Spatial information in European agricultural data management. Requirements and interoperability supported by a domain model

Katalin Tóth *, Andrius Kučas
European Commission, Joint Research Centre, Institute for Environment and Sustainability (IES), Monitoring Agricultural Resources Unit—H04, TP266, 26B, Via Fermi 2749, I-21027 Ispra, VA, Italy

A R T I C L E   I N F O

Article history:
Received 7 December 2015
Received in revised form 23 March 2016
Accepted 21 May 2016
Available online 4 June 2016

Keywords:
IACS
Domain model
LPIS
Interoperability
Policy modelling
UML

A B S T R A C T

Data compatibility and system interoperability are fundamental for crosswalks and collaboration between domains. The most frequently used references for information sharing are time and location. In order to understand the requirements, fundamental processes, and core information concepts of a domain, a comprehensive, but standardised documentation is needed. In spatial data infrastructures models presented in Unified Modelling Language (UML) are widely used to facilitate the uptake of standards and valorise best practices of various communities.

The European agricultural decision makers must deal with many components described in the Common Agricultural Policy (CAP) in order to optimize data integration and achieve transparency. Geospatial information plays a key role in the implementation of this policy, which comprises the establishment and maintenance of the Integrated Administration and Control System (IACS).

In the past the IACS of the MS has principally served a single high level business case: to run a correct administration of the CAP. However, the recent reform and the synergies with environmental and societal policies create increasing expectations for IACS, which impact upon system interoperability and data usability. These objectives can be achieved by establishing a framework that is standard based, allows flexible extensions, and that supports efficient implementation and information exchange between the stakeholders.

This paper presents the development process and describes the structure of a domain model we propose for IACS. This standard driven model was designed to translate requirements into technical elements, ensuring interoperability, and providing flexible extensions at the same time. For this purpose we set-up a formal requirement model, formalized use cases, and integrated ISO/TC 211 and INSPIRE UML based class diagrams. As such, this paper proposes a methodology to help guide how the policy can be implemented.

The importance of traceability from IACS and third party business rules to the information concepts was proven by simulation runs. The developed modelling approach yielded a reference for conformance testing, indicated critical points of potential errors, and highlighted the impact of eventual changes. Storing all concepts and implementation options of IACS in a unique framework helps to eliminate redundant efforts, provides a strong basis for developing various applications, underpins interoperability with other domains and enhances transparency of the CAP.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

The Common Agricultural Policy (CAP), being one of the oldest policies of the European Union, aims at maintaining a stable supply of safe and affordable food to the consumers and ensuring that EU farmers can make a reasonable living whilst preserving the rural landscape and natural environment (European Commission, 2012). The objectives, tools, and measures for the implementation of CAP are laid down in the legal acts, adopted by the European Council and the Parliament and, by delegation, by the Commission. Member States are responsible for the implementation of CAP, which comprises the establishment and running of the Integrated Administration and Control System (IACS).

The recently adopted new legal framework of CAP (European Parliament and the Council of the European Union, 2013a,b,c; European Commission, 2014a; European Commission, 2013a; European Commission, 2014b) sets up genuinely new require-
ments promoting agricultural practices beneficial for the climate and environment. These new requirements are also referred to as the “greening” of the CAP. Greening and the introduction of the geospatial aid application for farmers set new challenges to spatial information management in the domain.

Council Regulation (EC) No 73/2009 (Council of European Union, 2009) referred to the identification system of agricultural parcels as an information system established on the basis of maps or land registry documents or other cartographic references that make use of computerised geographical information system techniques. The user community translated this formulation as “the single GIS of IACS” (European Commission DG JRC, 2015b) and introduced the term of LPIS to identify the corresponding spatial dataset. LPIS has to support two fundamental processes: the application of the farmers and the controls of authorities, by localisation and quantification of agricultural land eligible for the EU support. This approach remained in place after the 2013–2014 year CAP reform.

Relating information to location plays a fundamental role in agriculture. Beyond such classical fields as agro-statistics, crop monitoring, land registry, or the management of the CAP this method is becoming more and more pertinent (Pelorosso et al., 2009). Among the emerging fields we can mention food security (Singh et al., 2014; Candele et al., 2014), rural development (Van Berkel and Verb, 2011; Pašakarnis et al., 2013), management of specialty agricultural crop management (Oczelik and Nisanci, 2015), precision farming (Möckel 2015; Réznik et al., 2015), and analysis of agro-ecosystems (Wainger et al., 2010). The greening of the CAP adds a further significant area where spatial data plays a key role.

Translating greening into spatial data terms requires the extension of the LPIS by new feature types to capture objects that provide ecosystem services and, at the same time, fulfil some administrative restrictions. These objects, which include ecological focus areas (EFA), support biodiversity, contribute to climate resilience and protect soils and habitats as well as ground and surface water. The tight implementation deadline (the EFA elements had to be mapped preferably until 2015, but not later than 2018) put in perspective the keep data at the source principle (Lemmen et al., 2015), which implies importing and reusing third-party environmental datasets in LPIS, rather than collecting new data.

Whilst on one hand the INSPIRE Directive (European Parliament and the Council of the European Union, 2007) and national spatial data infrastructures (SDI) may provide valuable input for the CAP, new data explicitly collected for populating EFA can be shared within the SDI. This may amplify the impact of freshly collected data fostering agro-environmental research and management, or contribute to the maintenance of the digital topographic databases. Needless to say that data sharing makes savings of public resources.

Undoubtedly the conditions for data discovery and assessing their fitness for purpose have substantially improved with the implementation of INSPIRE. Geoportals established at European, national, and regional levels make both metadata and spatial data services accessible (Kliment et al., 2013) in the domain of the spatial data themes listed in the annexes of the directive (European Commission and European Environmental Agency, 2014). INSPIRE also foresees a gradual data harmonisation. However, for the data themes relevant for greening the interoperability problems may persist until 2020, as the majority of environmental thematic datasets such as soils, natural risk zones, bio-geographic regions, habitats and biotopes and species distribution are included in Annex III of the Directive and so are not immediately addressed.

The experience of INSPIRE and the national SDIs show that the easiest way towards knowledge integration and interoperability leads through a framework that presents the shared “language” between communities (Laschuyt and Hekken van, 2001). This framework, which includes harmonised semantics, data structures, and spatial representation, relies on standards and other well-established practices of user communities. INSPIRE successfully used conceptual models for documenting standard-driven interoperability specifications for a large number of spatial data themes using the common principles of the Generic Conceptual Model.

The new requirements of CAP, the on-going domain related standardisation of geographic information in ISO/TC 211 and OGC and the success of the INSPIRE data specification process in 34 data themes (European Commission DG JRC, 2014; Geospatial World Forum, 2015) urged us to outline a generic domain model for IACS that targets a dual objective: to express the requirements of the CAP in technical terms, and to serve interoperability.

1.1. Background

Standardisation of geographic information in the agricultural domain is on the agenda in a number of organizations. The most relevant initiatives are the on-going work of the Agriculture Domain Working Group of the Open Geospatial Consortium (OGC, 2015), the development of agroXML (Association for Technologies and Structures in Agriculture, 2012), the accepted Land Administration Domain Model (LADM) standard (ISO/TC211, 2012b), the INSPIRE data specification on Agricultural and Aquaculture Facilities (INSPIRE Thematic Working Group Agricultural and Aquaculture Facilities, 2013) and the LPIS Core Model (LCM) (Sagris and Devos, 2010). The ultimate goal of these initiatives is to facilitate a platform independent exchange of geographic data using shared semantics and the Geography Markup Language (GML).

With the exception of agroXML, which focuses at the development of an appropriate XML extension, the above mentioned initiatives widely use the Unified Modelling Language (UML) to document the information concepts and in some cases (INSPIRE) the main use cases. The formalism of a conceptual schema language is necessary to turn such models into machine readable conceptual schemas (ISO/TC211, 2005a).

The LCM was developed to support the LPIS implementation of the Member States (MS), especially to provide an input for model conformance testing (Sagris and Devos, 2010). This conceptual model detailed the properties of reference parcels together with their relationships with other component of IACS, in particular, with Declarations and payments, Farmer register, and Cross compliance (Inan et al., 2010). Consequently the LCM presented a simplified collaboration diagram between selected classes of IACS.

Although IACS/LPIS implementations are subject to MS subsidiarity (Inan et al., 2010), the early LCM was successful in presenting uniform requirements across the sector. The content, structure, and the properties of the model were described by (Inan et al., 2010; Sagris et al., 2013). A detailed technical specification of LCM (LCM v.1.2) was given by (Sagris and Devos, 2010). The main feature types introduced by this model were the reference parcel, land cover type, farming limitation, farmer, aid application, agricultural parcel, farmer sketch and crop code.

The development of LCM was tightly linked with the LPIS quality assurance process (LPIS QA). In addition to conformance testing it served as a documentation basis of the so called eligibility profile, which accommodated specific implementation conditions in terms of land cover (Milenov and Devos, 2012). However the model did not address some other important usability aspects like validity of data or their state in the updating cycle.

It is important to note that LCM v1.2 used some standards. The attributes were specified as types of ISO 19103 Geographic Information – Conceptual Schema Language standard (ISO/TC211, 2005b). As the developments of LCM and the LADM were running in parallel, the LPIS community actively contributed to this latter. Annex H of LADM (ISO/TC211, 2012b) presents LCM as a conforming profile.
The data model for the collaboration between LADM and LCM is discussed in details by (Inan et al., 2010). LCM v1.2 mentioned INSPIRE in the context of reusing external classes (Sagris and Devos, 2010). Also the INSPIRE model/schema testing methodology was recognised as an input to LPIS (Sagris et al., 2013). However, no formal link between the INSPIRE conceptual models and LCM v1.2 was established. Nevertheless the standard and model driven specification methodology, the use of XML, XML Schema Definition (XSD) application schema, Geography Markup Language (GML), as well as the choice of the conceptual schema language (UML) unambiguously showed the impact of INSPIRE.

2. Challenges and methods

The CAP regulations include rules and requirements both for the business processes and the supporting information infrastructure (IACS) generally aiming at efficiency and transparency. These goals are very similar to those of enterprise architecture (EA) frameworks, especially when enterprise is defined as “any collection of organizations that has a common set of goals. For example, […] a government agency, […] or a chain of geographically distant organizations” (The Open Group, 2013). EA frameworks motivated us to take a holistic view on CAP and model it in three interrelated segments of Agile EA modelling (Gill, 2015) addressing the motivational concepts (requirement model), business layer (dynamic model) and the application layer (conceptual model).

There are a number of appropriate EA modelling standards starting from the overarching high level Archimate down to the detailed level BPMN, UML, FAML, SoaML (Agostinho et al., 2015), just to mention a few of them.

The completion of EA framework with the concept of interoperability added another technology enabler to our work, as the common understanding of requirements, business processes and information concepts are vital for collaboration between the numerous actors of CAP.

We selected a comprehensive UML based modelling approach for various reasons. Firstly, the provisions of the CAP regulations are implementation independent in terms of organisational structures and IT solutions. Therefore we did not deal with the technology layer of EA and limited the business layer to behavioural elements presented in our dynamic model. On the other hand we needed a suitable extension for storing the requirements as well as a detailed level tool for the application layer, i.e. for our conceptual model.

UML is widely used in mainstream information technology as well as in geographic information modelling (Belussi et al., 2006), comprising the geographic information standards of ISO TC 211, OGC, INSPIRE and LCM v1.2. Furthermore, the majority of UML tools allow modular development, which makes such products resilient yet very flexible when there is a need to adapt to changes. Since the requirements basis continues to evolve, this latter was a very important aspect of our choice.

Before starting we reviewed the predecessor LCM v1.2 model to see whether it could be extended and also made an inventory of potential third party datasets that may fill information gaps. In the course of this work we immediately observed that the new CAP extended not only the scope, but also increased the implementation flexibility of the Member States. So we concluded that a genuinely new model that embraces all requirements and interoperability aspects of spatial data has to be developed. Nevertheless we were keeping at hand the class models of LCM v1.2 and for sake of continuity we reused its concepts when no semantic clash with the evolving regulatory text or the foundation standards occurred.

Selecting a new model development was justified by other reasons too. Firstly, we wanted to structure, formalise, and integrate all best practices that were widely spread within the LPIS community. The related materials were dispersed in documents of various style (web pages, downloadable documents, document templates, figures, question and answer sessions, presentations, citations to standards, XML files, etc.). These materials are usually available in the Wikicap (European Commission DG JRC, 2015b).

Secondly, we realised that with a formal quality model and integrating feature level metadata (Nogueras-Iso et al., 2004; Batcheller, 2008; Batcheller and Reitsma, 2010) not only conformance testing but also information on data usability (Devillers et al., 2010) can be improved.

Lastly, we wanted to fulfil the increased users’ expectations in terms of system interoperability (Paškauskis et al., 2013; Martin et al., 2014; Řezník et al., 2015). Therefore the IACS domain model should provide a standard-based framework, which takes into account the law in force, allows flexible extensions, and supports efficient information exchange between various domains, first of all with environmental science and management.

So we propose a domain model for spatial data within IACS, which includes structured requirements in the requirement model, system use cases in the dynamic model, and information concepts in the conceptual model. In order to see what spatial data are needed and in which subsystem, we performed a preliminary analysis of the main regulatory provisions. In addition to LPIS, controls and the geospatial application, the new CAP regulation makes other implicit references to spatial information management methods, for example, in the collective implementation of EFA, or in accounting areas with natural constraints.

Depending on the applied subsets of standards and modelling constructs the resulting model may be very different even when the same conceptual schema language is used. This difference is not limited to the layout, but also impacts on semantics. Therefore, before starting working on the domain, we agreed on the applicable UML grammar and notation. In order to properly deal with the location component, we used the GML profile of UML.

We also restricted the applicable UML diagram types and modelling constructs. Our final modelling toolset together with the applicable documentation layout was documented in the reference model (Kučas and Tóth, 2015b). A generalised overview of concepts of this reference model is given in Fig. 1.

In the reference model we defined the following diagram types:

- Package diagram for representing the structure of the model;
- Class diagrams to document the information concepts;
- Use case and activity diagrams to describe the high and low level business processes;
- State machine diagram for describing life cycle status of information concepts and data.

The conceptual model describes the logical structure of the system. The building blocks are classes. The properties of a class were described by attributes and/or relationships. For describing a class we also used methods (operations or behaviour) and constraints. The latter were formulated both in plain text and Object Constraint Language (OCL). We also applied feature type, data type and code list stereotypes to our classes. The stereotypes extended the semantics, but not the structure of the classes.

The particularity of our conceptual model is the bi-directional navigability of the associations, which is justified by the backward traceability that is required for every transaction in IACS. Consequently the connectors include named roles with multiplicity, and possibly, constraints at each ends. We used the following association types:

- Requirement model: aggregation and composition;
- Dynamic model: control flow, use, associate, invoke, extend, realise, trace;
- Conceptual model: association, aggregation, composition, nesting.

The reference model also defined rules for code list extensions. The 'none' value emphasize the strict requirements where users are not allowed to extend the referenced vocabulary. When the value is 'narrower', the vocabulary may be extended by semantically narrower terms. If the value is 'any', extensions with any additional term is accepted. These governance rules are fully in line with the corresponding implementing rule of INSPIRE (European Commission, 2011). We did not define own enumerations, but occasionally imported them with external schemas.

Due to the meticulous preparatory work, external peer review process, and improvement iterations the reference model was frozen and we did not apply any "on the fly" modification in course of the domain related work.

We started the work on the domain model with setting up a generic overview of CAP, distinguishing the requirements, the system use cases and the information system. These areas were further structured according to the IACS subsystems laid down in the regulations. At this point we arrived to a framework that allowed a modular model population. Having defined where spatial data reside, how they interact, and which of them are "urgent" from a requirements point of view, we prioritised our tasks. Firstly we targeted the detailed modelling of the LIPS subsystem. The second priority was to deal with those feature types of the Integrated Control and the Aid Applications and Payments subsystems that carry spatial data. The Beneficiary, Entitlement, and Animal registers, as well as the Farmer Advisory System remained placeholders.

The development process followed the INSPIRE iterative methodology (INSPIRE Drafting Team “Data Specifications”, 2008). Each cycle consisted of requirement analysis, use-case development, drafting the first cut application schema, as-is and gap analysis, and schema refinement followed by review and testing. The outcome of review and testing provided input for the next iterative cycle. We re-performed only those steps that were justified by a specific input. The overview of the development process with the inputs and outputs is presented in Fig. 2.

In course of modelling we were driven by two generic and inter-related principles: the holistic approach and reusable components. The holistic approach meant that none of the components was developed in isolation. Exploring the co-dependencies was a permanent task that yielded realisation and tracing relationships. We set up and maintained a glossary from the very beginning of work.
that helped us to eliminate semantic clashes and duplications. Due to the specific setup of the IACS we introduced the definitions of all feature types, technical and legal terms, as well as a vocabulary for abbreviations.

In order to model the functional requirements of the domain we had to make some preparatory work. First, we worked on the legislative text to transform it into a structure suitable for importing in our software tool. We divided the legal acts in approximately 9000 logical statements. After the import we discarded those that were not related to the domain (e.g. references to EU treaties) and classified the rest to build up a hierarchy. For completeness we analysed...
the implementation guidelines and the Wikicap pages (European Commission DG JRC, 2015b).

The requirements were classified as functional and non-functional ones. Out of the non-functional requirements we treated extensibility by standard modelling constructs. The rest of them (performance, reliability, security) are implementation-dependent, so we did not address them. The functional requirements became subjects of direct realisation in the model as use cases or information concepts.

In the second phase of preparations, in order to reuse their elements, we imported 38 application schemas, namely ISO/TC211 geographic information standards, the INSPIRE Generic Conceptual Model, and Annex I-II-III application schemas. They were mainly used for defining property types in class diagrams.

One of the most critical steps of our work was to correctly identify our use cases as they are “the basic concept for specifying requirements of […] information systems.” (Langlands and Edwards, 2009). A business use case describes in a technology-free terminology the business processes and the actors to achieve their goals. It was obvious for us that the CAP regulations and the related implementing guidelines of the European Commission represented our principal business cases. It is important to emphasize that their provisions were taken as axioms. They were not discussed or changed, only analysed and structured to realise them in system use cases and information concepts.

Our domain model supports two high level system use cases. Firstly it may guide the information model development of local implementations. Secondly it provides a support for conformance testing of the LPIS component of IACS both in matters of workflows and the related information concepts. So the model, on one hand, can be used by system developers who work on local IACS implementations. On the other hand it can be used by domain experts, who are in charge for the maintenance and the quality control of the system. The latter group mostly benefit from the step by step guidance documents derived from the detailed documentation of the related activity diagrams.

Following the dependency path from the initial requirements to the use cases and then to the information concepts we could check the completeness of the model. In the course of the development we were continuously performing the “as-is” and gap analysis, confronting the freshly developed application schema with existing technical solutions, such as ISO TC/211 standards, INSPIRE application schemas, and LCM v1.2. In the case of information gaps we searched for solutions in other domains further strengthening information infrastructures and interoperability. For example, we proposed to import the code list of crop types from the community farm structure survey that is well known in agricultural statistics (European Commission DG ESTAT, 2015).

We also organised testing with stakeholders’ involvement to evaluate how faithfully the requirements were reflected and how technical standards and best practices of the domain were applied. We set up a collaborative tool to keep the discussions and the decisions transparent. After each consultative phase we had a new release of the model. Currently we are working on the third release of LCM.

3. Results

The overview of the CAP legislative documents and the supporting guidelines as well as the revision of the LCM resulted in the IACS domain model, downloadable in native format and accessible in html (Kučas and Tóth, 2015b,c). The entry point to the model is the high level overview (Fig. 3) diagram, which systemises the ideas about the reformed CAP in terms of requirements, users’ interactions, and information concepts.

In this interwoven structure every functional requirement points to at least one element, either in the dynamic or the conceptual model. Through the realisation and traceability functions it is possible to arrive from the requirements to the machine-readable conceptual schemas (XSD), or vice versa.

The overview diagram provides full navigability between the components. Thanks to the built-in hierarchies a vertical navigation is possible; i.e. the user can go deeper in any part of interest by a simple click. The horizontal navigability between the components is implemented through the realisation and traceability mechanism. The html export (Kučas and Tóth, 2015c) of the model makes accessible these functionalities for those users that do not have the modelling software.

The use cases and the life-cycle status of data were included in the Dynamic model, which also contains the actors (Fig. 4). We defined ten high level use cases, out of which we elaborated three in details (in Fig. 4 they are of light yellow colour).

The high level use cases were broken in lower level use cases and the latter were decomposed in use cases of lowest level, i.e. units of work that are meaningful from business point of view. A lower level use case is presented in Fig. 5, which includes eight lowest level use cases. It is worth noting that semantic relationships between the use cases are represented by the “invoke” stereotyped relationships, which means that performing some of the use cases automatically triggers some embedded workflows.

The lowest level use cases were further detailed by activity diagrams. Two examples of activity diagrams are shown in Figs. 6 and 7. The invoking mechanism represented through the stereotyped connector in Fig. 5, which is also visible in the related step of the activity diagram in Fig. 6. This means that an instance of the reference parcel feature type can be updated only when the decomposable activity step (Apply 2% stability threshold shown in Fig. 7) has been performed.

Using our software tool we performed simulations for the complete dynamic model of the LPIS subsystem. All the steps included in the workflows ran correctly invoking all interrelated processes and pointing to the necessary elements in the conceptual model (see Fig. 7). The decision points highlighted the variants and their impact on the information flows. The logical sequences of low level use cases and the fully documented steps therein clarified the related pre- and post-conditions. The use cases and the activity diagrams provided input for detailed technical guidelines on system maintenance and quality assurance (European Commission DG JRC, 2015a).

From the information point of view the main outcomes of our model are the conceptual models (frequently referred to as application schemas). According to our priorities we prepared the detailed conceptual model for the LPIS subsystem as shown in Fig. 8.

In order to respond to the greening requirements of CAP we had to introduce new feature types. The most important is the EFA that has to identify, localise, and quantify the area of those objects that deliver ecosystem services. We also supported the collective implementation of EFA (i.e. more farmers can maintain the same EFA when specific conditions are met) by introducing a number of auxiliary feature types (EFA regions, mandated EFA areas).

The extension of the LPIS model was a good occasion to review the conceptual model for the quality assessment framework (QAF). We defined quality measures for EFA elements and integrated all concepts of quality management from sampling through to conformance testing. The QAF is presented as a leaf in the LPIS model.

We supported interoperability by reusing components from INSPIRE and ISO/TC 211 standards, in particular the Conceptual schema language (ISO/TC211, 2005b), Conformance testing (ISO/TC211, 2000), Spatial schema (ISO/TC211, 2003), Temporal schema (ISO/TC211, 2002) and Data quality (ISO/TC211, 2013).
Fig. 3. Overview of the IACS domain model.
Some more specific attributes and value types of the quality model were taken from a statistical standard (ISO/TC69/SC5, 1985), while life cycle information, validity, document citation, and legislative citation came from INSPIRE. We restricted user-defined value types to those code lists that contain values prompted by the CAP regulations. The extensive use of standards was one of the instruments that reinforced internal consistency of the conceptual model.

The second way to achieve internal consistency was sharing components within the model. Following the general principle “specify once, use several times” we reused classes (for example

Fig. 4. High level use cases of the IACS domain model.
Fig. 5. Example of a lower level use cases of the LPIS subsystem broken down in lowest level use cases.

code lists) not only within the same application schema, but also in application schemas of different subsystems. These elements received a special visibility in a separate, the so-called “Base types” schema (Kučas and Tóth, 2015a). This solution is similar to INSPIRE, where the shared elements reside in the Generic Conceptual Model. The shared elements helped us to avoid redundancy and gave an unambiguous indication where and how the system integration in IACS happens.

In our model we had to respond to the implementation flexibility that Member States enjoy by the force of the European law. This is not limited to the hardware and software components, but is extended to adapting the information content according to the local natural conditions and agricultural practices. We implemented the extensibility using various methods. Firstly, in accordance with the reference model we used code lists with different extensibility governance as defined in the method part.

Secondly, we allowed multiple representation geometries using the generic GM.Object type from ISO 19107 (ISO/TC211, 2003). In case of the EFA feature type this solution represents an implementation choice as foreseen by the law. In case of reference parcels the geometry is use case specific: for sampling in LPISQA a simplified (GM.Point) representation is satisfactory, while in other cases (e.g. application process) full planar localisation (GM_Surface) is required. We differentiated these use cases by specifying constraints on the geometry attribute.

Thirdly, we specified constraints on attribute values to indicate how and where users may insert locally valid limitations. Such limitations are applicable to certain EFA types (for example hedges,
or field margins), where the width cannot exceed certain locally specified values.

Fourthly, we acknowledged that some properties may not be conceptually defined in a specific system. For example, we encouraged the use of global unique identifiers, but assigned a 0.1 multiplicity to show that this practice is not required by the law.

As stated before, our conceptual models are also presented as machine readable XSD application schemas with XML code lists.

Fig. 6. Activity diagram of the “Update reference parcel” use case.
The schema storage and management approach are inherited from INSPIRE, which supports development of registries as suggested by the (European Commission, 2013c). This platform independent output is handy at automatic model validation, or at development of logical schemas for physical models.

4. Discussion

“In the context of a lot of challenges and the greening of the CAP [...] there is also a need for [...] new or innovative solutions combined with appropriate technologies, such as Information and Communication Technologies (ICT), satellite navigation support systems (Geographic Information Systems and remote sensing) for improved management of farm input (e.g. precision application and irrigation) or new management tools” (Singh et al., 2014).

Inspired by the above citation we position the proposed IACS domain model as a tool that facilitates the technical understanding of CAP and helps to insert IACS in ICT and SDI. The biggest methodological challenge of model development was to address the requirements of the CAP and interoperability at the same time. Our framework also underpins the requirement of sharing core semantics and fundamental data structures of a specific field (Association for Technologies and Structures in Agriculture, 2012; Tóth et al., 2012; Ferrè et al., 2014).

There are only a few domain models that extensively use spatial information. The two examples that are generally accepted by the user communities are the LADM (ISO/TC211, 2012b) and the INSPIRE consolidated model repository (European Commission, 2015). The first is interested in rights, responsibilities and restrictions affecting land and the geometrical (geospatial) components thereof. The second deals with information needed to EU environmental policies and environmental impact assessment. Strong associated to INSPIRE there are other initiatives such as the European Location Framework (Eurogeographics, 2015) and the
European Union Location Framework (EULF) (Pignatelli et al., 2014).

The above initiatives focus on their respective domain from information point of view. They provide detailed specifications for the conceptual models, while the requirement and the business process parts are outlined with short narratives or occasional use case diagrams. Even when use cases are documented in diagrams, they are not connected to the conceptual model, missing the opportunity to create a self-explanatory model as envisaged by (Agostinho et al., 2015).

Interoperability can be enabled in the design phase by the consistent use of an agreed set of standards, like it happened in case of INSPIRE. We followed this good practice and directly applied the UML conceptual models of geographic information standards.

The interwoven requirement, use-case, and conceptual models of the IACS domain model present a number of advantages. Firstly, this structure gives a full overview of the domain from business and information points of view. Secondly, the model is an instrument for embedding IACS in spatial data infrastructures by sharing fundamental standards, semantics, and information modelling concepts. Thirdly, the compact UML presentation and the XSD application schemas are handy starting points for system developers. Finally, the overview model with the linked plain text documentation of the elements is an inclusive communication medium for non-IT/GIS users.

The tracking mechanisms built in the model (Fig. 9) is a novel solution not only in policy modelling, but also in geographic information science. The explicit representations of dependencies give justifications for non-trivial system developments, especially in matters of subsystem integration in IACS. They are also useful in assessing the impact of eventual changes on the system as a whole. Since simplification of the reformed CAP is on the agenda (Council of the European Union, 2015), (Special Committee on Agriculture, 2015), this feature may become important in policy planning and monitoring.

The explicit tracing between the elements of the domain model may help to identify errors and weaknesses of system implementations. The discovery of missing links that cause interruptions in the information flows represent inputs not only for system design, but also for system audits.

Promoting feature level metadata increases the usability of data within IACS. The life cycle information (when an instance of a feature type was created/archived) or dates of their legal validity are important parameters for selecting right data populations for controls or for data quality assessments. Based on the internal consistency rules the combination of life cycle and validity properties may automatically trigger some business process (e.g. retroactive recovery of ineligible payments). This demonstrates that efforts paid on comprehensive and careful conceptual modelling can yield simple but effective technical solutions that improve transparency of CAP implementation.

According to Ferré et al. (2014) domain models “may also go beyond physical objects and processes to involve the cognitive and social reality, e.g. by including regulations concerning restriction areas”. In our model business processes and the related information
Fig. 9. Generic tracing mechanisms between the requirement, dynamic, and conceptual models of the IACS domain model.
areas where a more detailed description of agricultural land can add further clarity to better understand the processes.

In order to successfully implement greening measures, “one of major research needs to be performed to define what an ecological focus area is and how this can be determined in the field” (Singh et al., 2014). Our UML model contains not only the requirements and the definitions but also the technical translation of EFA implementation choices, represents a basis for cataloguing EFA types across Europe. This can be also used for a very detailed land cover mapping of the rural landscape, where for the sake of interoperability the application of ISO 19144-2 standard (ISO/TC211, 2012a) is highly recommended.

Since the development of the IACS domain model is not finished some shortfalls remain that have to be addressed in the future. The volume and the urgency of work constrained us to set clear priorities and strict sequences of work. Consequently, some parts of the model represent mere approximations of the information content. However, the modular structure permits to continue the work upon any ‘placeholder’ element. The ultimate goal is to address all functional requirements.

Another issue of discussion might be the selection of UML driven methodology and the applicable UML profile. UML is the de-facto standard formalism not only for the analysis and design of software but also for documenting standards. This technology is applied in TC 211 of ISO, or in INSPIRE. GML is adopted specification of the OGC and is rapidly emerging as a world standard for the encoding, transport and storage of all forms of geographic information (Lake, 2005). However, the expressiveness of the UML can go undetected by the user in complex diagrams, and cause various forms of inconsistencies or redundancies (Berardi et al., 2005). In case the risk may occur that responsible bodies, especially domain specialists may misunderstand the model.

In order to minimise this risk we decided to follow a strategy, where UML and text documentation go hand in hand. We attached extensive documentation (definition, description, examples, and notes) to each element that could be directly exported together with the diagrams to various communication media: to the html view and text reports. The latter can be inserted in a feature catalogue, traditionally used by domain experts, or can be published on the widely used Wikicap (European Commission DG JRC, 2015b).

The third aspect where we need to carry out further research relates to the rules of modularity and scalability at conceptual profile development. We recognise that standard workflows are needed for quick and effective consideration of local conditions, which would be a useful input for construction of logical models. This task should also include schema mapping (completeness, integration), since data management, data exchange and data integration are well-known contexts in which schema mapping plays a central role (Rull et al., 2008; Papotti and Torlone, 2009; Calvanese et al., 2013). Currently we are working on model conformance testing of the LPIS subsystem that together with the selected encoding (XML and GML) will reinforce data management and exchange between user communities.

5. Conclusions

The proposed IACS domain model offers a holistic view of the Common Agricultural Policy from business and information points of view describing the relevant static and dynamic aspects of the domain. The model addresses two high level use-cases. The first deals with operating the Integrated Administration and Control System according to the rules defined in the related EU regulations. The second is about enabling interoperability with mainstream ICT and GIS. This double goal can be achieved through a careful realisation of requirements by standard-driven information concepts. This model, on one hand, may serve as a reference to check completeness and modalities of implementation of local IACS implementation. On the other hand, the standard-based spatial data concepts may contribute to application development in agro-environmental domain and inserting them to spatial data infrastructures.

Conformity with standards and the wide spread implementation of IACS (mandated in the 28 MS of the EU and also used by candidate and accession countries) qualify the LPIS subsystem to be used as reference data. Linking information to LPIS by object referencing further amplifies interoperability and convergence between domains.

The modular structure of the model and the dedicated place-holders allow subsequent development and extensions where the reference model, the imported schemas and the interlinked shared semantics facilitate the consistency of further work. Presenting the domain in an interlinked unique structure helps to indicate critical points, potential errors and the impact of changes in the system as a whole providing input for local system design, audits, and policy analysis. The storing of requirements and implementation options in a common also, underpins interoperability with other domains and enhances the transparency of the CAP.

6. Disclaimer

The model features, described in this article and accessible through the given links, express and share a methodological approach. They do not in any manner replace the formal texts of the CAP legislation.

Acknowledgements

This work was carried out within the institutional work programme of the Joint Research Centre of the European Commission. Colleagues in the GTCAP project helped significantly through discussing the business details.

We are grateful to Valentina Sagris, Clemens Portele and Rokas Bartkevičius, who performed the external peer review of the UML reference model. Paolo Pizzoli coded the legal text for importing and Katalín Bódis provided helpful comments on an earlier version of this manuscript. The detailed reviews of Pavel Milenov, Neil Hubbard and the two anonymous reviewers were most valuable contributions resulting in improvements. Last but not least we would like to thank Wim Devos for encouraging us to write this article.

References

Agostinho, C., Ducq, Y., Zacharewicz, G., Sarraipa, J., Lampathaki, F., Poler, R., Jardim-Goncalves, R., 2015. Towards a sustainable interoperability in networked enterprise information systems: trends of knowledge and model-driven technology. Comput. Ind. Available at: http://www.sciencedirect.com/science/article/pii/S0166361515300191 (accessed 21.08.15.).

Association for Technologies and Structures in Agriculture, 2012. agroXML. Available at: http://www.agroxml.de/ (accessed 08.09.15).

Batcheller, J.K., 2008. Automating geospatial metadata generation–An integrated data management and documentation approach. Comput. Geosci. 34 (4), 387–398. doi:10.1016/j.cageo.2007.04.001.

Batcheller, J.K., Reitsma, F., 2010. Implementing feature level semantics for spatial data discovery: Supporting the reuse of legacy data using open source components. Comput. Electron. Urban Syst. 34 (4), 333–344. doi:10.1016/j.cru.2010.02.008.

Belussi, A., Negri, M., Pelagatti, G., 2006. An ISO TC 211 conformant approach to model spatial integrity In: F. Roddick, R.V. Benjamins, S. Si-Said, (Eds.), Chief Advances in Conceptual Modeling—Theory and Practice Springer-Verlag Tucson, AZ, USA 2006; 100–109 Available at: http://download.springer.com/static/pdf/160/bok/978-3-540-47704-4pdforigin.tar. http://link.springer.com/book/10.1007/11908833&fileId=exp=1440769239_ae=static/pdf/160/bok/978-3-540-47704-4.pdforigin.tar?2.
