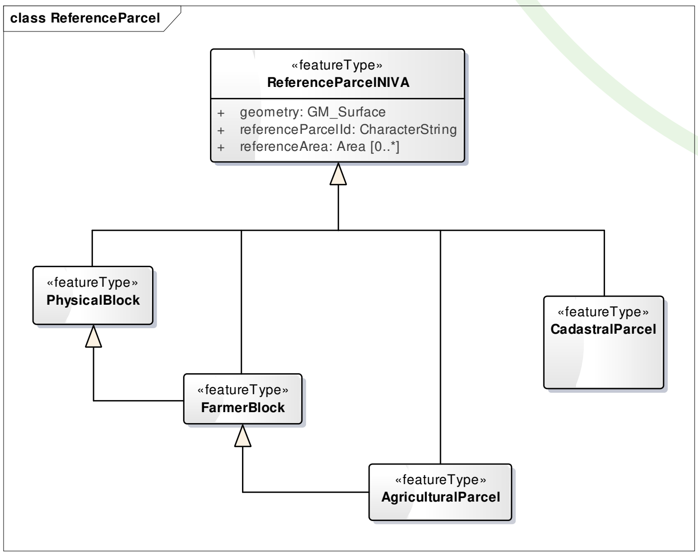


Figure 4: Initial NIVA data model for reference parcels

The JRC data model also includes a feature type called AgriculturalArea and dedicated to carry the Land Cover information. From the investigations conducted by NIVA, Paying Agencies were not all using a specific layer for Land Cover as many of them were using the Agricultural Parcels or the Reference Parcels.

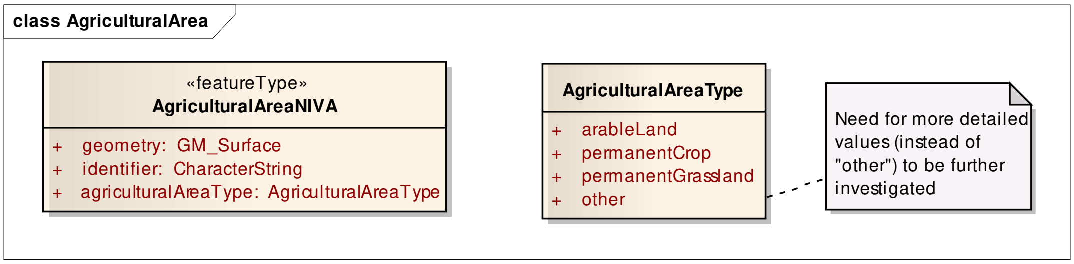


Figure 5: Initial NIVA data model for agricultural areas

- **What we have learned from NIVA experiences and investigation**

The variety of national practices regarding Reference Parcels has not been an issue in the NIVA project. The reason is likely that this feature type has not really been useful for the Use Cases, the main feature of interest being the agricultural parcel.

UC2 about prefilled application has developed a preliminary parcel delineation tool; this tool is based on the analysis of satellite images (e.g. raw or super-resolution Sentinel-2) and the results are provided to farmers for their declaration. The result assessment has shown that this tool was of specific interest in case of reference parcels being physical blocks (that provide very coarse information to locate the agricultural parcel) or cadastral parcels (that entail non-relevant boundaries).

For CAP post-2022, DG AGRI is proposing that there should be only one value of LandCover {arable land, permanent crop, permanent grassland} per Reference Parcel. However, there is strong reluctance from Member States who are afraid of too much effort in addition of the Area Monitoring System to be set up.

The investigations about IACS data sharing have shown that external users have no requirements regarding Reference Parcels but some interest for the Land Cover information (that is anyway under an INSPIRE theme). This Land Cover information is considered of interest mainly at European level.

WP3 recognizes the benefits of the DG AGRI proposal (one Land Cover value on each Reference Parcel):

- It is consistent with the process recommended with AMS (check basic lines)
- It would lead to Reference Parcels more helpful for farmer declarations
- It would remove the need of a specific layer to carry Land Cover information, such as the AgriculturalArea feature type and therefore it would contribute to simplify and harmonise the national practices, what would be quite beneficial for IACS data publication at European level

Nevertheless, WP3 also understands the feasibility issues raised by Member States. It is not within WP3 competence and role to decide on the cost-benefits of such a recommendation.

- **Proposed model**

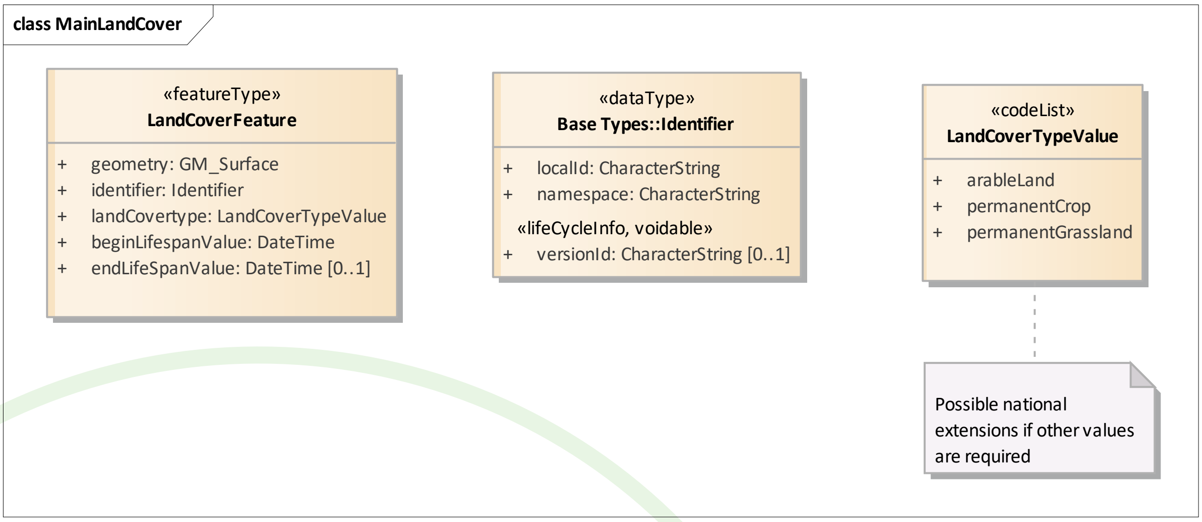


Figure 6 NIVA proposed model for main Land Cover information

NOTE 1: In current IACS, the LandCoverFeature may be the ReferenceParcel, the AgriculturalParcel or it may be located in a specific layer (AgriculturalArea). As explained above, the three options are expected to be used during couple of next years. On longer term, the use of ReferenceParcels is proposed as best candidate.

NOTE 2: The information described in above illustration is expected to be captured by Paying Agencies and doesn’t include information provided by farmers; therefore, all attributes may be used for data publication.

3.4 Non-productive elements

- **Initial context and NIVA proposal**

The starting point was the JRC data model that includes two feature types (LandscapeFeatures and EcologicalFocusArea) that have different impact regarding CAP payments, landscape features influencing the eligible area whereas EFA may receive specific payments. This approach entails some duplication as for instance, a hedge is considered both as a landscape feature and an EFA.

The survey about semantic interoperability has shown different practices from Member States, generally implying the capture of landscape features and non-eligible elements.

NIVA proposed a topographic view on landscape features and non-eligible elements in the alternative model below.

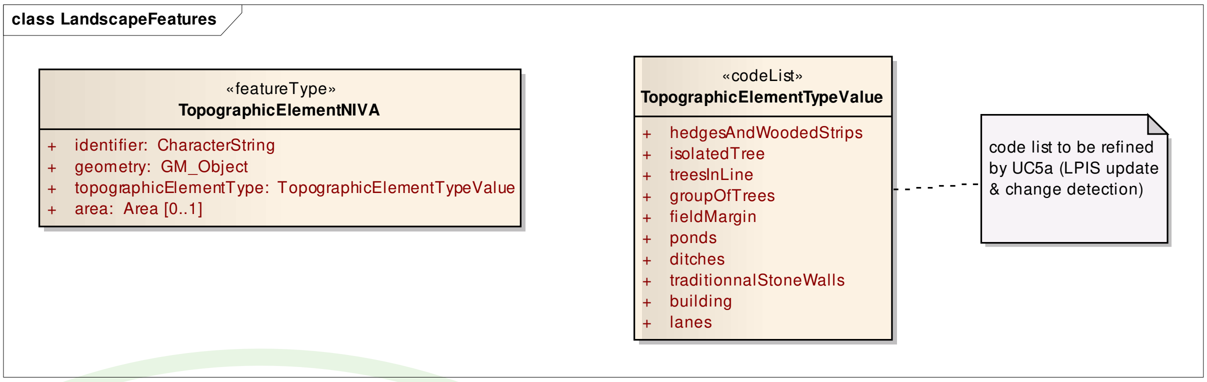


Figure 7: Alternative NIVA data model for landscape features and non-eligible elements

- **What we have learned from NIVA experiences and investigation**

UC5a (LPIS update & change detection) has developed tools for automatic detection on non-productive elements that are necessary to update LPIS and the boundaries of reference parcels.

In first phase, a survey was conducted among Paying Agencies to collect the lists of non-productive elements used by Paying Agencies; it proposed a detailed list (around 40 values) of landscape elements such as isolated trees, scrubs, buildings, pylons, ponds, stone walls, rivers and asked to fill the payment status: non eligible / eligible / EFA. This survey has shown both some commonalities but also a wide variety of practices:

- The payment status doesn't depend only on the nature of the landscape element but also of other conditions (size, length, width)
- Whereas isolated trees are generally considered as eligible elements and sometimes as EFA, the rules regarding other vegetation landscape features (hedges, trees in line, groves, scrubs, afforested areas) are quite diverse depending on countries
- In most cases, artificialized landscape elements are not eligible. A frequent exception concern traditional stone walls; there may be more specific exceptions concerning only a few countries (e.g. kitchen gardens, small areas around pylons).
- Regarding hydrography, ditches and ponds are eligible in several countries whereas lakes and rivers are most often considered as not eligible.

The survey did not include detailed questions about the capture of these various landscape elements.

Based on the results on this survey, UC5a decided to adopt a modular approach. There is a specific algorithm for each of the main categories of these non-productive elements. Denmark and France who are the co-designers of these tools have chosen different approaches:

- France has focused on the detection of buildings and high vegetation (using DTM and DSM)
- Denmark has chosen to detect trees in lines, trees in groups, ponds with surrounding vegetation, artificial covered areas, unfortified roads and pathways.

These two lists don't form a partition of potential non-productive elements, as there are overlaps and gaps. The choice was about dealing with main elements rather than trying to address all of them.

The biodiversity indicator designed by UC1b (agro-environmental indicators) requires as input data, semi-natural elements. Some discussion took place to decide if the landscape features and non-eligible elements of IACS could be used directly; this answer was negative, mainly because IACS is covering only agricultural areas whereas the information was required on whole territory but also because there was some fear that, even on agricultural areas, the IACS data is not reliable enough and above not homogeneous enough in Europe (various modelling practices, no common classification).

The landscape features and non-eligible elements of IACS have nevertheless been used for the computation of agricultural area (also required as input data). As some features are captured as points or lines, it was necessary to apply a buffer to find their approximate area. The case of areas with prorata information was also identified as an issue but no practical solution has been found.

The biodiversity indicator was designed at Tier 1 level, requiring only the quantity of semi-natural elements. A more accurate indicator (biodiversity indicator Tier 2) would require not only the quantity but also of semi-natural elements but also their nature (vegetation, water, bare soils).

The investigations about IACS data sharing have shown some user requirements regarding this kind of information. Unfortunately, until now, the corresponding data is rarely openly published and available for potential users.

The JRC discussion paper about Interoperability and Data Harmonization investigates which is the most appropriate corresponding INSPIRE theme for each piece of spatial IACS data. The candidate themes are Land Cover, Land Use and Regulated Areas. One of the conclusion is that there is likely limited demand at European level for the landscape features and EFA and so, no strong need for a common code list: publishing data in "as-is" form and documenting the national or regional code list could be a reasonable approach.

- **Proposed model**

The NIVA proposal is based on following principles:

- A topographic point of view with a feature type "LandscapeFeature" gathering all the landscape objects not devoted to agricultural purposes and located within or close to reference parcels, independently of their impact on CAP payments
- The proposed model may be used for data production; collect first according what can be seen on Earth surface and then enrich it by the attribute paymentStatus whose exact description is let to Paying Agencies.

- Nevertheless, the main aim of this model is to enable efficient and relatively harmonized data publication ; the most relevant corresponding INSPIRE theme is considered to be Land Cover.
- A fully harmonized classification of the nature of landscape features doesn't seem achievable and doesn't look strongly required; however, it should be possible to agree at least on a high level classification. The model is proposing a very simple code list based on the highest levels of the EAGLE matrix (https://land.copernicus.eu/eagle). However, other harmonised classifications may be envisaged. For instance, a study conducted in the context of IACS65 has investigated the impact of landscape features on ecosystem services; it has proposed a generic classification of these landscape features: grassy, woody, wet, stony.
- Users may get misleading information if the landscape features are published alone as the information may be captured in a different way on permanent grassland, via the prorata attribute. Therefore, it is advised to provide also the concerned permanent grasslands.
- The attribute grasslandProrata is expected to provide the percentage of grassland agricultural and by deduction the percentage of landscape features (tress, scrubs …).

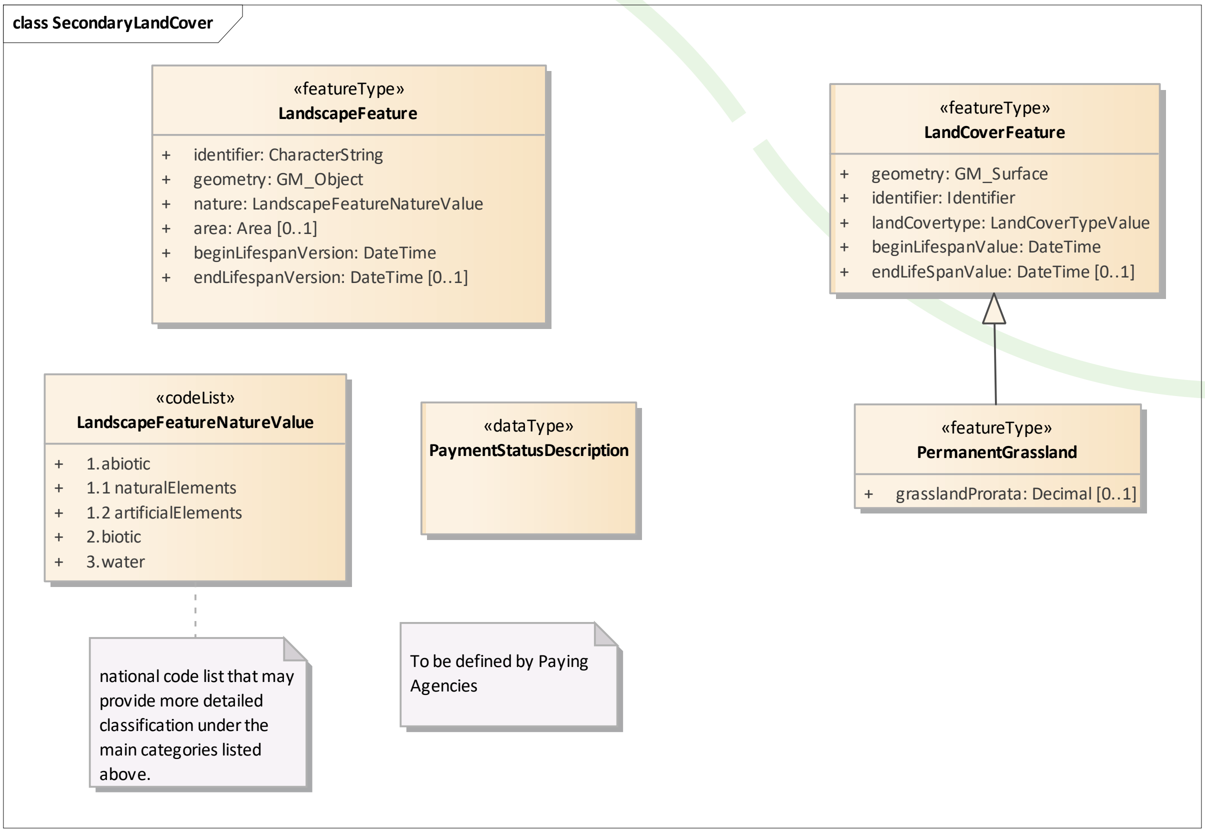


Figure 8: NIVA model for detailed (secondary) Land Cover information

NOTE: The model is allowing the capture of landscape features as GM_Object, i.e. as surfaces, lines or points. However, it should be recognized that it would be more user-friendly to provide it only as surfaces, even if these surfaces are a bit arbitrary (e.g. buffer around a point representing a tree or around the line representing a hedge). In case several geometry types are used, it is recommended to publish the data capture rules, e.g. the thresholds used between point or line and surface representation.

4. EO monitoring

4.1 Objectives

The general objective is to propose a conceptual data model to describe Earth Observation data and its use in the context of the new Area Monitoring System.

The target readers of this chapter are on one side the Paying Agencies and their technical partners implementing the AMS and on the other side, the European Commission, mainly JRC and DG AGRI who guide and control these national or regional AMS. It may be of interest also for researchers and any other stakeholders involved in AMS related activities.

This chapter aims mainly to provide common concepts that might be used to **document in a standardised way the complex EO monitoring processes.** In operational context, using these common concepts would be to ensure traceability and transparency of the control method chosen by the Paying Agencies. It may be part of a Quality Assurance approach.

Some of the base types for EO monitoring might also be reused to build a few bricks of an operational Area Monitoring System.

In scientific context, benefit would be a help to build methodologies in order to compare various EO monitoring methods and their results. These common concepts may be used also to identify gaps in current solutions and to decide on researches to be conducted.

In parallel, deep investigation has been conducted about how to get Analysis Ready Data from satellite images (mainly Sentinel 1 and 2): this work is documented in deliverable D3.5 Recommendations for standardised connections between IACS and other applications.

It is advised to read both the D3.5 chapter about EO data and this current chapter about "Base types for EO monitoring". They might also be used as **training material.**

4.2 Scope

The scope is limited to use of EO data in the context of new Area Monitoring System. It is understood as the use of temporal series of satellite images, such as Sentinel-2 and Sentinel-1.

In other words, the use of more traditional ortho-images coming from air-born sensors or from mono-date satellite images is not included in this model.

The use of secondary evidences (geo-tagged photos, FMIS data) is also out of scope of this data model.

4.3 Status

A first version of the UML main was initiated with the main objective to propose a modelling of the EO monitoring process in order to display the potential steps and options, to show the ones already done by Sen4CAP and to identify gaps. This gap analysis has helped the UC1a team to decide which functionalities missing in Sen4CAP should be developed in the NIVA tools: the purpose was to bring added value to the Sen4CAP results instead of duplicating efforts.

This first version of the data model "Base types for EO monitoring" was the basis for the related chapter in deliverable D3.2 Common Semantic Model v2 (M12).

The first version was discussed among NIVA partners, especially those dealing with EO data. The current model is also enriched from the work conducted for D3.5 about getting ARD.

This model doesn't pretend to be an exhaustive and definitive description of EO monitoring, but rather a starting point for future standardisation activities. More especially, most of the code lists included in this model provide just some examples and would need to be completed and updated according to technical evolutions.

4.3 Overview

The model is composed of 5 packages.

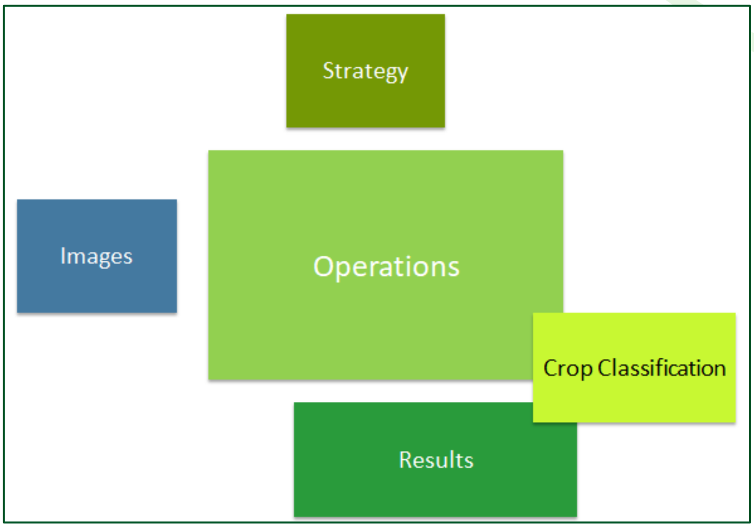


Figure 9: The 5 packages of data model "Base types for EO monitoring"

The "Strategy" package intends to describe the link between the payment scheme to be checked and the EO processes used to perform this control. It is modelling the theoretical strategy for EO monitoring.

The “Operations” package is documenting the concrete operations to be set up in order to ensure EO monitoring. It is the main part of the model.

The package “Images” aims to describe more in detail the EO sources used in the monitoring processes. It is strongly related to the chapter about EO data in D3.5.

The package “Results” provides some considerations about result reporting and assessment.

Crop classification is a specific case of process, very widely used. The package about crop classification may be seen as a kind of zoom on this specific process.

4.3 Package Strategy

This package deals with theoretical monitoring processes. It aims to describe the high level methodology to be decided when implementing EO monitoring.

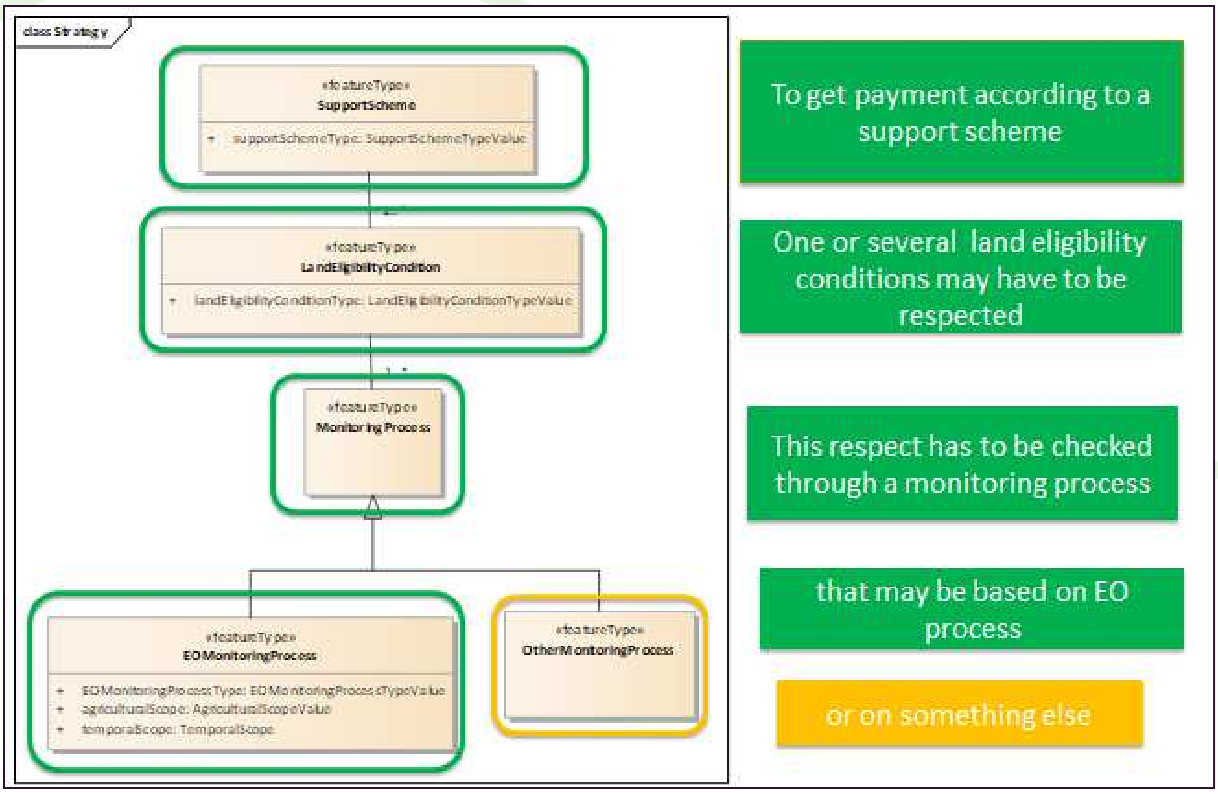


Figure 10: From support scheme to monitoring process

NOTE 1: Land eligibility conditions are obviously candidates to be checked by EO monitoring process. However, there may be some exceptions that will require other monitoring processes. For instance, it looks difficult to check some GAEC through EO monitoring.

NOTE 2: The part about Other Monitoring Process is out of the scope of this data model and won't be detailed.

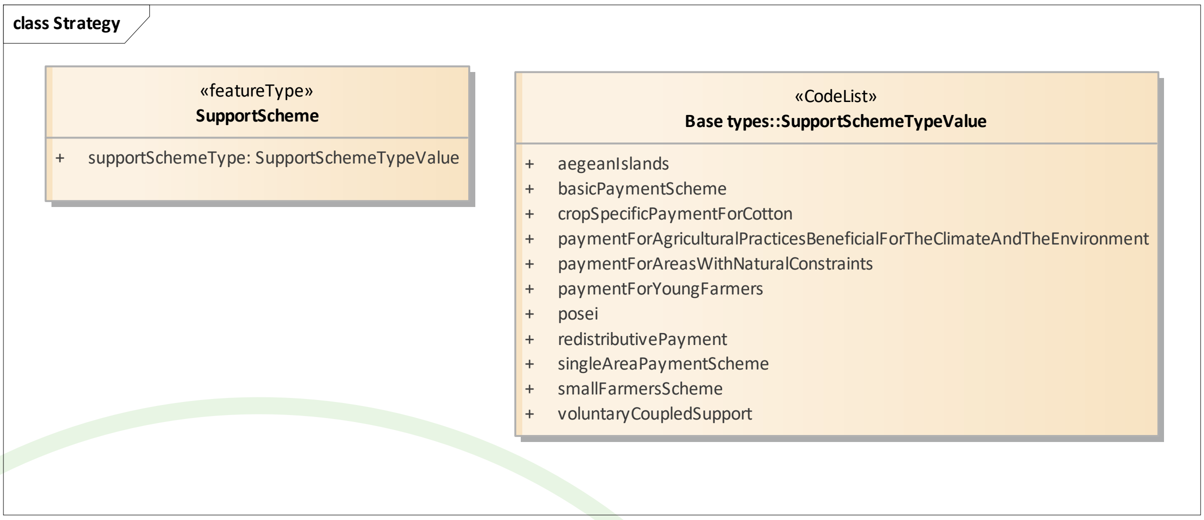


Figure 11: The possible support schemes in current PAC

NOTE: The code list about types of support schemes has been extracted from the JRC data model (CAP 2014-2022). It should be updated to take into account the support schemes of the post 2022 CAP.

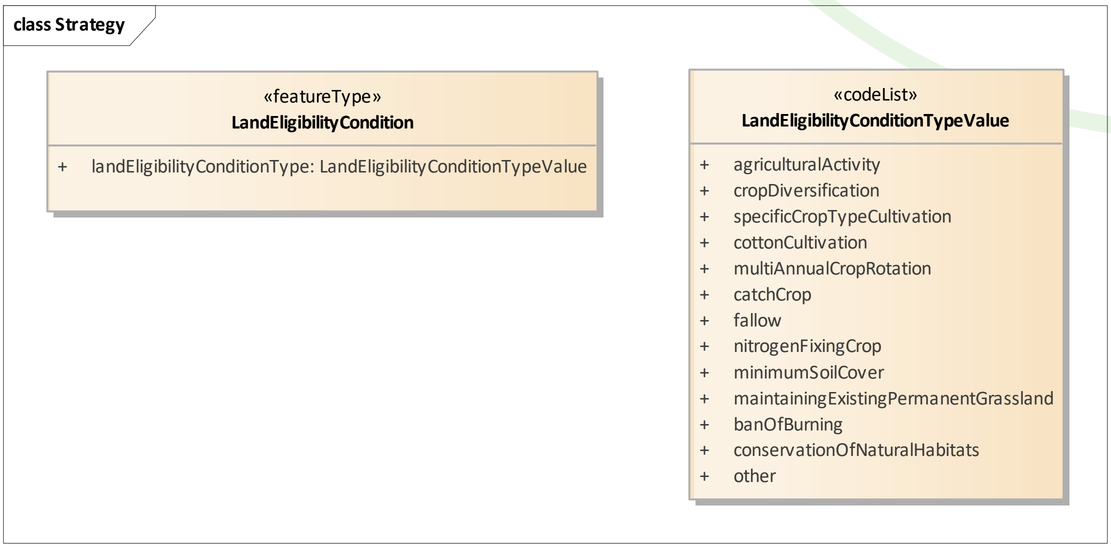


Figure 12: Some examples of land eligibility conditions

NOTE 1: The code list about land eligibility conditions is extracted and generalized from the JRC data model; it aims to provide main examples of the agricultural practices that may be monitored using EO data. This list doesn't pretend to be exhaustive.

NOTE 2: There are obviously links between the support scheme and their related land eligibility conditions. These links are not explicitly described in the current data model. However, a matrix showing these links would be quite useful. It is a potential follow-up action for future after NIVA.

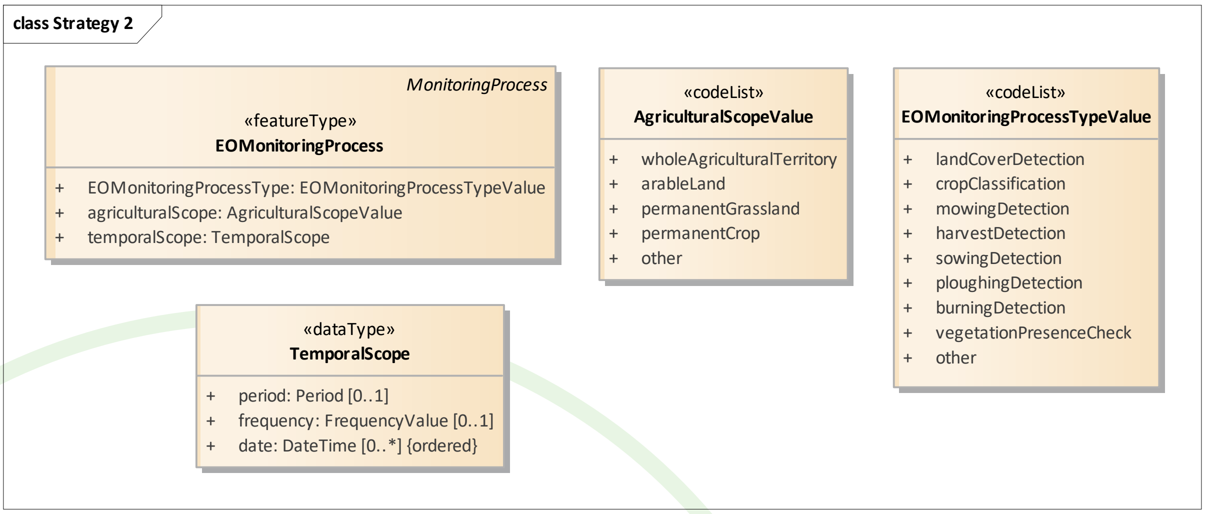


Figure 13: The EO Monitoring process

The EO monitoring process should be described by:

- How it is performed, i.e. its type (the purpose of the computing process); the code list provides some examples of frequently used processes.
- When it is performed, i.e. its temporal scope. This should be done by indicating the period when the EO monitoring process is done and the frequency (e.g. every 10 days) or the list of dates
- Where it is performed, i.e. its agricultural scope. In practice, it is expected that most EO monitoring processes will take place either on whole agricultural territory or on one of the land cover area {arable land, permanent crop, permanent grassland}

NOTE: Once fixed, the strategy is supposed to be relatively permanent. However, the temporal scope may have to be adapted to the meteorological conditions of the agricultural campaign.

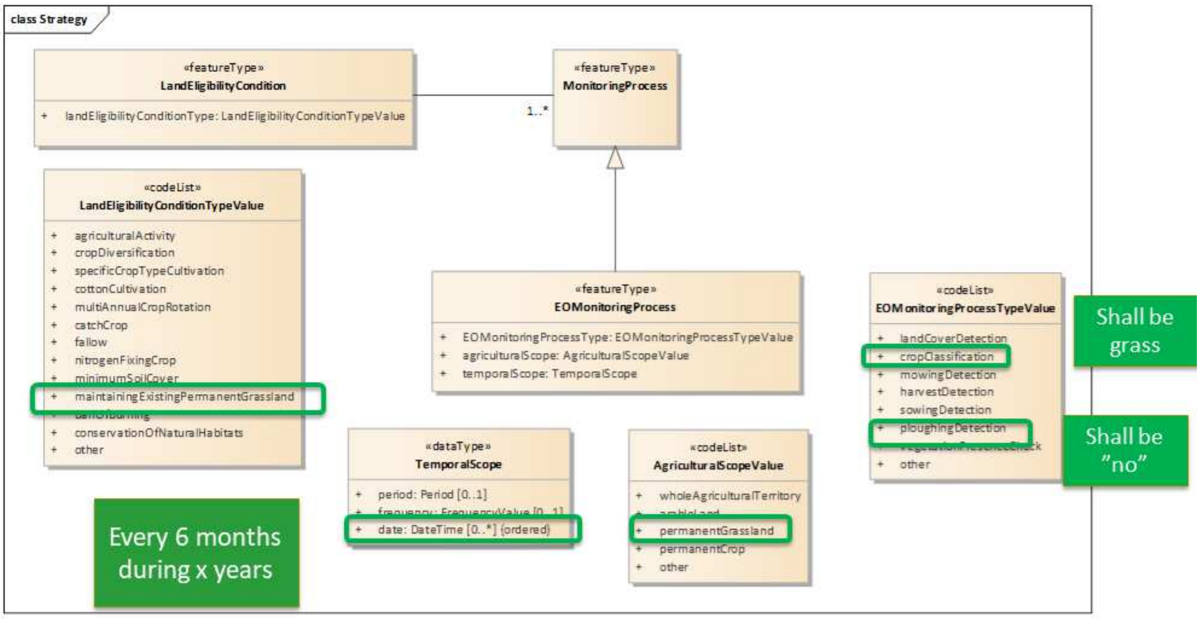


Figure 14: Potential example about maintenance of permanent grassland

The above figure illustrates a fictitious but potential example. To check the land eligibility condition about maintenance of permanent grassland, the Paying Agency decides to use 2 EO monitoring processes: one about crop classification (with expected result being "grass") and one about ploughing detection (with expected result being "NO"). Both processes are applied on permanent grasslands (as agricultural scope) and they are performed every 6 months during a couple of years: there temporal scope may be supplied as the ordered list of planned dates.

4.4 Package Operations

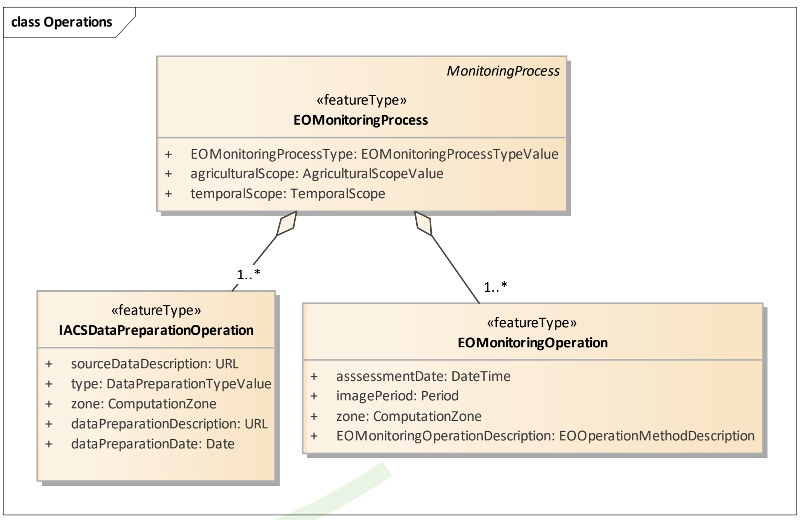


Figure 15: EO Monitoring process is implanted through concrete operations.

This package deals with concrete EO monitoring operations, performed on real data at a given date and on a given zone as shown by Figure 15.

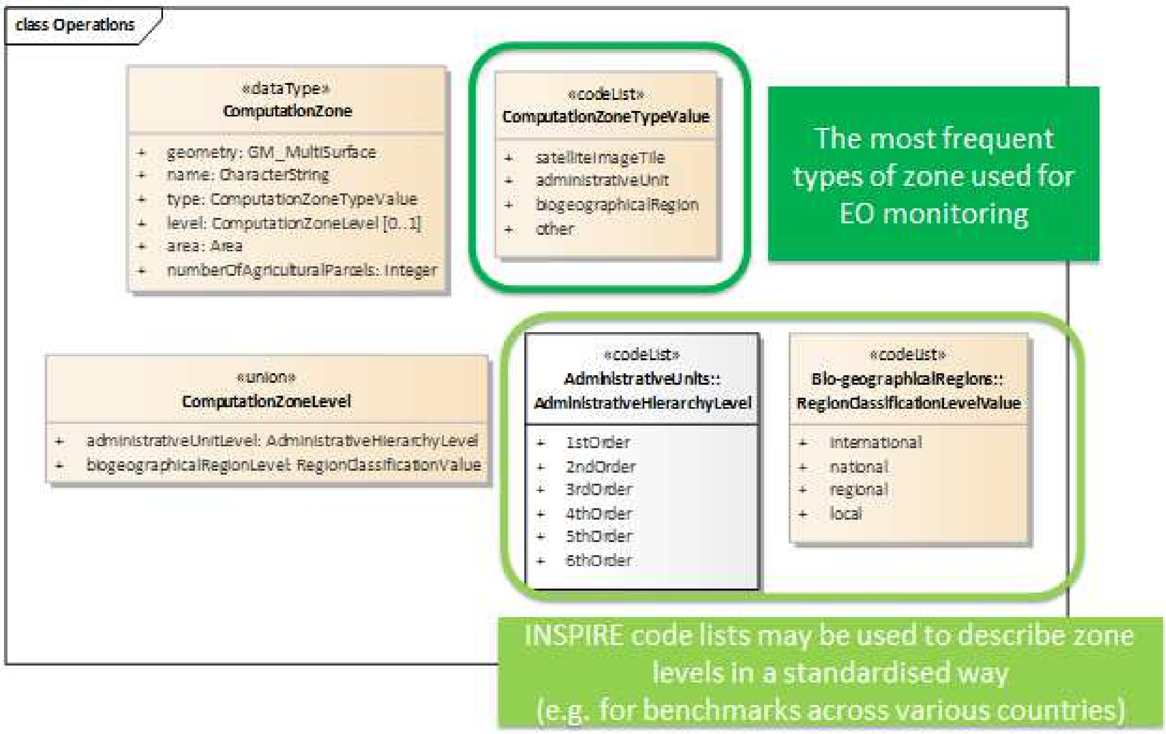


Figure 16: Detailed description of the computation zone

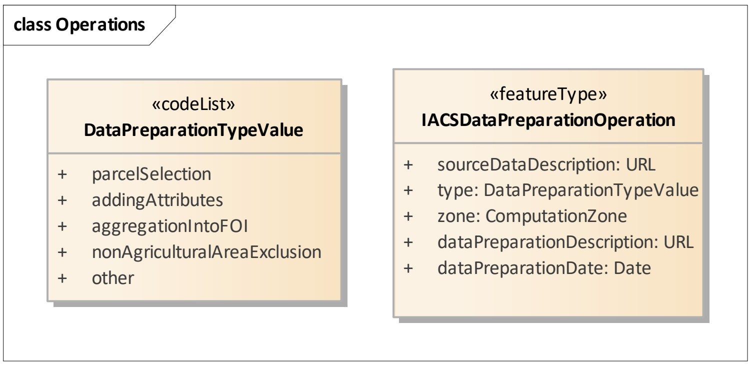


Figure 17: Preliminary IACS data preparation

Before performing an EO monitoring operation, it is generally necessary to carry out some preliminary preparation of the IACS data (generally agricultural parcels). Most common ones are selection of relevant parcels (e.g. by excluding the too small or too narrow ones), the aggregation of agricultural parcels into Feature of Interest, the addition of attributes to record pre-prepared information (e.g. crop grouping) and the results of the main steps of the computation process, the exclusion of non-agricultural areas such as landscape features (hedges, ponds, buildings ...).

NOTE: Figure 15 makes clear distinction between IACS data preparation operations and EO monitoring ones. It is the common case as, in practice, most of IACS data preparation operations may be carried about using data base management system and GIS. However, EO data may also be used for IACS data preparation: this is typically the case with the preliminary parcel boundary detection tool developed by NIVA Use Case on prefilled application (UC2).

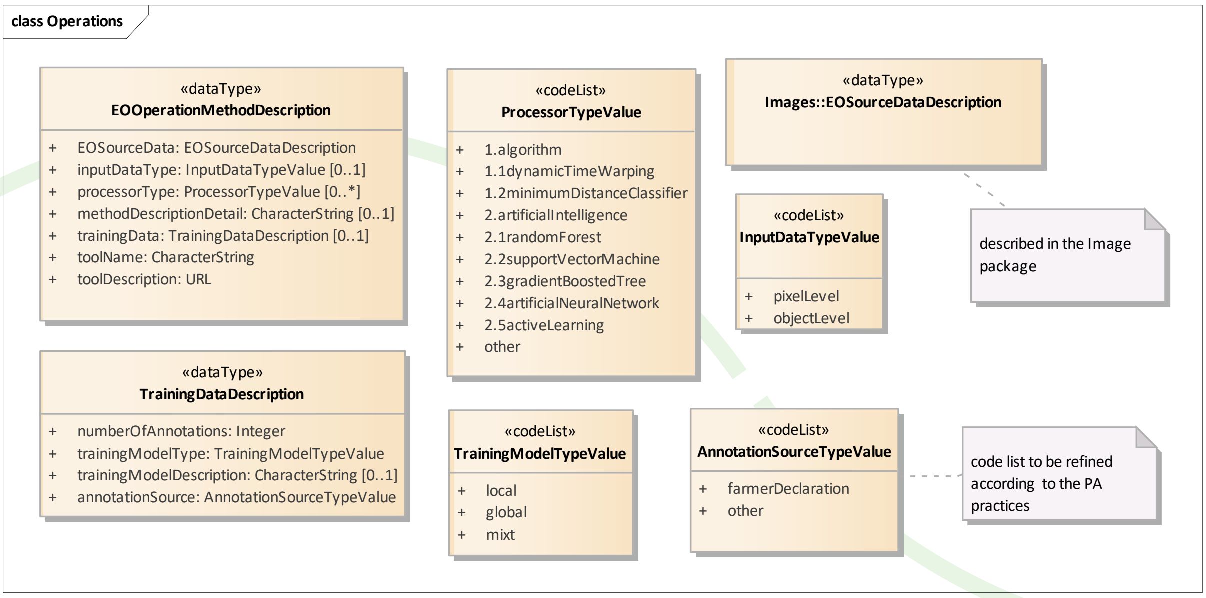


Figure 18: Description of an EO monitoring operation

In addition to the EO monitoring operation context (assessment date, period of images, computation zone), the method should be carefully documented, through the data type “EO Monitoring Operation Method Description”:

- Which kind of satellite images have been used and how : EOSourceDataDescription (details will be provided in the “Images” package)
- The type of input data used by the tool: a tool working on temporal series provided at object level requires less computation power than a tool working at pixel level but offers less flexibility to deal with heterogeneous parcels
- The kind of processor : the code list ProcessorTypeValue provides some frequently used examples
- In case Artificial Intelligence (machine or deep learning) is used, description of the training data
    - Number of annotations to get size of the training sample

- Training model type : local (only data from concerned zone and period), global (data from wider zone and/or period) or a mixt (global with stronger representation of local data)
- Training Model description: more detailed information; it might be useful especially in case of global or mixt training models
- Annotation source about how the annotations have been captured. A frequent source is the farmer declaration but other methods are possible. This information should ideally be provided as a code list (to be defined according to the monitoring context)

- More detailed information, such as method description (as free text), tool name and tool description.

NOTE: This data model proposes to document the EO Operation Monitoring Description at operation level, i.e. for each operation. This choice is relevant in the context of research, benchmark comparisons. However, in the operational context of Paying Agencies, it is very likely that most attributes of the EO Operation Monitoring Description will be the same for the operations dealing with same issue but running on different zones or at different times. Therefore, it is quite possible to factorise this information at higher level (e.g. on EO Monitoring Process).

4.5 Package Images

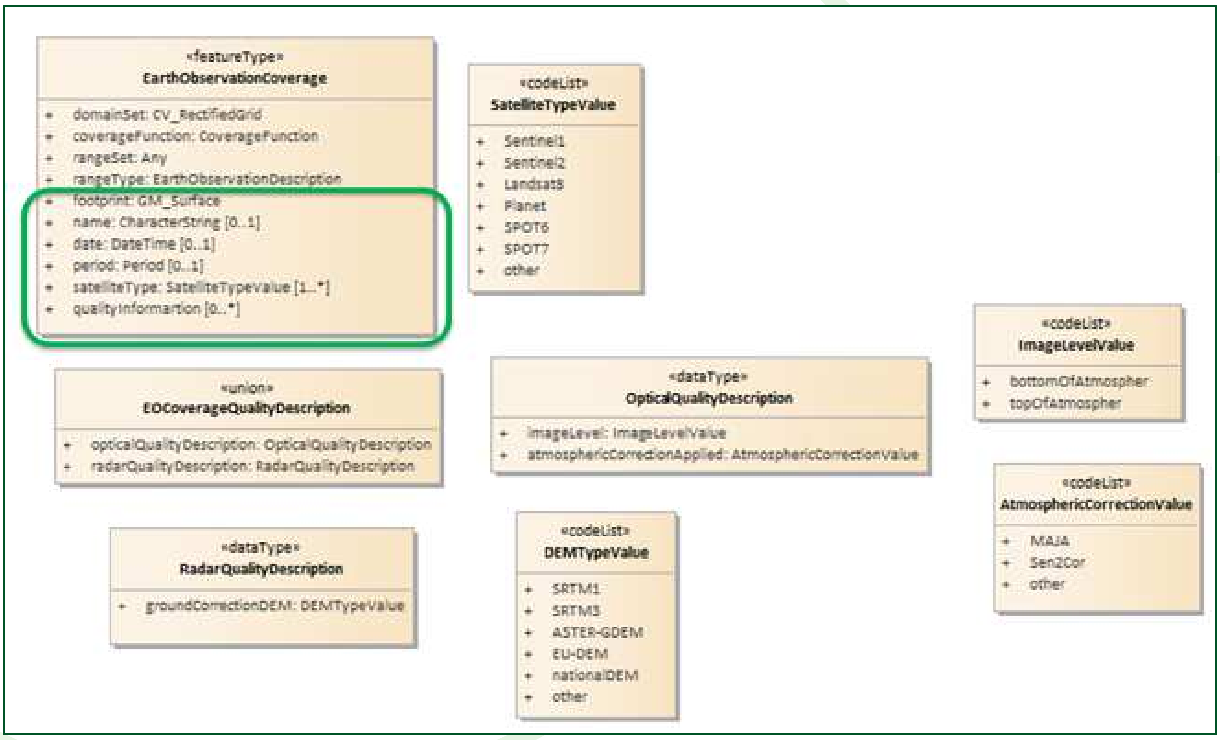


Figure 19: Information carried at image level

The satellite images are modelled as coverages (as in INSPIRE). This modelling choice is purely formal and isn't expected to have any practical impact on the EO monitoring of CAP.

The feature type Earth Observation Coverage may be used for mono-date images (in this case, the attribute "date" should be filled) or for multi-date images carrying temporal series (in this case, the attribute "period" should be filled).

The images may provide from various types of satellites. The code list Satellite Type Value provides the satellites most widely used (from the survey conducted for deliverable D3.5). It doesn't pretend to be exhaustive.

An attribute is dedicated to the quality information of these images. For optical images, the main quality information is about the radiometry, i.e. if atmospheric corrections have been applied (top or bottom of atmosphere images) and if yes, which kind of correction has been applied. The code list Atmospheric Correction Value gives as examples MAJA and Sen2Cor that are the algorithms more widely used for the correction of Sentinel-2 images. For radar images, such as Sentinel-1, the main quality information is about the choice of the DEM used for ortho-rectification as it is impacting the georeferencement accuracy.

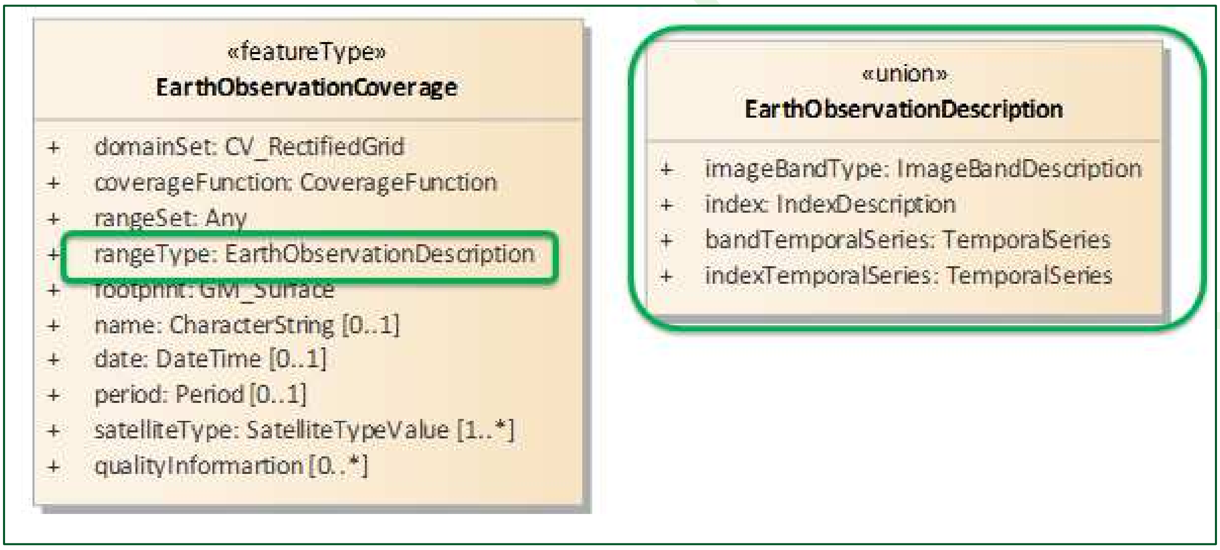


Figure 20: The range type documents the information carried at pixel level

The information carried at pixel level may be a band or an index in case of mono-date images or temporal series in case of multi-date images.

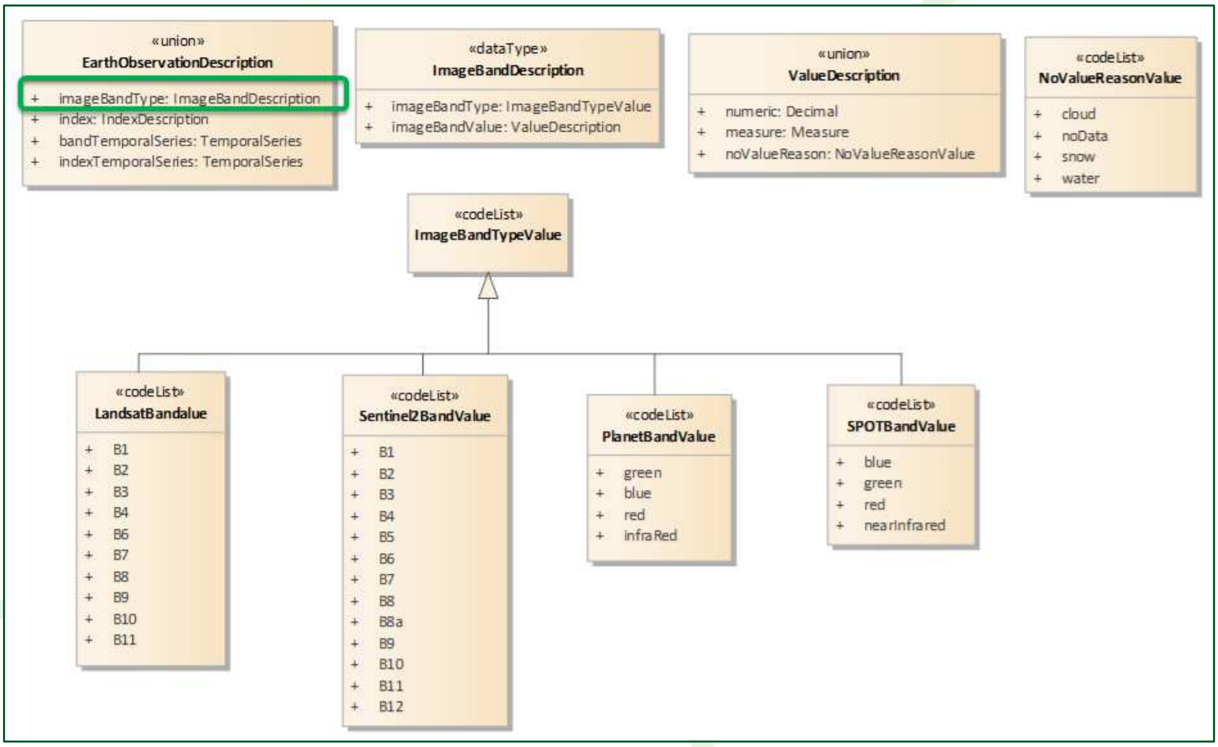


Figure 21: The case of (optical) satellite band

The information about Image Band Description (at pixel level) consists in the band name (the code list Image Band Type Value supplies possible values for the most widely used satellites) and in the value itself; for optical satellite, this value is generally a number (decimal) but it may also occur that there is no valid information, due to various issues (in practice, mainly cloud) documented by the noValueReason attribute.

The model may of course be extended to add other sources of images, such as images derived from supra-resolution and fusion processes or High High Resolution images other than Planet or SPOT.

NOTE: The code list related to Sentinel-2 provides the whole set of available bands. However, the bands B1, B9 and B10 are not expected to be used for EO monitoring as they refer to atmosphere and not to land.

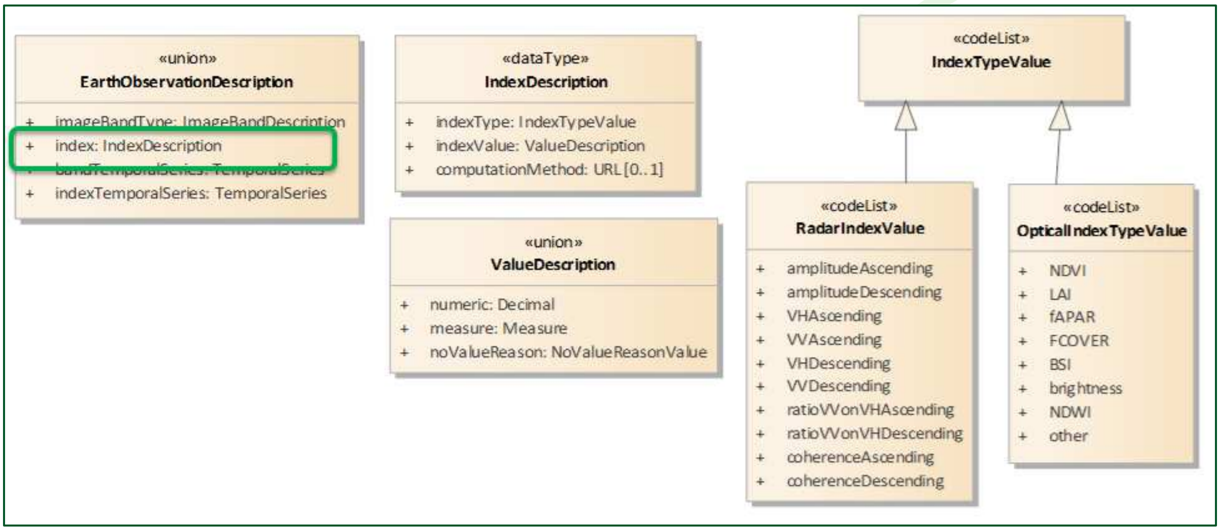


Figure 22: The case of indexes

The information about Index Description (at pixel level) consists in the index name (the code list Index Type Value supplies possible values for radar data and some of the most widely used indexes for optical images) and in the value itself; this value is generally a number (decimal) but some indexes are expressed as a measure (e.g. LAI).

In case of optical images, it may also occur that there is no valid information, due to various issues (in practice, mainly cloud) documented by the noValueReason attribute.

NOTE: Concerning radar indexes, the coherence applies in practice to a set of 2 images and may be computed from SLC products whereas the other indexes come from mono-date images in GRD products. For simplicity reasons, this distinction is not done in the model.

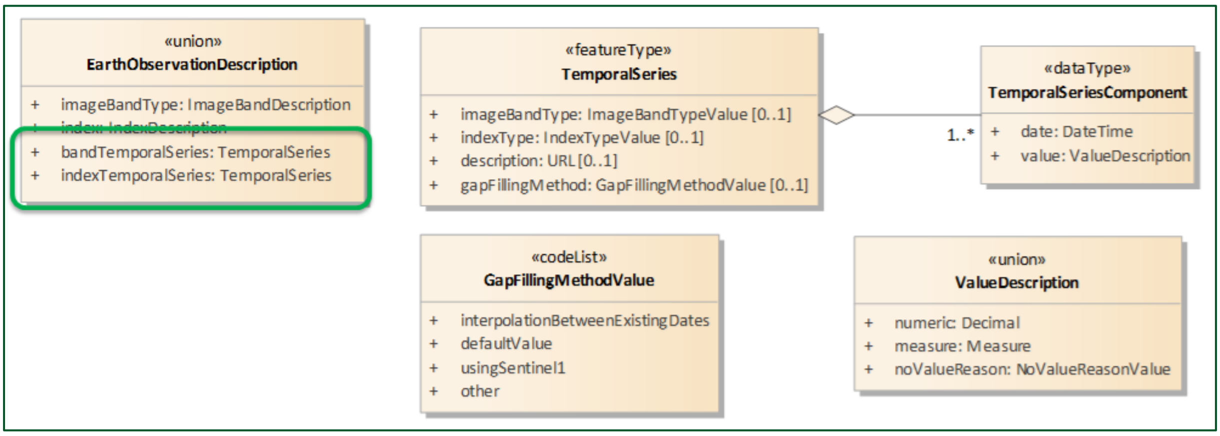


Figure 23: The case of temporal series

A temporal series is a set of temporal series component, i.e. of dates with corresponding value of a given satellite band or index.

These temporal series may be just set of raw data but they may be also prepared (e.g. by resampling in order to get regular frequency, such as every 10 days). This should be explained through the attribute “description”. A frequent case is the gap filling of Sentinel-2 series in order to mitigate the lack of observations due to cloud issue. A code list is proposed to document the method used to fill these gaps.

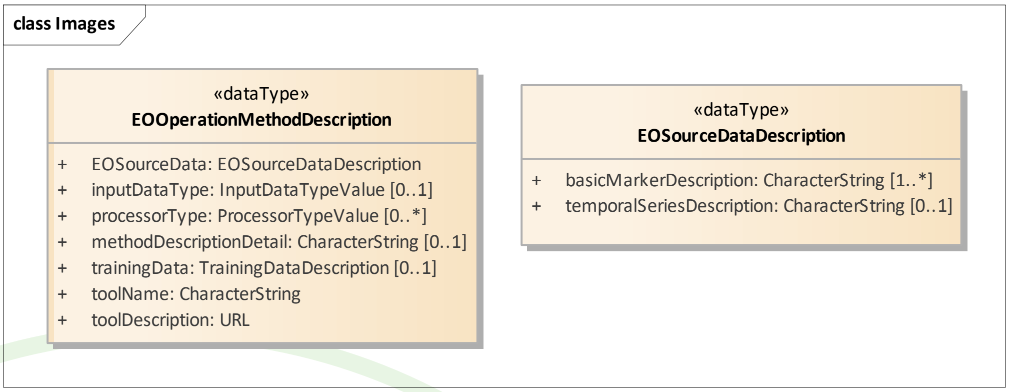


Figure 24: Main information about Earth Observation sources

The data type EO Source Data Description aims to provide a summary of the key information to be documented r for the description of an EO (monitoring) Operation:

- The description of basic markers, that may be an image band or an index; this description should include the marker name and the satellite (e.g. NDVI from S-2)
- A coarse description of the time series (e.g. raw observations /resampling, gap filling method if any).

4.6 Package Results

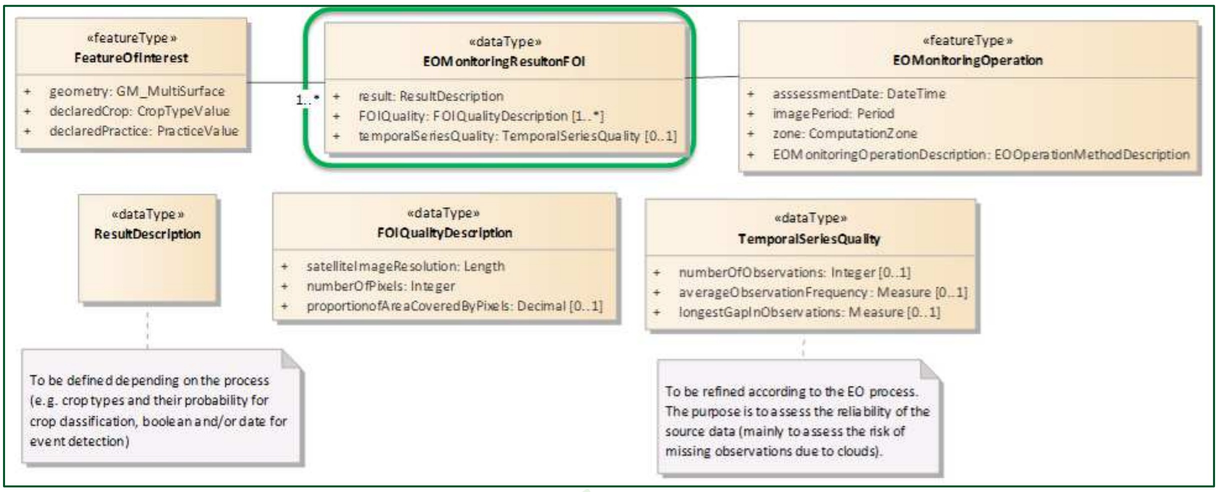


Figure 25: Results of EO monitoring Operations on Features of Interest

The results of an EO Monitoring Operation have to be recorded on each concerned feature of interest. This data model recommends recording not only the result itself but also some quality information to document the reliability of the process.

The parcel geometry is among the main factors impacting the result reliability, especially in case of small, narrow or complex shape parcels or other FOI. This is why the model proposes to document for each concerned satellite resolution, the number of pixels taken into account by the EO Monitoring Operation and the proportion of the area covered by these pixels compared to the whole area of the FOI.

In case of optical images, the characteristics of the temporal series are also a key factor impacting the result reliability, as there may be many gaps in observations due to clouds. The model is proposing a few examples of potential quality criteria of a temporal series. These criteria have to be adapted to the EO process. The NIVA experience has shown that this is not an easy task to get a simple quality indicator of temporal series.

NOTE: Some results should also be stored at more global level, for each EO Monitoring Operation. However, in general the relevant results to be stored at operation level are those coming from the EO Monitoring process but also from the decision process (how to convert EO results into traffic lights). This part is out of scope of this data model.

4.7 Package Crop Classification

This package aims to provide some more details on the crop classification that is a process widely used for EO monitoring.

- **IACS Data preparation**

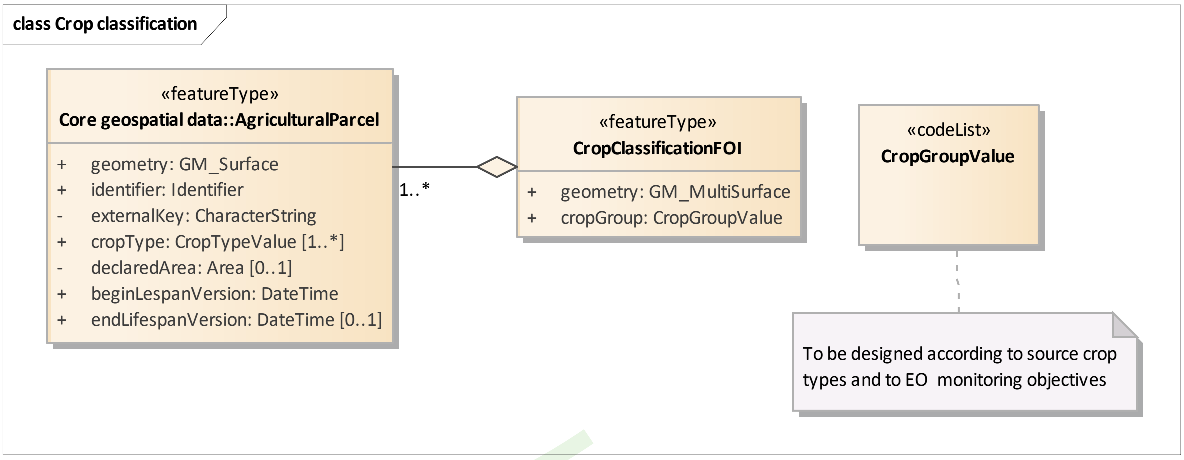


Figure 26: IACS data preparation in case of crop classification

In case of crop classification, the data preparation consists mainly (but not only) in grouping crop types into crop type groups. Agricultural parcels may also be merged into features of interest.

- **Training data**

For crop classification, the training data is generally coming from farmer declaration. This implies that all FOI may be used as training data; therefore, there are many possible options about the choice of the sample data to be used for training and calibrating the data model.

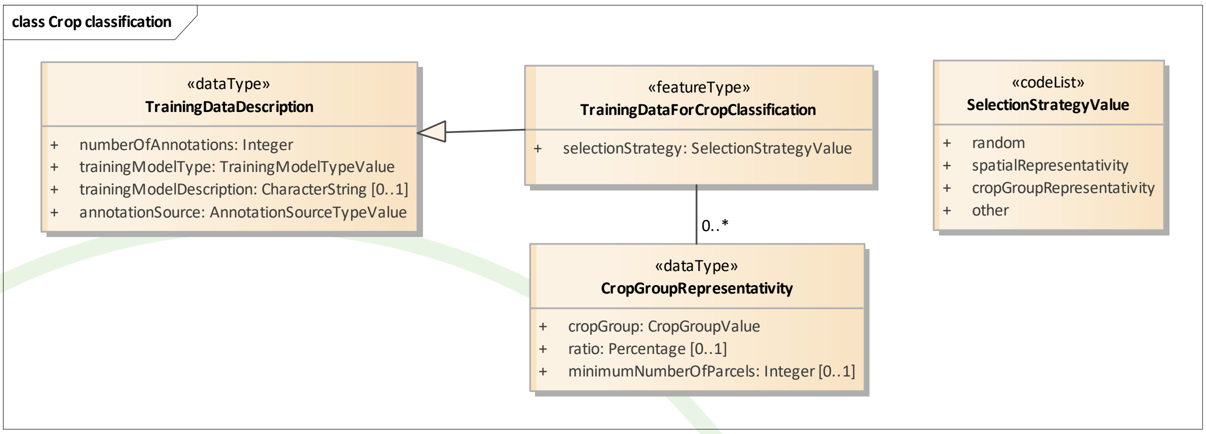


Figure 27: Selection of training data for crop classification

In addition to the common information of the Training Data Description, this model advises to document the selection strategy for crop classification. Examples of possible values are provided in the code list Selection Strategy Value.

A widely used strategy is to ensure that all crop groups all correctly represented in the sample data set. This is generally done by defining a percentage for frequent crop groups and a minimum number of parcels (or other FOI) for rare crop groups.

- **Results**

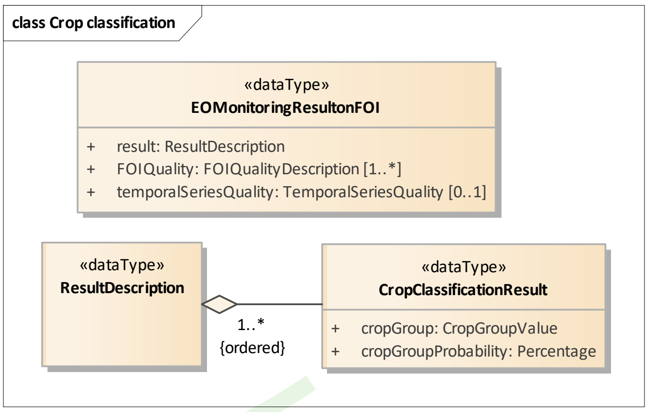


Figure 28: Results of crop classification at Feature of Interest level

The results of crop classification shall be provided on features of interest; this is generally done through an ordered list of crop group with its probability.

In addition, in the case of crop classification, it is also possible to provide some quality information about the crop classification itself. In practice, most of the "training" data is used for training itself whereas a smaller part is generally devoted to the machine or deep learning calibration. This calibration phase enables to generate a confusion matrix (as shown on figure29).

This confusion matrix may then provide quality indicators. The most frequently used are the overall accuracy (whole set of data) and then, for each crop group, the percentage of false negatives and false positives.

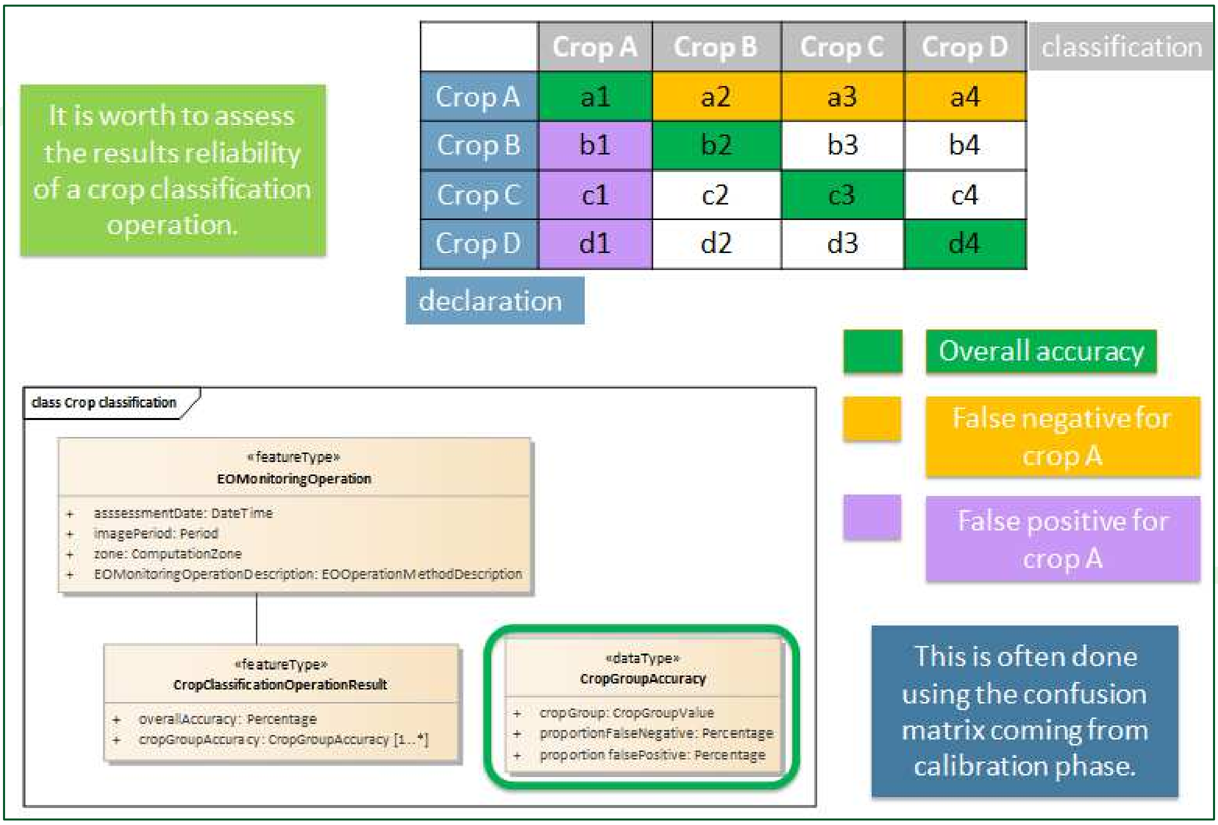


Figure 29: Results of crop classification at EO monitoring operation level

NOTE: Other metrics may be used to measure the quality of classification methods, such as:

- Recall : it is the ration between number well-predicted positive (true positive) compared to the number of positive that should have been detected (true positive + false negative)
- Precision: it is the ration between number well-predicted positive (true positive) compared to the total number of the positive that have been detected (true positive + false positive)
- F1 score that is harmonic average between recall and precision.

5. Exchanges between FMIS and IACS

5.1 Context

One of the objectives of the NIVA project is to provide common tools to prepare and implement the new CAP. In this context, it appeared that getting access to FMIS data would be very useful. The access to FMIS data was tested by several NIVA Use Cases:

- Machine data combined with FMIS data may be used as secondary evidences for the new Area Monitoring System (UC4b)
- FMIS data is necessary to provide the Farmer Performance results (UC1c)
- Some FMIs data is also necessary to compute carbon indicators Tier 2 and 3 (UC1b)
- The Farm Registry is expected to be used as common receiver of FMIS data (UC3).

Providing common solution to access FMIS data is quite challenging, as it raises several interoperability issues. The deliverable D3.5 Recommendations for standardized connections between IACS and other applications includes a chapter about exchanges between FMIS and IACS that is providing general information about FMIS and that has capitalized the NIVA experiences on this topic, until M24 (May 2021).

D3.5 Recommendations for standardised connections between IACS and other applications (niva4cap.eu)

This chapter aims to provide a quick and final overview about what has been done in the NIVA project about data exchange between FMIS and IACS with some focus on semantic interoperability but readers willing more details are strongly invited to look also at chapter 3 of D3.5 and to the tool documentation on the NIVA project GitLab: nivaeu · GitLab

The common data model for Farm Registry is described in detail in next chapter of this document.

5.2 Identifying required data

There are 2 ways to identify required data:

- First define which data is required by the Use Cases or more generally by Paying Agencies for CAP management and then investigate potential data sources, including FMIS

- Investigate what is (generally) available in FMIS and decide what could be useful for the Use Case or for Paying Agencies.

In practice, these methods are not exclusive and both have been used in the NIVA project.

Use Case requirements

The content of the table below provides mainly from the Use Case presentations done during Madrid meeting (in March 2020) or during virtual “Dublin” meeting (in May 2020). Therefore, this is only a provisory assessment of the FMIS data required from FMIS.

FMIS data requirements from NIVA Use Cases	
Use Case	Requirements for FMIS data
UC1a	The Decision Support System will have to take into the secondary evidence, namely geotagged photos and machine data that may be available through FMIS.
UC1b	To compute indicator on carbon storage and on nitrate lixiviation Tier II (and III), data is required about: - Use of fertilizers (organic and mineral) : nature + quantity - Export of straw (out of the farm): quantity
UC1c	FMIS data will be required for computing farm performance indicators. Indicators are not yet defined but some FMIS key topics have been identified: - Fertilizer use - Manure stack data - Plant protection product use
UC2	No data required from FMIS
UC3	The requirements are provided by the draft data model for Farm Registry. Key elements : - Producer - Farm identification - Location (crop parcels….) - Practices: crop labours and activities, fertilizers, phytosanitary products …
UC4a	UC4a provides an application for geotagged photos ; these geotagged photos may be stored in the farmer FMIS.
UC4b	UC4b is developing tools for transfer of machine data to FMIS and to IACS with focus on cases where EO data may be not enough to decide on eligibility: - Catch crop : o Which parcel has been processed? o Which crop (mixture) is sown? o When was it sown? o What amount has been sown per hectare? o When is the field plowed / the crop converted? - Strip crops: geometry + crop type
UC5a	A priori, no need for FMIS data
UC5b	UC5b would use data from Farm Registry for the Seamless Claim.

Table 5 – Estimated UC requirements for FMIS data

- **Paying Agencies requirements**

As leader of UC1b, ARIB has launched a questionnaire about data exchange between FMIS and IACS, targeted to PA. Answers were received from most NIVA partners and, in addition, from Austria, Czech Republic, 3 Lander in Germany, Luxembourg, Malta, Slovakia and Sweden.

Two questions were related to data content expected from FMIS by PA:

- Question 5: If data is being transferred from FMIS to IACS, what kind of FMIS data is shared with IACS?
- Question 10: From your perspective as a Paying Agency, is there any agricultural data on farm level which is of particular interest to you but is currently not readily available?

Main results are summarised below:

FMIS data requirements from PA		
Data	**Question 5** (current situation)	**Question 10** (wished in future)
Geotagged photos	Malta Spain (Valencia)	Germany
Use of fertilizers (organic and mineral)	Denmark Netherlands Slovakia Spain (Valencia)	Austria Estonia Spain (Catalonia – Navarra)
Use of plant protection products (e.g. pesticides)	Netherlands Slovakia Spain (Valencia)	Austria Estonia Netherlands Spain (Catalonia – Navarra)
Farming operations (with their dates) : ploughing, sowing, planting, showing, harvesting	Netherlands (cultivation dates) Slovakia (cultivation dates) Spain- Valencia (cultivation dates)	Austria Estonia Spain- Catalonia &Navarra
Specific cultivation methods (high tunnels, hydroponic, …)	Czech Slovakia Spain (Valencia)	
Pasture and grazing Livestock and pasture management	Slovakia Spain (Valencia)	Czech Denmark (grassland) Estonia Netherlands Spain - Navarra
Yield data	Slovakia	Yield data
Farm data (type of farming, economic size, income, productivity …)		Italy Netherlands
Management programs (soil, water, carbon)	Netherlands	
Precision farming data	Netherlands	Germany

Machine data	Slovakia Spain (Valencia)	Lithuania (machine data) Spain - Catalonia
Animal tracking system		Spain (Valencia)

Table 6 – FMIS data requirements from Paying Agencies

NOTE 1: Data already provided through the geospatial aid declaration, such as “geospatial data” or “crop types” have been mentioned several times but are not reminded in the above table.

NOTE 2: This table is not fully exhaustive. As the objective is to provide an overview, some aggregation has been done, with the risk of removing some details or to exclude unclear answers.

- **Available FMIS data**

Detailed analysis has been conducted in chapter 3.3 of deliverable D3.5 Recommendations for standardized exchanges between IACS and other applications.

In summary, all FMIS include a digital farmer calendar where farmers record their practices; therefore it is possible to find the TOP 3 requirements identified by the NIVA project: agricultural practices in general with more details on fertilization and Plant Protection Products applications. However, this data being generally captured manually is prone to errors: it is possible to use it to facilitate farmer declaration but its lack of reliability doesn’t make it suitable for control.

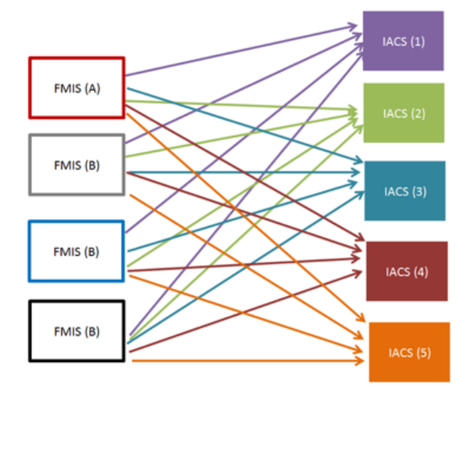
In addition of the digital farmer calendar, some FMIS have more specialized modules; in-field sensors and photograph applications looks the most promising for CAP monitoring.

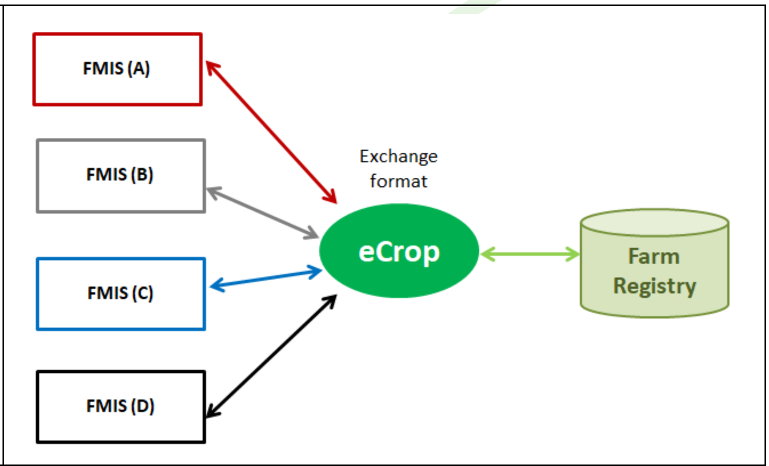
5.3 General principles

To enable common solution for data exchange between various FMIS and various IACS, the NIVA project has proposed:

- The use of the UN CEFACT eCrop standard for data exchanges (http://www.unece.org/fileadmin/DAM/cefact/brs/BRS_eCROP_v1.pdf)
- A common data model of Farm Registry that is expected to be a component of future IACS system.

The benefits of this standardised approach are shown in following figure.





Potential data exchanges between FMIS and IACs before the NIVA project	NIVA proposal for more standardized exchanges between (new) IACS and FMIS

Figure 30: Benefits of NIVA proposal for data exchange between FMIS and IACS

In practice, the eCrop standard is quite wide and rather flexible so it is necessary to use dedicated profiles when exchanging data between FMIS and IACS; this exchange is done through eCrop messages.

The profiles are mainly a selection of the features and attributes of interest but they may also include some extensions. In general, the Paying Agency willing to receive some specific FMIS data has to define the profile whereas it is up to the FMIS editor to implement it in its software export modules.

In addition, the eCrop standard does not include any code lists. Therefore, the choice of code lists to be used has to be agreed between the Paying Agency and the FMIS editor(s). In practice, the development and use of national or even better pan-European standards are expecting to facilitate this issue.

5.4 NIVA tools and associated eCrop messages

During its investigation of the data exchange between FMIS and IACS, the NIVA project has developed several tools and has tested them in the specific context of their Use Case or Work Package.

The tool documentation and the eCrop messages that have been defined are available on the NIVA GitLab. In case a Paying Agency is willing to exchange data with FMIS for a similar topic as in NIVA, it is advised to have look at the eCrop messages prepared by NIVA, to analyse them and if considered as relevant, to adopt or adapt them rather than beginning from scratch.

- **Tool developed by Use Case Farmer Performance (UC1c)**

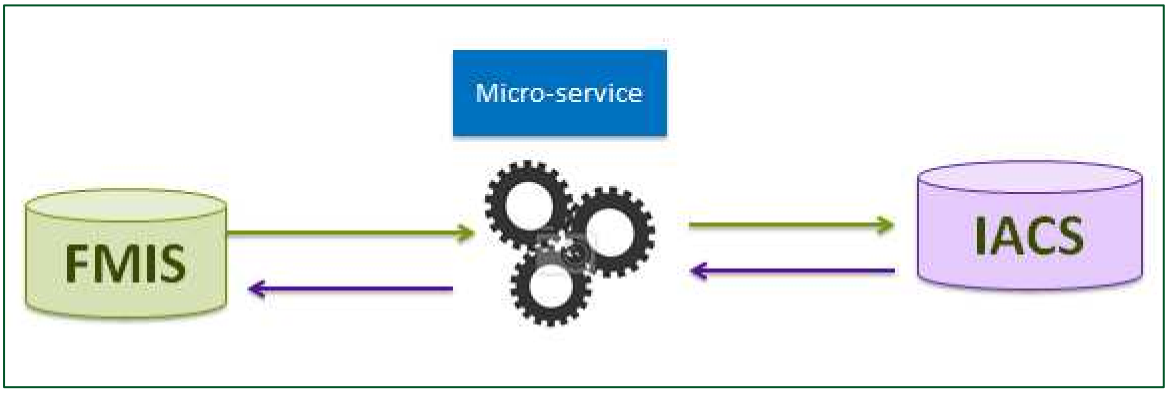


Figure 31: The IACS-FMIS exchange tool developed by UC1c Farmer Performance

This tool was designed with the main purpose to get in an easy and automatic way the FMIS data necessary to compute the information displayed by the Farmer Performance dashboard, i.e. the target data is about agricultural practices with focus on use of fertilizers and Plant Protection Products.

However, the tool may be used for data exchange in opposite direction, in other words from IACS to FMIS. In the UC1c context, it was tested to exchange data about geometry and properties of fields.

The tool has been developed and tested by the NIVA Estonian team. Profiles of eCrop messages are documenting in following document:

https://gitlab.com/nivaeu/uc1c-public-api/-/blob/master/ercrop%20doc/ecrop_fmis_iacs.pdf

For each direction of data exchange, this document provides:

- The general structure of the eCrop profile; the additional features or attributes are highlighted in a different colour. This is the case for instance of the reference parcel as this concept is very specific to IACS and generally not included in FMIS
- The description of the features and attributes; the table provides definition and explanations about the additional features and attributes and about the standardised ones used in a more specialised context
- Example(s) of eCrop messages with real data.

However, in a second step, UC1c decided to move to a different technical solution. eCrop message format gave inspiration for data exchange message structures, but eCrop is also too verbose for simple data exchange and has quite a bit of unnecessary elements and parameters whereas some other useful parameters are missing. In several cases, it can be easier and cleaner to just use a custom API. UC1c workgroup initially worked on adapting eCrop messages, but switched to a simple custom REST API. The API has been developed and tested by the NIVA Estonian team and is part of the farmer data dashboard. The available APIs can be seen and tested from the uc1c-backend Swagger page.

- **Tool developed by Use Case Machine data (UC4b)**

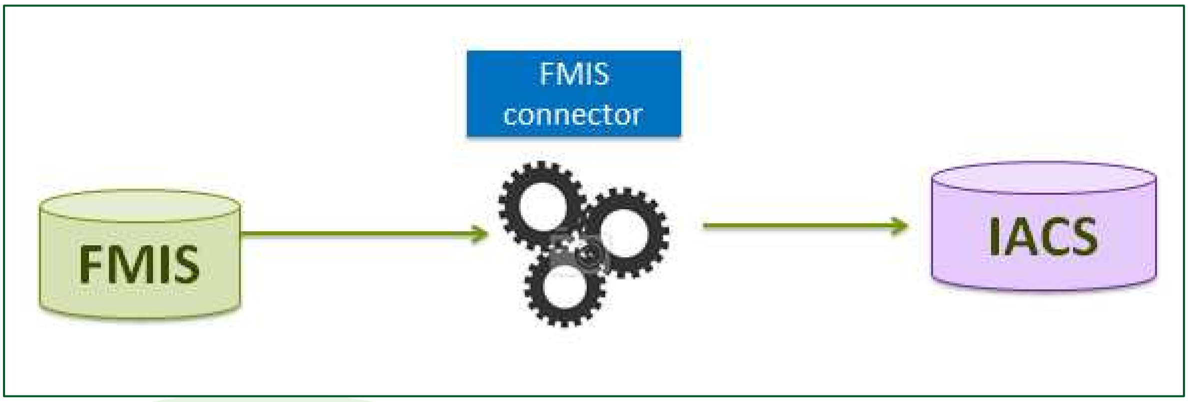


Figure 32: The IACS-FMIS exchange tool developed by UC4b Machine data

The tool developed by the NIVA UC4b is part of a wider data exchange flow:

- In first step, the machine as applied log data is sent from the machine to the FMIS (using ISOXML standard IS0 11783-10:2015 or generic shp-files) then combined with FMIS data to provide meaningful information.
- In second step, the resulting information is transmitted from FMIS to IACS using eCrop messages upon consent by the data owner (as shown in figure 32).

The concerned machine data is expected to be used for the new monitoring system of CAP payments, especially in cases when satellite images are not good enough solutions. This is for instance the case when there are strong temporal constraints (e.g. control of catch crops) or when the resolution of Sentinel images is not fine enough (e.g. control of strip crops that might be adopted as new eco-schemes to enhance biodiversity).

The tool has been developed and tested by the NIVA Dutch and Danish teams. Profiles of eCrop messages are available on: GitLab (table below)

The documentation includes:

- the conceptual document (designed model)
- json schema (implemented model)
- a document explaining the minor differences between them and the cause of these differences
- some anonymized sample UC4 Ecrop messages (based on SEGES testfarm).

https://gitlab.com/nivaeu/uc4b_ecropservice/-/blob/release-2/docs/schema/eCrop%20used%20by%20NIVA%20Final%20(conceptual).docx	This document includes the general profile used by NIVA and more specific ones used by UC4b (release 1 and release 2)
https://gitlab.com/nivaeu/uc4b_ecropservice/-/blob/release-2/schema/20210712-NivaMessageSchema-v2.json	JSON model for release 2 of the UC4b profile, it supports not only polygons but also point, multipoint …
https://gitlab.com/nivaeu/uc4b_ecropservice/-/blob/release-2/schema/ecropmessage-24-point.json	Concrete example provided by SEGES to exchange point data (personal information has been removed)
https://gitlab.com/nivaeu/uc4b_ecropservice/-/blob/release-2/schema/ecropmessage-24-polygon.json	Concrete example provided by SEGES to exchange polygon data (personal information has been removed)
https://gitlab.com/nivaeu/uc4b_ecropservice/-/blob/release-2/docs/schema/Differences%20message%20book%20and%20final%20JSOn%20schema.docx	Short document providing rationale for the differences.

Table 7 ECrop message documentation for machine data

NOTE: The differences between conceptual and implemented models have appeared during last phase of testing, i.e. too late to be corrected. However, these discrepancies are quite minor as they apply only to some changes in cardinality of elements.

- **Tool developed by task about common components (WP4)**

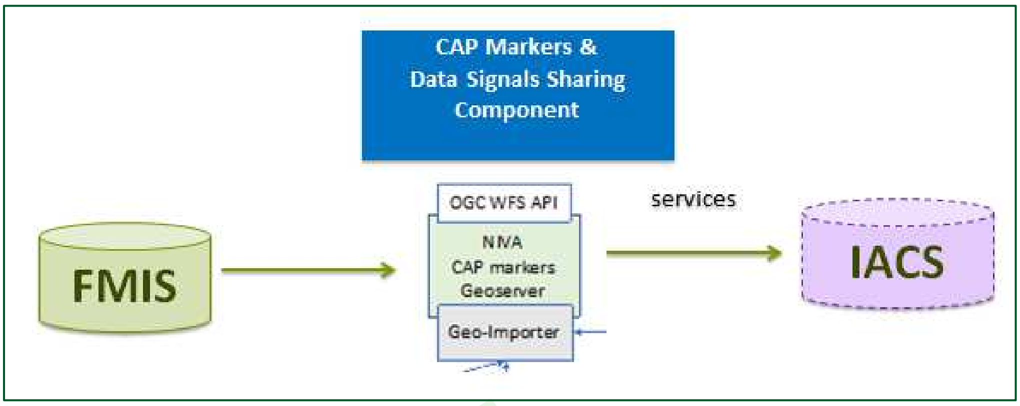


Figure 33: The CAP Markers & Data Signal Sharing component developed by WP4

This WP4 component was designed with the main purpose to provide an import module for the Decision Support System of Use Case UC1a about EO monitoring & Traffic lights: it is expected to enable easy access to CAP markers for EO monitoring but also to FMIS data and geotagged photos if case secondary evidences are required.

This component has been developed and tested by the Greek team. The FMIS export has to be designed by the FMIS edito, it may use the FMIS data model or a standardised model, such as eCrop messages.

6. Farm registry data model

6.1 Context and objectives

- **What is a Farm Registry?**

The Farm Registry is an official registry that should be integrated in IACS, due to the new CAP context and that describes and records where the farm is, what the farm produces and how the farm produces.

- **What are the objectives of the Farm Registry?**

The Farm Registry is one of the necessary IACS components for the future seamless claim system to be put in place by Paying Agencies, due to the new CAP requirements.

In addition, the Farm Registry may be used as source data to derive common indicators required by European policies or other external data, such as INSPIRE data on theme Agricultural or Aquacultural Facilities.

- **How is the Farm Registry supposed to work in practice?**

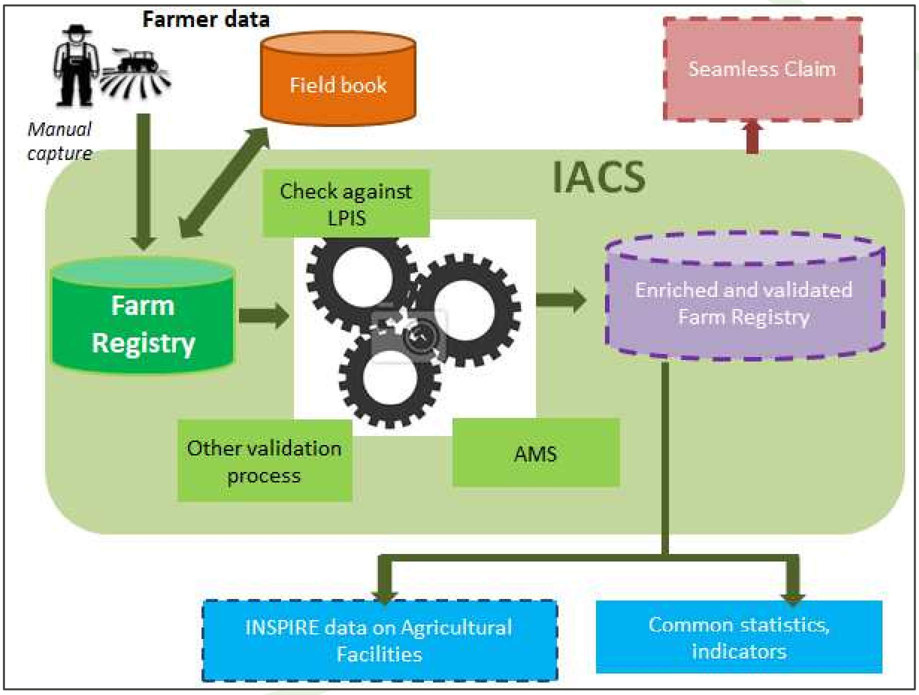


Figure 34: The objectives of the Farm Registry

The Farm Registry is mainly devoted to store the information received by farmers in a continuous way. So, in a first step this information is expected to come either from manual data capture from farmers or from automatic data exchange from Farmer Information Systems (Field Book, current FMIS solutions, FAST tools ….)

In a second step, the Farm Registry content is submitted to several validation processes such as comparison with LPIS for geometry checking or EO monitoring or other validation processes for checking the activities.

Once validated and possibly enriched, the Farm Registry should be integrated in the whole IACS where it will provide very useful information for the Seamless Claim system, enabling farmers to get payments without having to make a declaration. It may also be used to derive various output data (common indicators, INSPIRE data).

NOTE: The trend is to get data from farmers as automatically as possible. Several NIVA Use Cases have investigated the feasibility of automatic exchanges between IACS and FMIS (see previous chapter).

- **Why a common Farm Registry data model?**

The Farm Registry is based on common requirements mainly due to the new CAP, therefore it is obvious that all future national or regional Farm Registries in EU should have common content. The common data model aims to capture this common content in a structured way.

This common data model is expected to be a good starting point for Paying Agencies: when designing their future Farm Registry, they may build it by adapting the NIVA proposal rather than beginning from scratch. In other words, this common data model should be considered as an help for PA rather than as a mandatory standard.

- **How was this common data model designed?**

The Farm Registry common data model has been initially elaborated by the NIVA team working on UC3 (Farm Registry). This team lead by Spain has investigated what should be the content of a Farm Registry and it has proposed a physical data model for a relational databases, i.e. the data model is composed by a set of tables. More details about the UC3 data model may be found on the NIVA GitLab : nivaeu / uc3 · GitLab and in WP2 deliverables. In addition, the NIVA UC1c (Farmer Performance) has extended the UC3 data model in order to make it suitable for the derivation of its dashboard for farmers.

The WP3 on Harmonisation and Interoperability has recognized the value of this common data model and has decided to adapt it for integration in this deliverable. The content and modelling choices have been analysed, the UC3 and UC1c models have been compared to other models dealing with farm description, such as the INSPIRE model about Agricultural Facilities or its extension proposed by the FOODIE project. As result, the WP3 adaptation is mainly formal, main purpose being to propose a generic conceptual model.

6.2 Data model scope and principles

- **Scope**

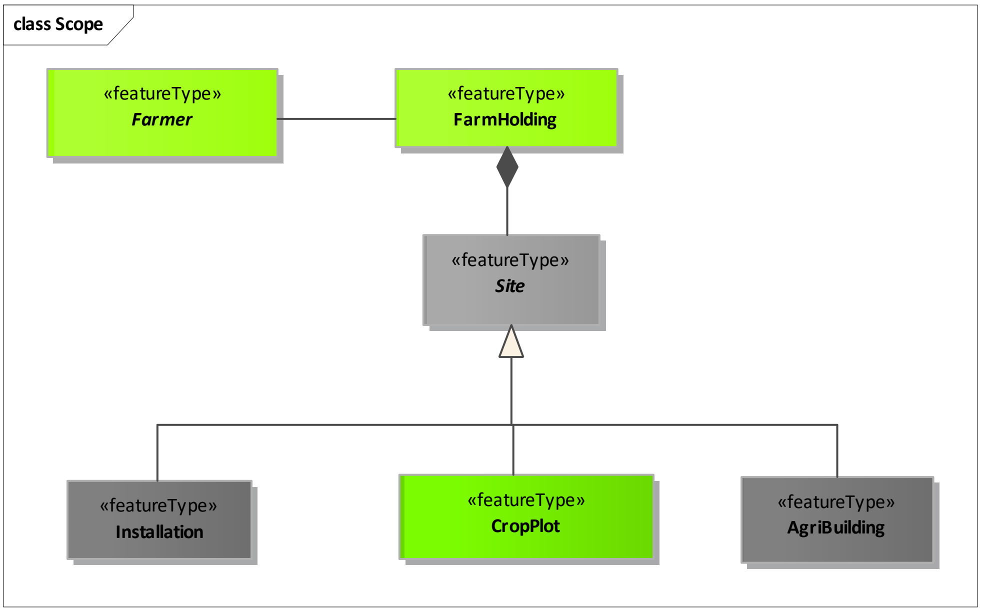


Figure 35: Comparison between NIVA and INSPIRE data models

According to the extended INSPIRE data model, a Farm Holding is composed of Sites that may be Agricultural Buildings, Installations or Plots, each site having the ability to carry information about activities and animals.

The NIVA model has a more limited scope as it is including only a simple description of the Farm itself, its relation with the Farmer managing it but it has a strong focus on the description of the Crop Plots and on the crop related activities.

- **Modelling conventions - Principles**

In the proposed data model, all the attributes have the public status shown by the "+" symbol. This is just a convention that does not imply that the whole content of the Farm Registry should be made publicly available. The issue of IACS data sharing is addressed in the WP3 deliverables dedicated to legal interoperability, namely D3.7 Guidelines to cope with legal issues and D3.8 Profile of priority data for external applications.

From the testing experience of UC3 data model, it appeared that not all the attributes will be available or even required in all countries. This may depend for instance of the eco-schemes put in place by the Member States and more generally about the choices of the National Strategic Plans. There is no indication of mandatory information in this data model, i.e. the model may adapted both by simplification (removing useless attributes) or by enrichment (adding more attributes or more detailed values in code lists). As a result, the multiplicity of attributes is just reflecting their existence in real-world and not if they are expected to be filled to all Paying Agencies.

The UML data model uses 3 main patterns:

- Feature types : they are the main entities in the data model, used for the classes of objects and storing the relatively static, permanent information (e.g. localisation) as direct attributes
- Data types have been used as complex attributes to provide information in a structured way (e.g. an address) or as classes attached to a feature type and providing dynamic information (e.g. crops, animals, activities)
- Code lists provide the list of possible values of an attribute.

The following colour coding has been used:

- Features types in the scope of NIVA data model are represented in **green (primary information**) ; features types outside the scope of NIVA data model are in grey
- Data types used as classes attached to a feature type and providing dynamic information (e.g. crops, animals, activities) are represented in **yellow (secondary information)**
- Other patterns (data types as complex attributes, code lists, notes) are represented in light **pink (details – last level of information)**

6.3 Data model content

6.3.1 Farmer description

- **Definition**

A farmer is a natural or legal person, or a group of natural or legal persons who exercises an agricultural activity and whose farm is lo is located in the territory of a European Union Member State.

- **UML diagram overview**

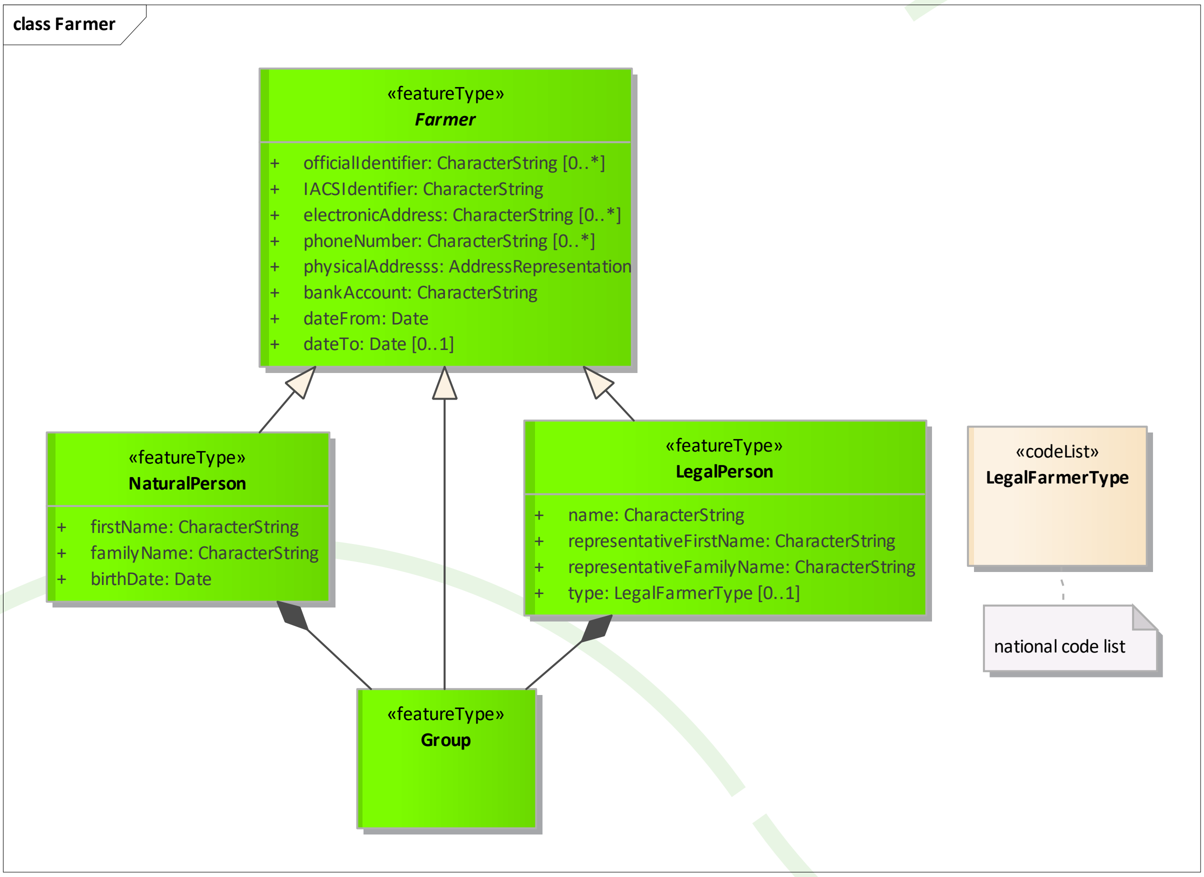


Figure 36: Information about Farmer

- **Model description**

In conformity with the definition, a Farmer may be:

- A natural person (NaturalPerson)identified by his or her first name, family name and birth date
- A legal person (LegalPerson) identified by its name, by the first name and the family name of its representative and if relevant, by its type (national code list documenting the possible legal types of moral persons in agricultural sector)
- A group (Group) of natural or moral persons.

The common attributes to these various types of farmers are factorised in the abstract feature type Farmer and are the following:

- Identifiers
  - Official identifier(s) generally coming from national official register(s) and enabling link with this register
  - IACS identifier used for the management of data by the Paying Agency. The IACS identifier of Farmer is expected to implement the link between Farmer and Farm. It is strongly advised to adopt an anonymous identifier, i.e. an identifier without any

personal information in order to be able to provide data about the farm without providing any data about the farmer

- Contact information
    - Electronic address(es)
    - Phone number(s)
    - Physical address.
- Bank account (for managing the payments)
- Temporal information: dateFrom when farmer is registered in the dataset and dateTo when the person is no longer a farmer (retirement, death, change of job …)

NOTE 1: The above model may be used to generate or to replace current IACS information about Beneficiaries.

NOTE 2: The above model is only about common basic information in Europe about farmers. It may be extended to include more specific information required in national context. For instance, in some countries, there may be need to add the father's name to the first and family name of natural persons.

NOTE3: The model includes the possibility that a farmer may be a group of natural or legal persons, as stated in the official definition of farmers. If this case doesn't occur in practice in a given Member State, the feature type "Group" may be removed from the model.

NOTE 4: The model proposes the INSPIRE data type AddressRepresentation as a standardised way to structure the address. However, in practice, the address may be modelled in a simpler way, according to the national address system. The eCrop standard is also proposing a common way to define addresses that may be used in replacement of the INSPIRE data type

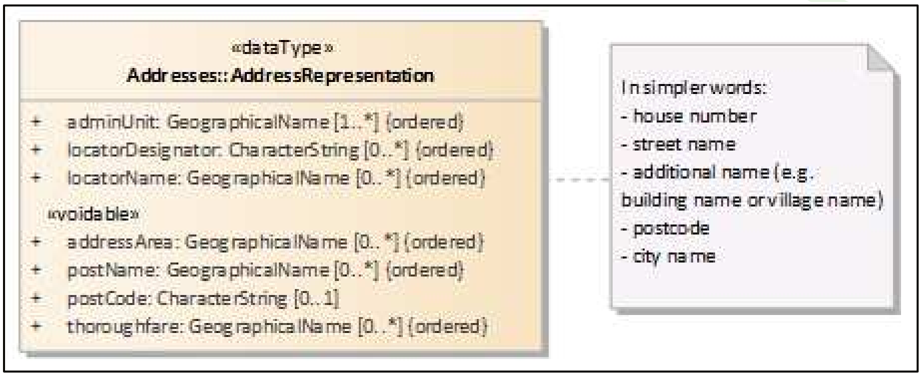


Figure 37: The INSPIRE modelling of addresses

6.3.2 Farm Holding description

- **Definition**

A farm holding is the set of all the sites used for agricultural activities, managed by same farmer and situated within the territory of the same Member State. It is a single unit both technically and economically; in general this is indicated by a common use of labour and means of production.

NOTE 1: There are unfortunately two official definitions of Farm Holding, one proposed by the EU regulation about Integrated Farm Statistics and one coming from INSPIRE theme on Agricultural and Aquaculture Facilities. The main difference between these 2 definitions comes from the notion of “single unit” that is present only in the Integrated Farm Statistics.

NOTE 2: The definition proposed by NIVA is an attempt to combine the advantages of the Integrated Farm Statistics (Farm Holding seen as a single economic and technical unit) and the advantages of INSPIRE (Farm Holding seen as a set of sites).

- **Basic information**

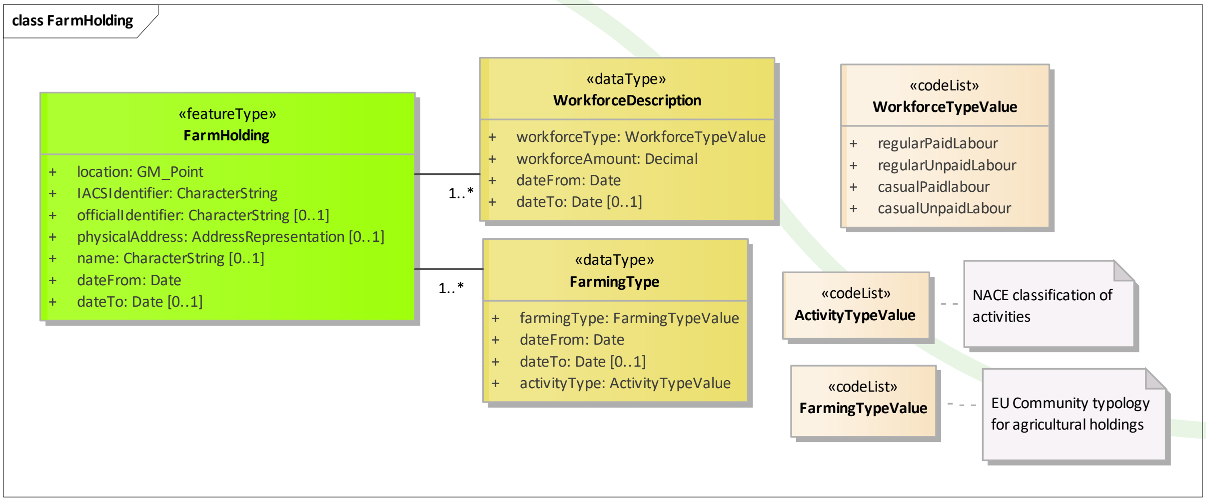


Figure 38: Basic information about the Farm Holding

The FarmHolding has following attributes:

- Location: coarse location of the farm as a point. This point is dedicated to facilitate the publication of farm related data. It may be the physical address of the farm or derives from the sites composing the farm (e.g. barycentre of the sites, centroid of the bounding box around the sites)
- IACS identifier (without any personal information)
- official Identifier(s): identifier of the FarmHolding in any other official information system (e.g. Business Register)
- a name (if it is the case in real-world) and a physical address; it is especially interesting to register the farm address if it is not the same as the farmer address. As for Farmer, the model is proposing the INSPIRE data type AddressRepresentation for this attribute type.

- temporal attributes that document the start and end date of the FarmHolding, respectively dateFrom and dateTo

The above attributes are expected to be permanent (except the dateTo that will be filled only when the Farm stops to be active).

In addition, more dynamic information may be attached to the Farm:

- Farming type that is documented using the EU typology for agricultural holdings
- Activity type that is documented using the NACE classification (Statistical Classification of Economic Activities) ; only the values from 0.1 to 0.5 and their sub-values are relevant (0.1: growing non perennial crops, 0.2: growing perennial crops, 0.3: plant propagation, 0.4: animal production, 0.5: mixed farming). In addition, some sub-values of 0.6 : support activities to agriculture and post-harvest crop activities are also possible.
- Workforce type that is documenting the type of workforce (paid/ not paid, regular/casual) and the annual amount, expressed in AWU. Regularly employed work force is converted into annual work units. One annual work unit (AWU) is equivalent to one person working full-time on the holding.

For both attributes, the model enables to record various successive values, due to the temporal information provided by dateFrom and dateTo.

NOTE 1: The FarmingType should come from the Community typology for agricultural holdings (type of farming)

A list is supplied in the UC3 documentation and may be used as example. An extract is provided below. However, it is based on Commission Regulation (EC) No 1242/2008 of 8 December 2008 (Commission Regulation (EC) No 1242/2008) but this regulation is no longer valid.

Table Name: CO_PFARMINGTYPE			
151	Specialist cereals (other than rice), oilseeds and protein crops	364	Specialist tropical fruits
152	Specialist rice	365	Specialist fruit, citrus, tropical fruits and nuts: mixed production
153	Cereals, oilseeds, protein crops and rice combined	370	Specialist olives
161	Specialist root crops	380	Various permanent crops combined
162	Cereals, oilseeds, protein crops and root crops combined	450	Specialist dairying

Table 8 Extract of the EU typology (2008 version)

A new regulation about Farm typology is under preparation and draft may be found here:

EUR-Lex - 51985GC0377 - EN - EUR-Lex (europa.eu)

Once consolidated and voted, this new Farm Typology should be used.

NOTE 2: This data model includes both the Farming Type (Farm Typology code list) that is used by NIVA UC3 and NIVA UC1c and the Activity Type (NACE code list) that is used by the INSPIRE data model.

The Farm typology list is more detailed than the NACE list used for the activity part. If it is possible to get easy matching from the Farm typology list (once voted) to the NACE classification, it would be quite logical to store natively only the Farming Type and to derive the Activity Type from this more detailed information.

- **Information about animals**

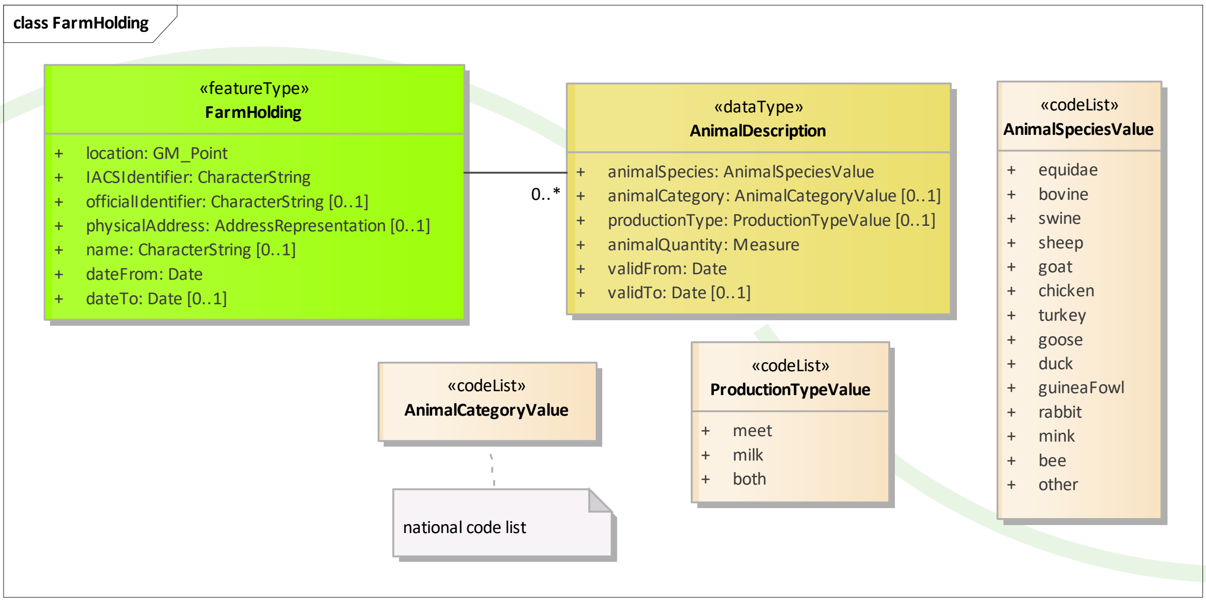


Figure 39: Animal information attached to the Farm Holding

Information about animals may be attached to the Farm Holding. There may be various kinds of animals, depending on their species, on their destination (for slaughter on not) and on their more detailed category.

For each kind of animal, the model proposes a data type describing:

- The kind of animal :
  - animalSpecies: the values of this attribute should be taken from a standardised national or European code list ; a potential example is provided by the model
  - animalCategory: this information is generally based on gender and age and is relevant mainly for cattle. The values of this attribute should be taken from a standardised code list, likely at national level

- productionType is about the animal production purpose. The proposed code list (meet, milk, both) is limited to the information that is generally available in Paying Agencies, i.e. information about cattle. It might of course be extended to include other animal productions (eggs, fur …)

- animalQuantity it will generally be expressed as the number of animals but in some cases, other units of measure may be more adapted (e.g. number of beehives for bees).
- dataFrom and dateTo provide the validity period of the information about animals.

NOTE: In the INSPIRE and the FOODIE models, the information about animals is attached to each Site (Buildings, Installation or Plot). This is a more accurate way to provide the information. However, the current model only addresses the information about animals that is or may be supplied in future to Paying Agencies and that is provided only at farm level.

6.3.3 Crop Plot description

- **Definition**

A crop plot is a continuous area of land, managed by same farmer, with the same crop on a given period of time.

- **Overview**

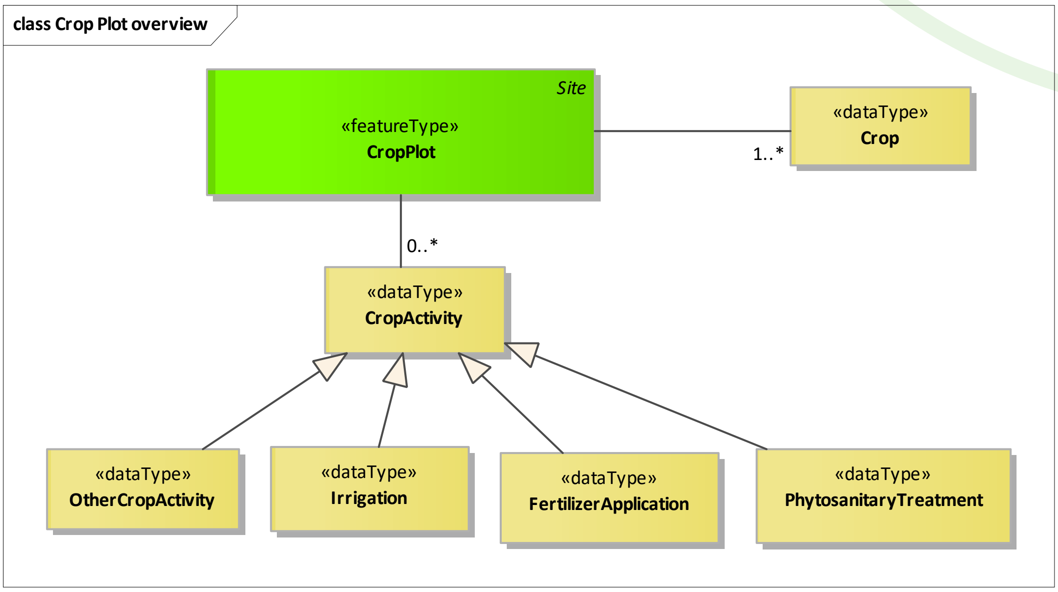


Figure 40: General description of Crop Plot

The feature type CropPlot captures the information that is relatively permanent, such as the geometry (see more details in next illustration).

Several crops may successively be associated to the CroPlot.

In a similar way, several agricultural activities may be associated to the CropPlot, such as irrigation, fertilizer application, phytosanitary treatment or other activities (OtherCropActivity).

NOTE: The data types corresponding to agricultural activities are attached to the CropPlot (geometry) rather than to the crop (Crop) for following reasons:

- The agricultural activity impacts mainly the current crop but some impact may remain for a longer term in the soil
- The payment is generally attached to the area ; so, attaching information to the geometry should make the payment checking and processing easier
- In case it would be necessary to attach this information to a specific crop, this might be done by post-processing by using the dates.

- **CropPlot attributes**

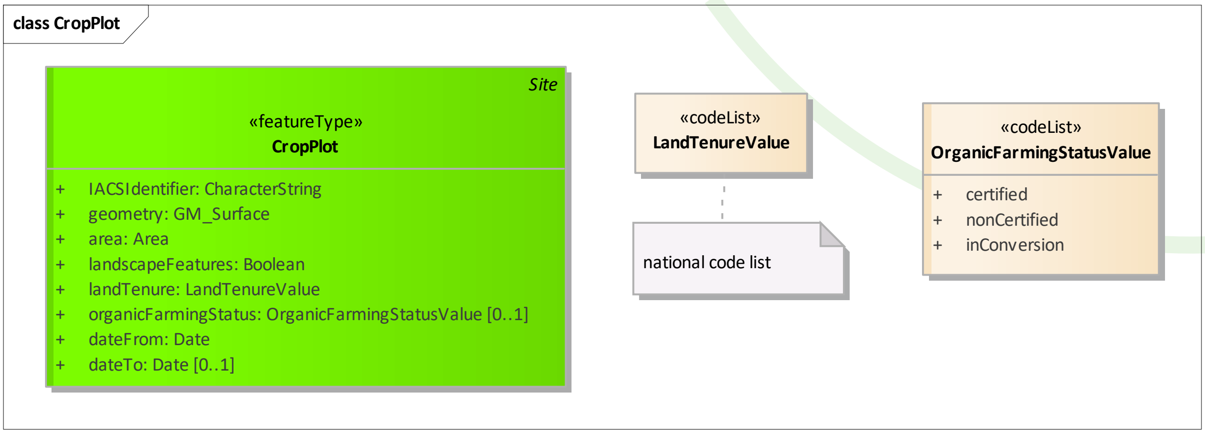


Figure 41: CropPlot details

The feature type CropPlot stores the relatively permanent information of crop plots, mainly the geometry, the presence of landscape features , the land tenure and the status of organic farming (if it is the case). The temporal attributes dateFrom and dateTo document the period of validity of the CropPlot feature. The attribute IACSidentifier is supposed to be an anonymous one, without any personal information about the farmer.

NOTE: The possible values of land tenure depend on national regulations; this is why a national code list is proposed by this model. An example of possible code list is provided in the UC3 data model.

6.3.4 Crop description

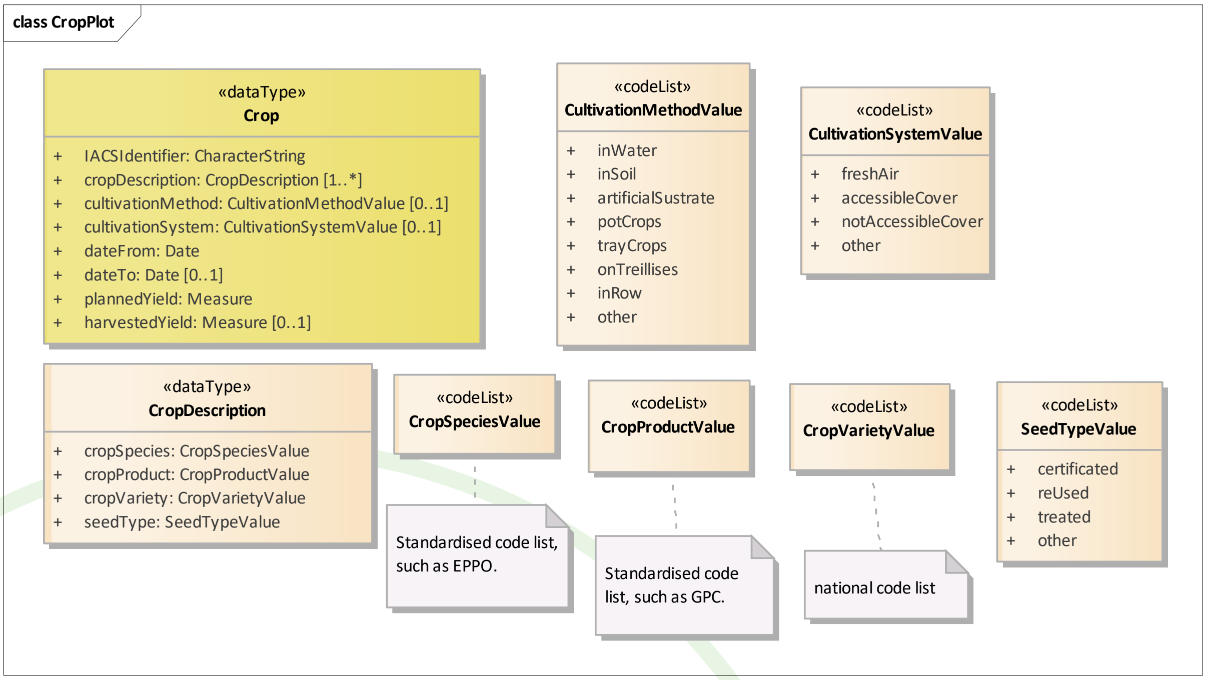


Figure 42: Information about the Crop

The data type Crop represents the crop cultivated on a given CropPlot feature on a given period of time. It includes the following attributes:

- Crop description
    - cropSpecies : multiplicity of crop species is [1..*] as there may be several species at same time, e.g. in case of catch crops
    - In addition, it is possible to document the crop product, the crop variety and the seed type
    - For the cropSpecies and cropProduct, it is recommended to use the standards identified by previous version of this deliverable (D3.2 - M12), i.e. respectively EPPO and GPC. As these attributes are not independent, a better way would be to use a combined standard, such as AgroConnect (Dutch standard). However, AgroConnect would need to be extended to include all crops in Europe.
    - For the cropVariety, the model recommends to use official national code lists; the variety depends on the species so some constraints should be added in the model and checked automatically to avoid inconsistent data capture
- Information about cultivation system and method
- Information about yield: plannedYield and harvestedYield; both are expressed as Measure, i.e. as a number and an unit of measure.

- Temporal attributes (dateFrom and dateTo) document the period when the crop has been cultivated or is being cultivated.

6.3.5 Irrigation description

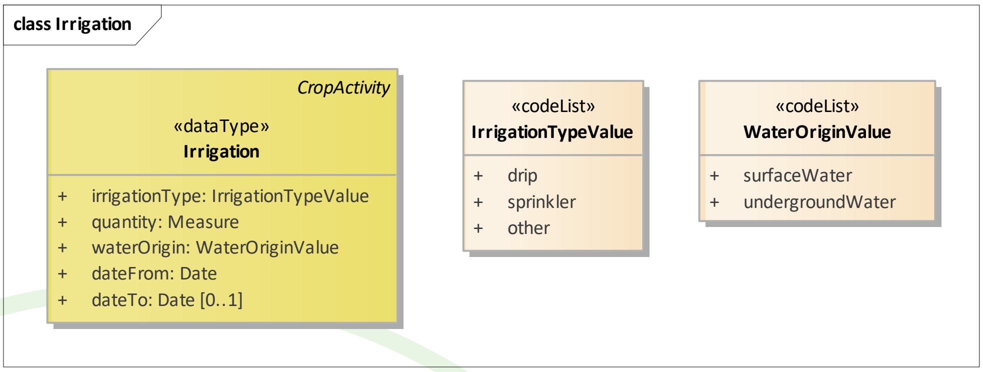


Figure 43: Information about irrigation

The data type Irrigation includes the following attributes:

- The period when the irrigation takes place provided by the "dateFrom" and "dateTo"
- The "Quantity"of water is expressed as a measure : it might be a volume (quantity of water on parcel) or a volume/ unit of surface (e.g. m^3/ha)
- The irrigation type indicates the technology used for irrigation (drip or sprinkler or other)
- The waterOrigin provides information about where the xwater was taken from (surface or underground).
-

6.3.6 OtherCropActivity description

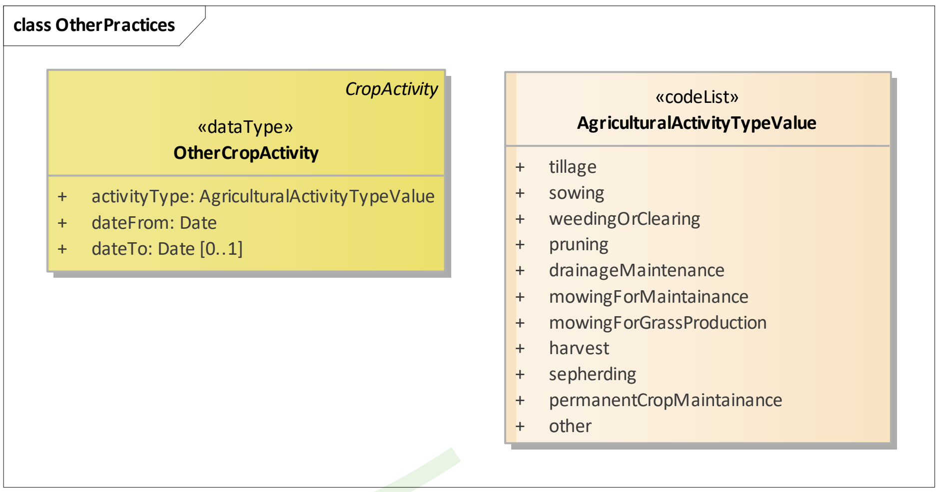


Figure 44: Information about the Crop

The data type OtherCropActivity concerns the activities others than (other than irrigation, fertilization or phytosanitary treatment) ; the data type includes the following attributes:

- The period when the activity takes place provided by the "dateFrom" and "dateTo"
- The type of the activity; possible values are provided by the code list AgriculturalActivityTypeValue.

6.3.7 FertilizerApplication description

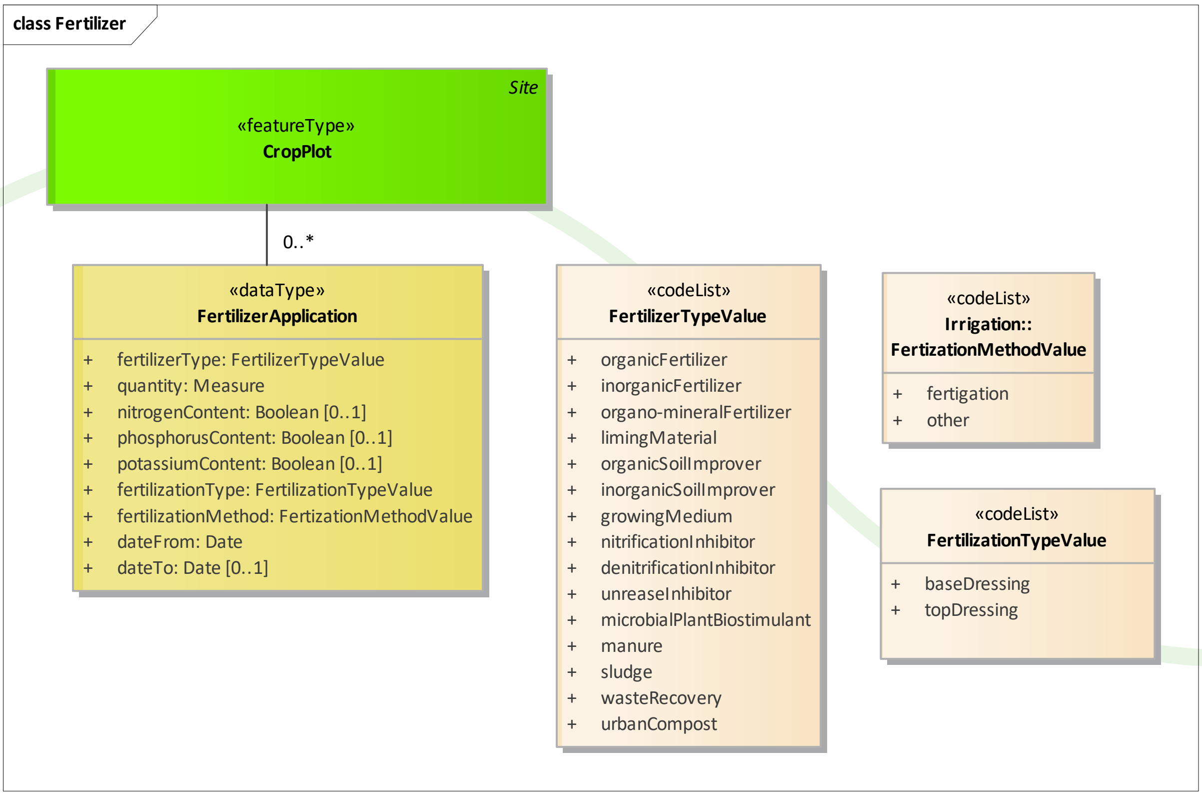


Figure 45: Information about the FertilizerApplication

The data type FertilizerApplication includes the following attributes:

- The type of fertilizer that has been used; possible values are provided by the (simple) code list FertilizerType Value
- In addition, it is possible to provide information if the fertilizer contains phosphorus, potassium or nitrogen
- Information about the fertilization practice:
  - The type indicates if fertilization was applied as top dressing or as base dressing
  - The method indicates if fertilization was applied combined with irrigation (fertigation) or by other means

- The attribute Quantity is expressed as a measure (i.e. a number and a unit of measure). It may be the total amount on CropPlot (as volume or mass) or the amount per unit of surface (e.g. a volume/ha or a mass/ha).
- The period when the irrigation takes place provided by the "dateFrom" and "dateTo"

NOTE: The UC3 data model proposes another attribute called "fertilizer" that provides quite more detailed information about the nature of the fertilizer. If considered as relevant, this detailed attribute may of course be added to the proposed model.

6.3.8 PhytosanitaryTreatment description

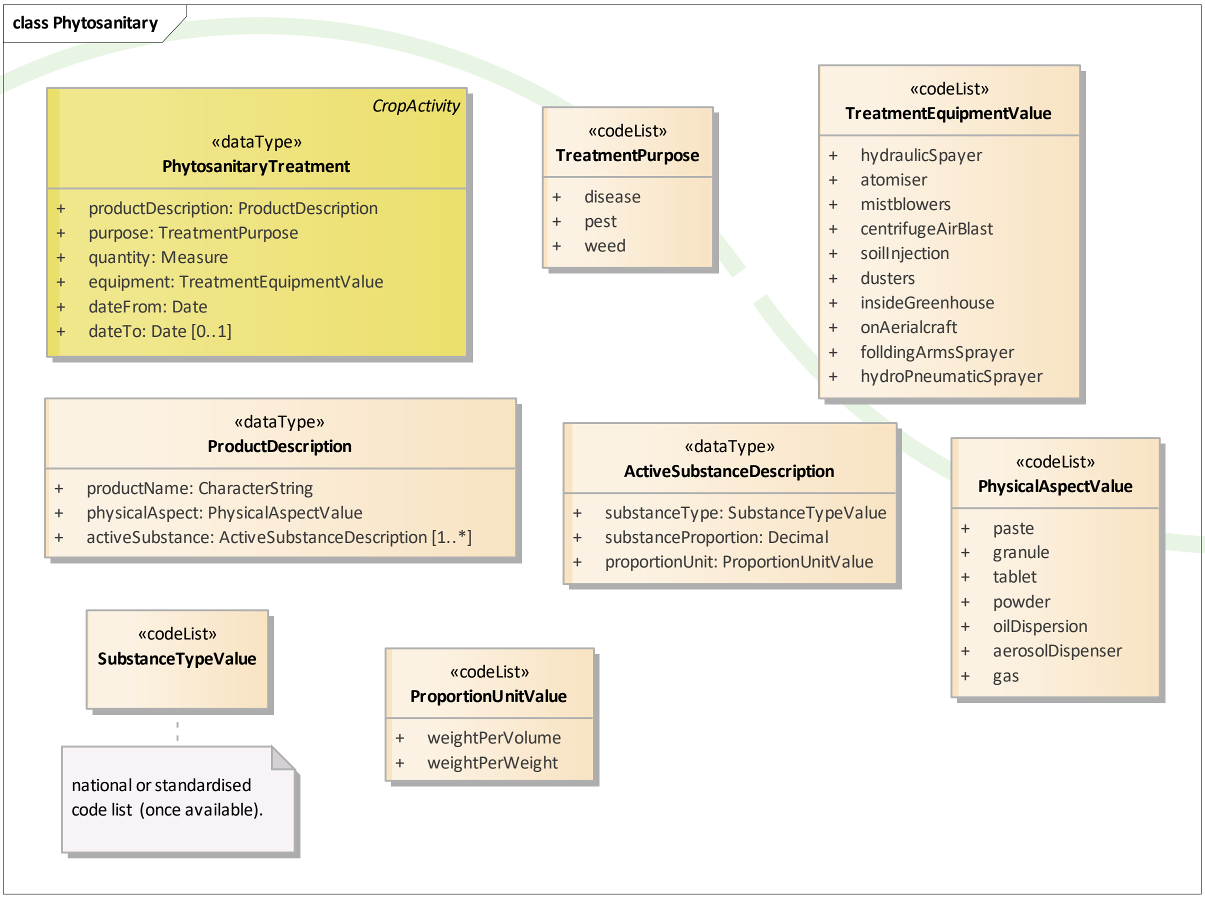


Figure 46: Information about the PhytosanitaryTreatment

The data type PhytosanitaryTreatment includes the following attributes:

- The description of the product that is itself composed of several sub-attributes
    - Name of the product (as free text)
    - The physical aspect (paste / granule / tablet / …)

  - The description of the active substance(s) present in the product: substance type, proportion (as decimal) and the proportion unit (weight per volume or weight per weight)

- The treatment purpose: the phytosanitary product may aim to fight against disease, pests or weeds
- The equipment used in the treatment
- The quantity of product spread on the CropPlot expressed as a measure (number + unit of measure)
- The period the phytosanitary treatment has taken place, through the attributes dateFrom and dateTo.

NOTE: The UC3 data model is proposing a more compact way to describe the product composition (the active substances used and their proportion).

6.3.9 Overview - Summary

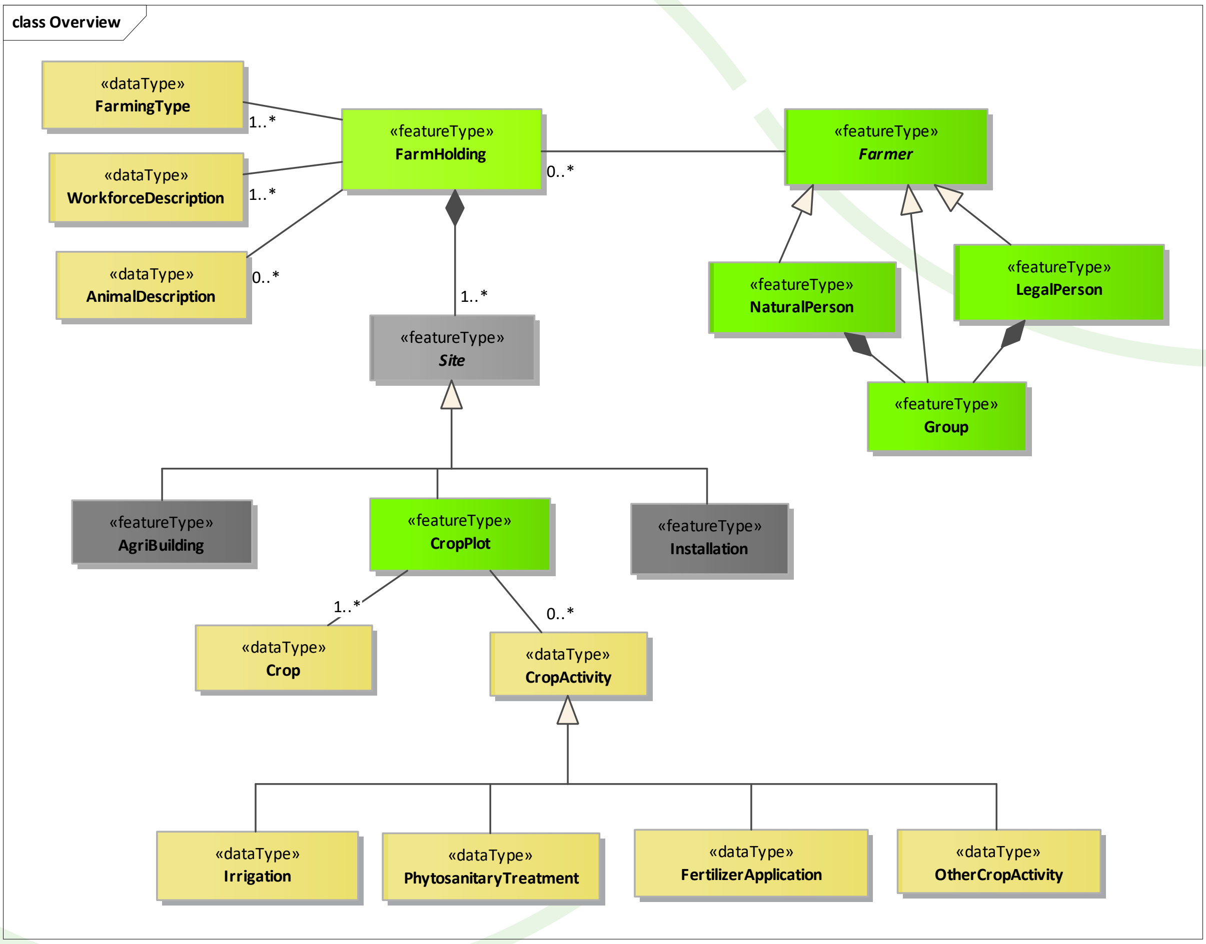


Figure 47: Overview of the Farm Registry data model

A farmer may be a natural person, a legal person or a group of natural or legal persons.

A farmer can manage any number [0..*] of Farm Holdings:

- It may be several (*) Farm Holdings as the Farm Holding is defined as an economic and technical production unit
- It may be none (0) FarmHolding in case of “deprecated” farmers whose data are still stored in the system though farmers have ceased their agricultural activities.

The FarmHolding feature type stores a few (relatively) permanent attributes such as identifiers but the dynamic information is stored in 3 data types:

- ActivityType
- WorkforceDescription
- AnimalDescription

As the FarmHolding is defined as an economic and technical production unit, it has normally only one ActivityType for a given period of time. However, this ActivityType may vary along time; this is why the multiplicity [1..*]is allowed, enabling to store historical data in the Farm Registry.

Whereas all FarmHoldings have at least an ActivityType and a Workforce to be described, not all of them are raising animals. This is why the multiplicity of AnimalDescription is [0..*].

The FarmHolding is composed of several (i.e. [1..*]) Sites that may be AgriBuildings, CropPlots or Installations. However, the NIVA data model deals only with the CropPlots.

The feature type Crop Plots stores the (relatively) permanent attributes, mainly the plot geometry but the dynamic data is stored in various data types: one for the Crops that are cultivated along time on the Crop Plots and a few ones for the Crop Activities : more details are provided for Irrigation, Fertilizer Application and Phytosanitary Treatment whereas a simpler data type describes the Other Crop Activities.

7. Conclusions

7.1 Looking back to the past

- **What has been achieved**

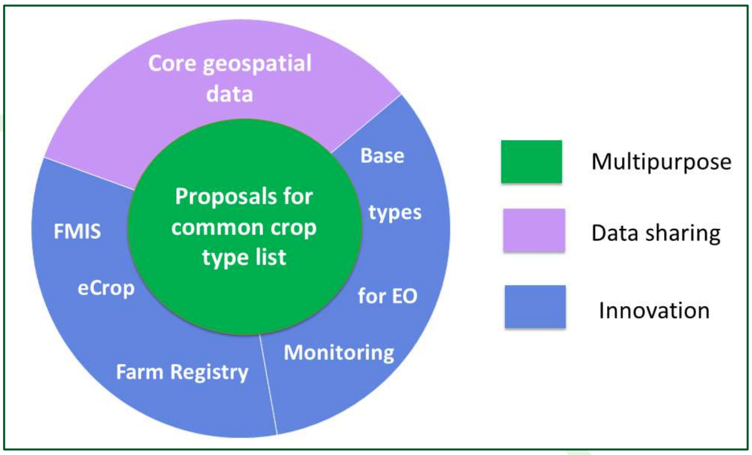


Figure 48: Overview of semantic interoperability achievements

Through the delivery of the three versions of D2.3 Common Semantic Model, the NIVA project has provided some solutions to facilitate IACS geospatial data sharing and to prepare the innovations required by the new CAP, as shown by figure 48.

- **Main lessons learnt**

Reduced scope but more benefits

The initial objectives of D3.2 Common Semantic model were to improve semantic interoperability of the IACS data required by the NIVA Use Cases. At the beginning of the project, the scope was, of course, rather wide and blurry as the Use Cases were defined only at very high level.

During the project time, it was been decided to focus the harmonization efforts on some key topics being of interest for several Use Cases rather than trying to deal with all the data required by all the Use Cases.

This choice has had two main advantages. The first one was obviously to reduce the amount of required work and to ensure feasibility. The second and main one was to make the recommendations less depending on the components developed by the Use Cases. In other words, the recommendations for semantic interoperability are expected to be helpful for any Paying Agency even if this Paying Agency is not using the NIVA components.

Different aspects of interoperability

The tasks of Work Package Harmonisation and Interoperability were organized according to the three main aspects of interoperability, i.e. semantic, technical and legal interoperability.

This organization of work was quite logical and useful at the beginning of the project; nevertheless, it appears to raise some issues when the project has progressed as some topics are concerned by several aspects of interoperability:

- Core geospatial data : semantic and legal interoperability
- Data exchange between FMIS and IACS: semantic, technical and legal interoperability
- EO monitoring: technical and legal interoperability.

As a result, the information about the key topics considered in the Common Semantic Model is scattered between various deliverables, what is not user-friendly. The reading advices presented in the introduction aim to mitigate this issue.

Moving context and objectives

The main purpose of the NIVA project was to develop innovative solutions to prepare the new CAP. However, in practice, these innovative solutions had to be tested using current IACS data.

Therefore, at the beginning of the project, the trend was about struggling for harmonisation of existing IACS data in order to support the Use Case teams. This is reflected in the first version of the Common Semantic Model.

Following the Use Cases progress, it became clearer that the goal of the Common Semantic Model is to propose solutions for the future IACS, the information system that will enable the implementation of the new CAP requirements. In other words, the purpose of the Common Semantic Model is to structure IACS data as they should be in near future; however, this near future has remained relatively blurry as the requirements linked to the new political context were not provided from the beginning of the project.

NIVA data model: priority to the physical world

A conceptual data model is expressing the Universe of Discourse, i.e. it is structuring the required information according to a specific point of view.

At the beginning of the project, the JRC data model of current IACS was considered and carefully investigated. Whereas the first version of the NIVA Common Semantic Model was rather encouraging the re-use of this model (or of a simplified version of it), this definitive version is clearly proposing a different point of view.

The JRC data model is mainly based on administrative considerations: the entry points of the structure are the requirements for payment administration and controls. In the way, this is quite logical as it is the main aim of IACS. However, this choice has some drawbacks as it leads to some complexity and even duplication in the modeling of the information coming from physical world (farmer activities) and necessary to check the CAP payments.

The NIVA project is proposing another approach where the entry point is the description of the physical world that is considered as the raw part of the system: Paying Agencies should record first the "real-world" data (LPIS, farmer data) and then they can apply a set of processes on top of this raw data to decide on the various CAP payments. In addition, this physical world point of view may also facilitate IACS data sharing by making IACS data both more understandable and less sensitive (as no information about payments).

7.2 Looking towards the future

- **Consolidation and adaptation/adoption of NIVA proposals**

The Common Semantic Model targets the community of stakeholders dealing with IACS data and processes, such as Paying Agencies and their technical partners, the European Commission (JRC, DG AGRI) and researchers.

The recommendations of this deliverable (and associated documents) may be implemented on a voluntary basis by Paying Agencies. The last Stakeholder Forum meeting of the NIVA project has provided a good opportunity to raise awareness.

However, the likelihood of adoption would greatly increase if these recommendations are recognized as relevant by the European Commission and more or less imposed to Paying Agencies.

More especially, the work about core geospatial data and common crop type list could be reused by DG AGRI for guidelines about IACS data interoperability (the current guidelines being dedicated to accessibility) and the data model about "Base types for EO monitoring" might be reused by JRC for setting up the Quality Assurance of the new monitoring system (for instance, common templates for reporting might be derived from the model).

- **Extension of NIVA proposals**

Wider scope

As explained in previous chapter, the NIVA Common Semantic Model is addressing only a few parts of the new IACS to be set up in order to implement the new CAP.

A more exhaustive data model would be quite useful for Paying Agencies, in a first step to provide common understanding and in a second step, for implementation guidelines. This is a potential task for a follow-up project; an update of the JRC data model might also be envisaged as first step.

Other interoperability tools

The Common Semantic Model is mainly composed on UML diagrams. The names of feature types, attributes, values of attribute have been captured in the data model but not their definitions. In general, the terms used in the data model are self-understanding; in addition, most of acronyms are defined in the Common Glossary and there is some explanatory text below the figures when it has been considered as necessary.

In other words, this model focuses on the data structure but does not provide a systematic recording of the definitions, as it would be the case with a Feature Catalogue or with an ontology.

The Open IACS project has worked on the publication of some IACS data (mainly LPIS – GSAA) as Linked Open Data. There would likely be place for progress by combining the experiences of NIVA about data harmonization and of Open IACS projects about semantic web technologies.