

Open Data Model for (Precision) Agriculture Applications and Agricultural Pollution Monitoring

Tomáš ŘEZNÍK
Department of Geography
Masaryk University
Brno, Czech Republic

Karel CHARVÁT
Baltic Open Solutions Center
Rīga, Latvia

Vojtěch LUKAS
Department of Agrosystems and Bioclimatology
Mendel University
Brno, Czech Republic

Karel CHARVÁT Jr.
WIRELESSINFO
Litovel, Czech Republic

Šárka HORÁKOVÁ
WIRELESSINFO
Litovel, Czech Republic

Michal KEPKA
Department of Mathematics
University of West Bohemia
Plzeň, Czech Republic

Abstract—Both agricultural and environmental domains have to manage many different and heterogeneous sources of information that need to be combined in order to make environmentally and economically sound decisions. Such examples may be found at the definition of subsidies, national strategies for rural development, development of sustainable agriculture etc. This paper describes in detail the development of an open data model for (precision) agriculture applications and agricultural pollution monitoring when aiming at identification of requirements from users from agricultural and environmental domains. The presented open data model for (precision) agriculture applications and agricultural pollution monitoring has been registered under the GEOSS (Global Earth Observation System of Systems) Architecture Implementation Pilot – Phase 8 in order to support the wide variety of demands that are primary aimed at agriculture and water pollution monitoring.

Keywords—open data; environmental monitoring; GEOSS; farm management information systems; cloud computing; INSPIRE

I. INTRODUCTION

The agriculture sector is a unique sector due to its strategic importance around the world. It is crucial for both citizens (consumers) and economy (regional and global) which, ideally, should make the whole sector a network of interacting organizations. Rural areas are of particular importance with respect to the agri-food and environmental sectors and should be specifically addressed within this scope. The different groups of stakeholders involved in the agricultural and environmental activities have to manage many different and heterogeneous sources of information that need to be combined in order to make economically and environmentally sound

decisions, which include (among others) the definition of policies (subsidies, standardization and regulation, national strategies for rural development, climate change), development of sustainable agriculture, ensure crop and animal food production, pests and diseases detection, etc.

In this context, future agriculture knowledge management systems have to support not only direct profitability of agriculture or environment protection, but also activities of individuals and groups allowing efficient collaboration among groups in agri-food industry, consumers, public administrations and wider stakeholders communities, especially in rural domain.

As stated by FAO [14], agriculture is the single largest user of freshwater resources, using a global average of 70% of all surface water supplies. Except for water lost through evapotranspiration, agricultural water is recycled back to surface water and/or groundwater. However, agriculture is both cause and victim of water pollution. It is a cause through its discharge of pollutants and sediment to surface and/or groundwater, through net loss of soil by poor agricultural practices, and through salinization and waterlogging of irrigated land. It is a victim through use of wastewater and polluted surface and groundwater which contaminate crops and transmit disease to consumers and farm workers. Agriculture and water pollution is therefore a subject of several national and/or continental legal acts, such as European Nitrate Directive [6], European Water Directive [8] or Common Agriculture Policy [7] applicable in all European Union Member States.

Regular monitoring through sensor networks is needed in order to support the evidence of agriculture and water

pollution. Such evidence stems from the legal frameworks as well as from farmer's and citizen's needs. Some examples thereof may be found in weather (soil and air temperature, soil and air moisture, wind speed), crop growth monitoring for precise application of fertilizers and pesticides, ground water level and quality observation, monitoring of nitrate leaching into water, etc. It is beneficial to use the sensors networks according to the OGC implementation specification adopted also as ISO 19156:2011 [15]. The O&M approach is the key for near real time obtaining of relevant data, its filtering according to the user's needs as well as publishing in the form standard OGC Web services and/or technologies like Google Earth API (Application Programming Interface).

The European projects called "Farm-Oriented Open Data in Europe" (FOODIE), funded between years 2014 and 2017, and "Farming Tools for external nutrient Inputs and water Management" (FATIMA), funded between years 2015 and 2018 address the above mentioned agronomical and environmental issues [25]. After presenting the methodology in this paper, the open data model for (precision) agriculture applications and agricultural pollution monitoring is presented in detail in textual as well as formalized way using the class diagram in the Unified Modelling Language. In the conclusion, benefits, opportunities and future development are mentioned.

II. METHODOLOGY

The primary inspiration for the methodology development lies in the existing international and European initiatives that aim at facilitating the exchange and access to a wealth of heterogeneous data sets related to the environmental and agricultural domains.

References to the main European policies were also included to define the context of the agriculture sector, such as Common Agricultural Policy [7] or Water Framework Directive [8]; and these have to be taken into account in the decision making process of the stakeholders. In this sense, call for global data collection for agricultural monitoring is analyzed by [26]. Principles of common agricultural policy are provided by [4]. Influence of Water Framework Directive on agriculture is discussed by [1]. Processing of biomass data from remote sensing is described in [22]. See [23] for information on Infrastructure for spatial data in Europe (INSPIRE) including the application schemas for agriculture and aquaculture. European nitrate directive and its influence on the farm performance were described by [21]. Discussion on accessing structural functionality and landscape service is provided in [27].

The results from relevant projects provide an overview of the different architectural approaches followed by various projects in the environmental and agricultural domain which represent the basis for designing FOODIE architecture and specifying its building blocks. For instance, an architecture for environmental and agronomical data that may be re-used for FOODIE purposes is advertised by [5], [2], [13] and [20]. Cataloging of the collected data follows the principles defined for environmental spatial data by [24]. Spatial data harmonization of openly available databases is further analyzed

by [3], while the mechanisms for its visualization is provided by [28].

The analysis of user requirements and the derivation of requirements on various software/hardware components cannot take place without having in mind a common FOODIE system architecture. Here, it rely upon agreed international standards such as ISO Reference Model for Open Distributed Processing (RM-ODP [18]). The RM-ODP defines, among other, the Information Viewpoint as the semantics of information and information processing and contains the information resources identified as a use case extension. The modelling basis has been extended by the standardization documents originating from the European INSPIRE (2007/2/EC) Directive. The underlying principles are described in [12].

To sum up, the FOODIE and FATIMA projects have a lot of similarities with the above mentioned initiatives. Data model and searching including metadata originates from INSPIRE. Parts of global initiatives called GEOSS (Global Earth Observation System of Systems) and COPERNICUS could be used and integrated as a part of FOODIE hub. It is important for the FOODIE and FATIMA implementation to establish a link with GODAN (Global Open Data for Agriculture and Nutrition) initiative, which is trying to define world Wide standards for Agriculture Open Data and CGIAR (Consultative Group on International Agricultural Research), which is active on a global scale in similar area as FOODIE and FATIMA.

The methodological point of view addresses the issues of user requirements, development of the proposal for the open data model, its verification through processes definition including business as well as environmental perspectives.

The FOODIE user requirements have been collected for three pilots, where each pilot contains three use cases (scenarios). The pilots foreseen in the funded project phase are:

- **Pilot 1: Precision Viticulture (Spain)** focuses on appropriate management of the inherent variability of crops, an increase in economic benefits and a reduction of environmental impact.
- **Pilot 2: Open Data for Strategic and Tactical Planning (Czech Republic)** aims at improving future management of agricultural companies (farms) by introducing new tools and management methods, which follows the cost optimization path and reduction of environmental burden, improving the energy balance while maintaining the production level.
- **Pilot 3: Technology allows integration of logistics via service providers and farm management including traceability (Germany)** focuses on integrating the German machinery cooperatives systems with existing farm management and logistic systems as well as developing and enlarging existing cooperation and business models with the different chain partners to create win-win situations for all of them with the help of IT solutions.

The user requirements matrix consisting of two hundreds of functional, information and non-functional requirements has been established for nine use cases. After having described the use cases in a semi-formal way using the template developed internally within the FOODIE project, formalization was conducted through the UML (Unified Modelling Language) using activity diagrams for processes and class diagrams for data modelling (e.g. see Fig. 1).

The presented methodology aims at repeatable development of the open data model. Such development is based on the iterative approach for incrementally growing concept and corresponding implementation based on user requirements. It also has to support the evidence on environmental burden. The open data model should support the evidence of all treatments that were used in a certain place as well as (where appropriate) to store relevant information on the application of those treatments. The stored data should together answer the questions like “What amount of which treatment was used in a certain place?”, “When it will be safe to apply another treatment?” or “Is the treatment registered and allowed in the European Union/Member State?”

The methodology also includes a definition of the scope, planning, actors and collaboration when focusing on how to perform the monitoring and management change, how risks will be managed and how the pilot execution will be evaluated attending to Key Performance Indicators described at FOODIE proposal stage. In addition, indicators to estimate the costs and benefits for the end users are also included.

The proposed methodology is applied to the three stages, *proof-of-concept*, *test* and *production*, which address the FOODIE pilots realization in an incremental way. This paper describes the results obtained from the *proof-of-concept* phase.

III. OPEN DATA MODEL

The INSPIRE data model for Agricultural and Aquaculture Facilities ([11], hereinafter AF) is composed by core information in relation to the geographical description of entities defined as “farming equipment and production facilities (including irrigation systems, greenhouses and stables)”. The AF data model is based on the Activity Complex model [10]. Within INSPIRE, “Activity Complex” denotes a generic name agreed across thematic domains trying to avoid specific thematic connotations such as “Plant”, “Installation”, “Facility”, “Establishment” or “Holding”. Such scope may be identified for this paper as the Nitrate Directive [6] or Water Framework Directive [8].

The basic feature types in the AF data model were preserved also for the open data model for (precision) agriculture applications and agricultural pollution monitoring. These are:

- *Activity Complex*: a whole area and all infrastructures it includes, under the control of an operator. In the AF theme, Activity Complex has a specialized representation named Holding.
- *Holding*: a whole area and all infrastructures it includes, under the control of an operator to

perform agricultural or aquaculture activities. It may be composed of one or more — Sites.

- *Site*: belonging to a holding, it is the geographical representation of land that constitutes a management unit. It includes all infrastructure, equipment and materials.

To sum up, the AF data model lacks more detailed levels of information. For instance, missing levels of information are an area for a certain crop or management zones for applications of fertilizers, pesticides, water etc. As a result, the AF data model cannot be re-used for the environmental monitoring since it does not contain the relevant information.

In order to ensure the maximum degree of data interoperability, the open data model for (precision) agriculture applications and agricultural pollution monitoring (hereinafter open data model) follows the INSPIRE generic data models, in particular the aforementioned AF data model, by extending and specializing them (see Fig. 1).

Thus, when taking the Figure 1 into consideration, the AF data model may be further specialized. For these purposes, there should be a feature on a more detailed level than “Site”. The main motivation is to represent a continuous area of agricultural land with one type of crop species, cultivated by one user in one farming mode (conventional vs. organic farming) or having the same environmental features (e.g. organic matter in soil, the amount of fertilizers and/or pesticides inserted in a certain time). Such feature is called “Plot” in the open data model, being its elementary reference item.

Note that the term “Plot” originates from the concept presented in the AF data specification. The precision agriculture domain uses the term “Management zone” for the concept that is in INSPIRE and open data model designated as “Plot”. Please bear in mind that both terms are meant as synonymous when reading the paper and the model.

Each “Plot” has a unique identifier to distinguish a “Plot” from any other “Plot”. Please, note that a “Plot” does not imply any explicit relation to the cadastral. For instance, a “Plot” may be only a part of a cadastral parcel. In other words, a cadastral parcel may contain from zero to many (0..N) “Plots”.

The parent entities named “Site” and “Holding” remain the same as defined in the underlying INSPIRE data model.

In order to illustrate the application of open data model, consider the following example. A company named “FoodieAgroProfi, Inc.” cultivates 4,000 hectares in the southern part of the country. This whole area is, according to the open data model specification, called a “Holding”. Such “Holding” is composed of three types of “Sites”: arable land, grassland and vineyards. Vineyards are composed of 164 explicitly geometrically-defined “Sites” where the wine grapes grow. Arable lands are composed of 120 geometrically-defined “Sites” where wheat, spring-barley and oil-seed rape are produced. These 120 geometrically-defined “Sites” are composed of 150 “Plots” according to the crop species. It means that a “Site” may contain several crop species.

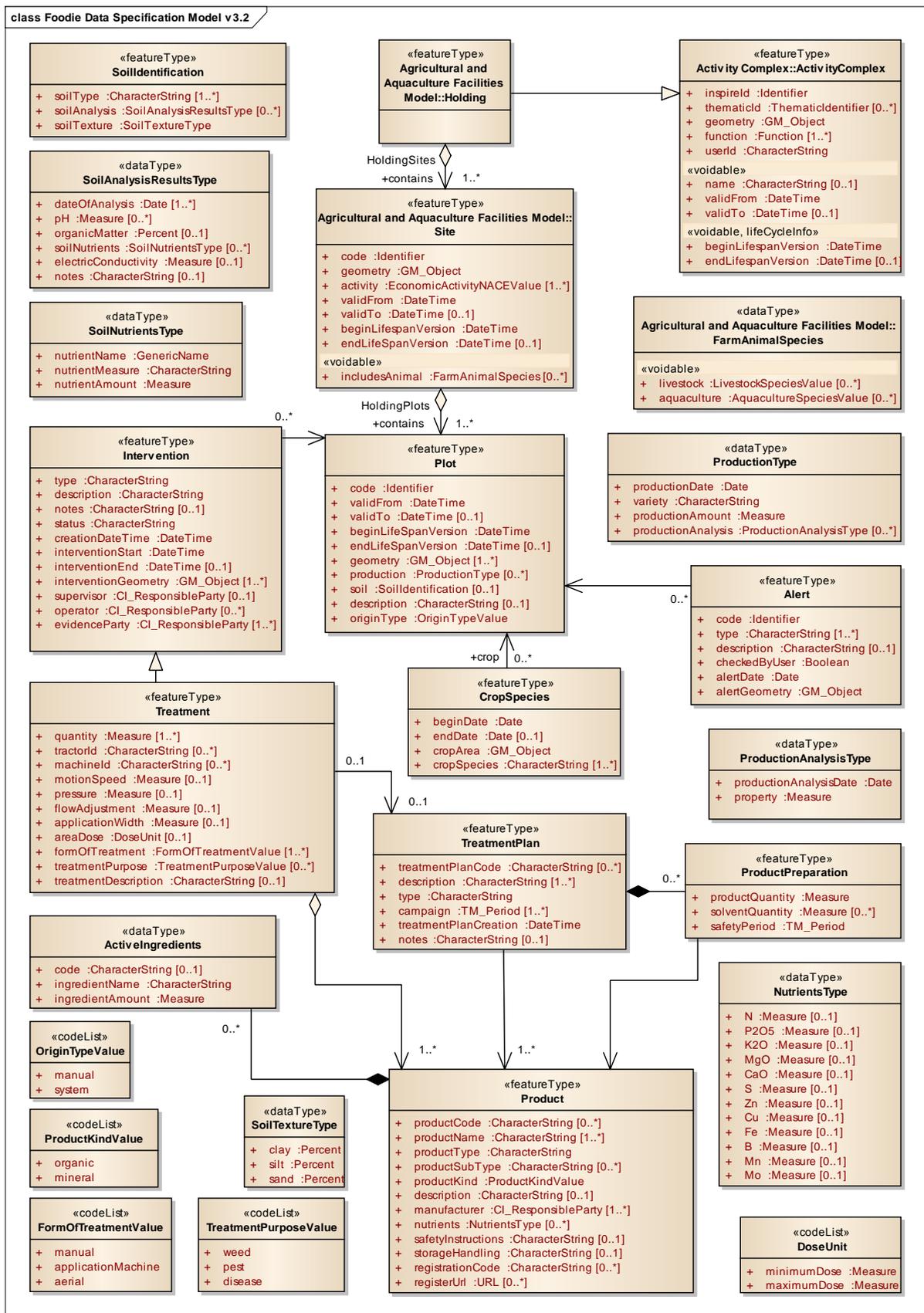


Fig. 1. The FOODIE core application schema based on the INSPIRE, OGC and ISO standardization frameworks.

A “Plot” is therefore the elementary spatial unit within the open data model. Note that the geometrical definition of a “Plot” may vary in time, typically from year to year. A “Plot” does not automatically mean relationship to the crop species. A “Plot” may be established from the intervention, soil type etc. points of view. In other words, a “Zone” would be more appropriate designation. See the explanation above for keeping the term “Plot”.

Regarding the initial import of data, it is possible to re-use the data already contained, at least, in an LPIS (Land Parcel Identification System) of any Member State of the European Union. In that sense the INSPIRE data model and the LPIS concepts are complementary approaches. Thus, for instance, LPIS data may be imported to the open data model on the level of “Site” feature which is equal to LPIS farmer’s block level. A farmer may then add the crop species information to the imported LPIS data to obtain data ready for the “Plot” level.

Open data model enables extensions through associations and/or attributes that may further specialize the “Plot” feature. The aim of extensions is to provide modularity and enable any farmer/external service provider to extend the data model according to his/her needs.

Detailed description of the open data model is presented towards the end of this section. It should be emphasized that the “ActivityComplex”, “Holding” and “Site” have been re-used from the AF data model, while the rest of the model is a new addition developed within the FOODIE and FATIMA projects. It is recommended to compare the following text with the Fig. 1.

The “Plot” is the key feature type in the open data model since it:

(1) is the level to which majority of agronomical and environmental data are related;

(2) acts as a mediator between diverse developed application schemas.

So far, three application schemas have been developed in order to support the complex modelling of the (precision) agriculture and environmental applications:

- FOODIE core application schema (presented in this paper);
- Sensor application schema (beyond the scope of this paper) covering the observations and measurements for sensors as well as human volunteers having e.g. a smartphone;
- Transport application schema (beyond the scope of this paper) covering the issues of fleet management of tractors and application machines.

The UML class diagram presented in Fig. 1 depicts the version 3.2 of the FOODIE core application schema. As such, it may also be re-used for an LPIS in any Member State of the European Union. The following paragraphs describe the main aspects of the open data model.

The developed open data model is as tightened to the standardization frameworks as possible. For that reason, it re-uses the data types defined in ISO standards (especially [16], [17] and [19]) as well as standardization efforts published under the INSPIRE Directive [9] (like structure of unique identifiers). Where feasible, the allowed units of the Measure data type should be limited to SI units or non-SI units accepted for use with the International System of Units.

The “Plot” feature type contains two kinds of data. The first kind could be considered as metadata about a “Plot” since it describes:

- *code*: unique identification of a “Plot” within the system; it is defined as an Identifier data type, i.e. a data type composed of a code (e.g. 178646A64BF), code space (MyFarm) and version in the form of date and time according to the ISO 8601 standard [19] (e.g. version 2015-02-18T18:09:07);
- *validFrom*: date (and time) when the “Plot” started to exist in the real world;
- *validTo*: date (and time) when the “Plot” no longer exists in the real world;
- *beginLifespanVersion*: date and time at which this version of the “Plot” was inserted or changed in the database;
- *endLifespanVersion*: date and time at which this version of the “Plot” was superseded or retired in the database;
- *geometry*: the geometry defining the spatial extent of the “Plot”;
- *description*: free text description of a “Plot” from a user point of view;
- *originType*: origin of the “Plot” when taking into consideration only two options: manual (created by human) and system (e.g. produced by the FOODIE platform).

The second kinds of attributes related to the “Plot” are intended to support the agronomical and environmental data, specifically:

- *production*: containing relevant production-related data, that is defined as a *ProductionType* data type comprising:
 - *productionDate*: date and time at which the information on production was inserted or changed in the database;
 - *variety*: an assemblage of cultivated individuals which are distinguished by any characters (morphological, physiological, cytological, chemical or others) significant for the purposes of agriculture;
 - *productionAmount*: the value of a physical quantity of the produced variety, together

with its unit according to the Measure data type as defined in ISO/TS 19103 [17];

- *productionAnalysis*: a general mechanism for analysis of all kinds of properties that may be related to the production; it comprises of:
 - *productionAnalysisDate*: date (and time) at which the analysis of production was conducted;
 - *property*: property that is being estimated by the analysis, together with its unit according to the Measure data type as defined in ISO/TS 19103 [17];
- *soil*: that is defined as a *SoilIdentification* feature type comprising:
 - *soilType* referring predominantly to the various sizes of mineral particles in the “Plot” expressed according to the national- and/or data provider- classification;
 - *soilAnalysis* that is defined as a *SoilAnalysisResultsType* data type comprising:
 - *dateOfAnalysis*: date (and time) at which the analysis (interpretation) of the soil properties was conducted;
 - *pH*: a measure of the acidity or basicity of a soil;
 - *organicMatter*: percentage of the organic matter component of soil, consisting of plant and animal residues at various stages of decomposition, cells and tissues of soil organisms, and substances synthesized by soil organisms;
 - *soilNutrients*: that is defined as a *SoilNutrientsType* data type, enabling to describe any soil nutrient, comprising:
 - *nutrientName*: full name of a (essential) nutrient like nitrogen or phosphorus;
 - *nutrientMeasure*: description of a method that was used to obtain the amount of the nutrient;
 - *nutrientAmount*: the amount of a nutrient, together with its unit according to the Measure data type as defined in ISO/TS 19103 [17];

- *electricConductivity*: electrical conductivity as the ability of a material to conduct (transmit) an electrical current in soil, together with its unit (typically milliSiemens per meter abbreviated as mS/m) according to the Measure data type as defined in ISO/TS 19103 [17];
- *notes*: any kind of further information on soil properties inserted by a user.

Furthermore, we may identify three associations between “Plot” on one hand and feature types *CropSpecies*, *Alert* and *Intervention* on the other hand.

- *CropSpecies*: identification of the planted crop species as a feature type comprising:
 - *beginDate*: date (and time) when the crop species started to be planted on the “Plot”;
 - *endDate*: date (and time) when the crop species ended to be planted on the “Plot”; typically date (and time) of harvest;
 - *cropArea*: the geometry defining the spatial extent of the crop species on the “Plot”;
 - *cropSpecies*: designation under which the crop species is commonly known;
- *Alert*: alerts automatically generated by the models integrated in the system as a feature type comprising:
 - *code*: unique identification of the *Alert* in the system;
 - *type*: type of *Alert* according to the user-defined classification, e.g. phytosanitary;
 - *description*: a brief narrative summary of the *Alert* content and rationale;
 - *checkedByUser*: indication whether the user is aware of the *Alert* or not (as Boolean values: true/false);
 - *alertDate*: date (and time) in which the *Alert* was created by the system;
 - *alertGeometry*: the geometry defining the spatial extent for which the *Alert* is applicable;
- *Intervention*: the basic feature type for any application with explicitly defined geometry comprising:
 - *type*: type of the intervention, e.g. tillage or pruning as a free text since it was not feasible to provide a common code list of all types of intervention, the types of intervention vary from country to country as well as from farmer to farmer;
 - *description*: human readable description of the type of intervention which may be a sub-

type of intervention or a broader description intended for the common understanding of the intervention;

- *notes*: any kind of further information on intervention inserted by a user;
- *status*: status of the intervention, e.g. approved, pending, completed as a free text since it was not feasible to provide a common code list of all statuses of intervention, the statuses and naming of an intervention vary according to the type of an intervention as well as from farmer to farmer;
- *creationDateTime*: date and time at which the intervention was inserted in the database;
- *interventionStart*: date (and time) when the intervention started in the real world;
- *interventionEnd*: date (and time) when the intervention ended in the real world;
- *interventionGeometry*: the geometry defining the spatial extent of the intervention;
- *supervisor*: a person or a body who has the power and authority to give instructions and guarantee the conducted intervention;
- *operator*: identification (typically at least name and surname) of a person (or several persons) who has conducted an intervention;
- *evidenceParty*: a person or a body who inserted the intervention into the database;

Intervention feature type has direct and indirect associations to the following types:

- *Treatment*: that is defined as a feature type comprising:
 - *quantity*: the value of a physical quantity of applied treatment, together with its unit according to the Measure data type as defined in ISO/TS 19103 [17];
 - *tractorId*: unique identification of a tractive vehicle for the machine applying the treatment; such unique identification is primarily used as a connection between the FOODIE core application schema and the Transport application schema; it is defined as an Identifier data type, i.e. a data type composed of a code of a tractor (e.g. TRCT00721), code space (e.g. Zetorka) and version in the form of date and time according to the ISO 8601 standard (e.g. version 2015-02-25T13:55:00);
 - *machineId*: unique identification of a machine applying the treatment; such unique identification is primarily used as a connection between the FOODIE core

application schema and the Transport application schema; it is defined as an Identifier data type, i.e. a data type composed of a code of a machine (e.g. MCHN005), code space (e.g. ManufacturerOfMyMachine) and version in the form of date and time according to the ISO 8601 standard (e.g. version 2015-02-25T13:56:04);

- *motionSpeed*: recommended speed for the application of the treatment that should be expressed together with its unit according to the Measure data type as defined in ISO/TS 19103 [17]; it is recommended to use km•h-1 as a *motionSpeed* unit;
- *pressure*: recommended pressure for the application of the treatment that should be expressed together with its unit according to the Measure data type as defined in ISO/TS 19103 [17];
- *flowAdjustment*: indication whether a flow adjustment was needed for the application of the treatment; should be expressed together with its unit according to the Measure data type as defined in ISO/TS 19103 [17];
- *applicationWidth*: a width in which a machine is capable of applying the treatment, should be expressed together with its unit according to the Measure data type as defined in ISO/TS 19103 [17]; it is recommended to use meters (m) as an *applicationWidth* unit;
- *areaDose*: the maximum application rate of the treatment, should be expressed together with its unit according to the Measure data type as defined in ISO/TS 19103 [17]; it is recommended to use kilograms (kg) or liters (l) per hectare as an *areaDose* unit;
- *formOfTreatment*: identification of the treatment application in the form of a code list comprising the following values: *manual* (i.e. conducted manually by a person), *applicationMachine* (i.e. automatic or semi-automatic application by a machine), *aerial* (i.e. aerial application from an airplane);
- *treatmentPurpose*: rationale why the treatment was used in the form of a code list comprising the following values: *weed* (i.e. to reduce plants considered undesirable in a particular situation), *pest* (i.e. a plant or animal detrimental to a human), *disease* (i.e. a pathological condition that affects part or all of the production);
- *treatmentDescription*: any further information related to the treatment that may facilitate understanding;

- *TreatmentPlan* that is defined as a feature type comprising:
 - *treatmentPlanCode*: unique identification of the treatment plan from the user point of view;
 - *description*: a brief narrative summary of the treatment plan content and rationale;
 - *type*: classification of a treatment plan from the user point of view; any classification system may be used;
 - *campaign*: a period to which the treatment plan was designed; typically a campaign may be an agronomical year or a season; a campaign is expressed as the *TM_Period* data type, i.e. an extent in time limited by two instances (beginning and ending), e.g. between 2015-03-01 and 2015-09-30;
 - *treatmentPlanCreation*: date and time at which the treatment was inserted in the database;
 - *notes*: any further information related to the treatment plan that may help for a better understanding;
- *ProductPreparation* that is defined as a feature type comprising:
 - *productQuantity*: the value of a physical quantity of the applied product, together with its unit according to the *Measure* data type as defined in ISO/TS 19103 [17];
 - *solventQuantity*: the value of a physical quantity for the solvent that was applied for the product application, together with its unit according to the *Measure* data type as defined in ISO/TS 19103 [17];
 - *safetyPeriod*: a period in which a dissolved product may be used; a *safetyPeriod* is expressed as the *TM_Period* data type, i.e. an extent in time limited by two instances (beginning and ending), e.g. between 2015-03-01 and 2015-03-15;
- *Product* that is defined as a feature type comprising:
 - *productCode*: unique identification of the product from the user point of view;
 - *productName*: designation under which the product is commonly known;
 - *productType*: type of the product, e.g. pesticide, as a free text since it was not feasible to provide a common code list of all types of products, the types of a product vary from country to country as well as from farmer to farmer;
 - *productSubType*: more detailed classification of a type of the product, e.g. biochemical insecticide, as a free text since it was not feasible to provide a common code list of all sub-types of product, the sub-types of product vary from country to country as well as from farmer to farmer;
 - *productKind*: origin of a product in the form of a code list comprising the following values: *organic* and *mineral*;
 - *description*: a brief narrative summary of the product;
 - *manufacturer*: identification of a producer of the product;
 - *nutrients*: identification of nutrients, i.e. chemical elements and compounds that are necessary for plant growth, defined as the *NutrientsType* data type comprising:
 - *N*: the amount of nitrogen, together with its unit according to the *Measure* data type as defined in ISO/TS 19103 [17];
 - *P2O5*: the amount of phosphorus pentoxide, together with its unit according to the *Measure* data type as defined in ISO/TS 19103 [17];
 - *K2O*: the amount of potassium oxide, together with its unit according to the *Measure* data type as defined in ISO/TS 19103 [17];
 - *MgO*: the amount of magnesium oxide, together with its unit according to the *Measure* data type as defined in ISO/TS 19103 [17];
 - *CaO*: the amount of calcium oxide, together with its unit according to the *Measure* data type as defined in ISO/TS 19103 [17];
 - *S*: the amount of sulphur, together with its unit according to the *Measure* data type as defined in ISO/TS 19103 [17];
 - *Zn*: the amount of zinc, together with its unit according to the *Measure* data type as defined in ISO/TS 19103 [17];
 - *Cu*: the amount of copper, together with its unit according to the *Measure* data type as defined in ISO/TS 19103 [17];
 - *Fe*: the amount of iron, together with its unit according to the

Measure data type as defined in ISO/TS 19103 [17];

- *B*: the amount of bismuth, together with its unit according to the Measure data type as defined in ISO/TS 19103 [17];
- *Mn*: the amount of manganese, together with its unit according to the Measure data type as defined in ISO/TS 19103 [17];
- *Mo*: the amount of molybdenum, together with its unit according to the Measure data type as defined in ISO/TS 19103 [17];
- *safetyInstructions*: information on the safe manipulation with the product;
- *storageHandling*: information on the safe storage of the product;
- *registrationCode*: unique identification of a product according to the national or any other relevant registration scheme;
- *registerUrl*: link to the national or any other relevant register where the product was registered;
- *ActiveIngredients* that is defined as a data type comprising:
 - *code*: unique identification of an active ingredient from the national or user point of view;
 - *ingredientName*: designation under which the active ingredient is commonly known;
 - *ingredientAmount*: the value of a physical quantity of applicable active ingredient, together with its unit according to the Measure data type as defined in ISO/TS 19103 [17].

The developed model was also transformed into the database schema for PostgreSQL (open source) database in version 9.3.6 together with its spatial extension PostGIS in version 2.1.0. Furthermore, the developed database schema was replicated in the Cloud (Infrastructure as a Service) provided by the Poznań Supercomputing and Networking Center (in Poland) using OpenStack as an Open Source Cloud Computing Software.

Verification of the data model has been conducted in two ways. The first one was population of the developed database schema by data from external systems, such as Land Parcel Identification System and cadaster. Information of a farmer and his/her parcels were successfully imported. The second verification has been conducted through historical data for last 10 years for two farms in the Czech Republic, called Vajglov (1'089 ha) and Tršice (1'291 ha). Currently, the developed data model is being verified on the present data for the season 2015.

IV. CONCLUSIONS

The open data model for (precision) agriculture applications and agricultural pollution monitoring has been successfully verified during demo tests in the *proof-of-concept* phase in the cloud environment. Such approach enables to disseminate the developed solution to all the interested stakeholders within the agricultural and environmental domains in the future. It may, therefore, replace proprietary data models used within the Farm management information systems. The biggest advantages are openness and interoperability with other systems such as (inter)national environmental and agronomical registers, customizability and scalability. The openness seems even more important than in the past since the commercial vendors propose that data measured by a commercial machine shall remain an ownership of such commercial vendor.

The presented open data model for (precision) agriculture applications and agricultural pollution monitoring has been together with the FOODIE platform registered under the GEOSS (Global Earth Observation System of Systems) Architecture Implementation Pilot – Phase 8 in order to support the wide variety of demands that are primary aimed at agriculture and water pollution monitoring. The data stored according to the developed data model may also be re-used for environmental monitoring of nitrogen, potassium, heavy metals leaching into water etc. The number of applications has not been limited since the data model enables to store any kind of data related to the environmental burden monitoring.

The developed data model was discussed with the Joint Research Centre (JRC) of the European Commission. The JRC is, among others, responsible for the development of the new data model for Land Parcel Identification Systems (LPIS). Implementation of the new LPIS data model is legally binding for the European Member States till 2018. We may state that the developed data model and the new LPIS model are complementary. In other words, it enables to import data from LPIS model into the developed data model as well as to export the corrected data back to the LPIS model.

The future work aims at explicit definition of processes in order to define the open and lightweight application programming interfaces (APIs) on the top of the open data model. Two types of APIs are foreseen for the development between 2015 and 2017.

The first ones are open APIs standardized by the International Organization for Standardization, World Wide Web Consortium and/or Open Geospatial Consortium. Their advantages lie in the area of interoperability, i.e. a possibility that the open data model may be integrated into a system accepting the above mentioned standards and best practices. On the other hand, we have to bear in mind the biggest disadvantage of open API – their performance criteria. It is common that the transmission of spatial data through the open APIs may require a long time to be finished for some requests.

For that reason, where appropriate, the open APIs are accompanied with the second type of provided APIs, i.e. the lightweight APIs. Their purpose is exactly the opposite. Lightweight APIs provide responses in a very short time; however, it comes at a price of interoperability. JavaScript

Object Notation (JSON) and similar formats are used at least to use the open standards for the data transmission between a server and web application.

One of the main open issues lies in the area that affects Big Data in all its forms. Especially farmers usually distrust the companies aggregating data. Farmers are afraid, that their sensitive detailed data may be misused. Future development would, therefore, be on the technological level as well as on the personal level to ensure the usefulness of the open data model as well as FOODIE platform in daily life.

V. ACKNOWLEDGEMENT

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No. 621074 called "Farm-Oriented Open Data in Europe" (FOODIE), from Horizon 2020 research and innovation programme under grant agreement No. 633945 called "Farming Tools for external nutrient Inputs and water Management" (FATIMa) and from the project No. MUNI/A/0952/2013 called "Analysis, evaluation and visualization of global environmental changes in the Component Earth's Spheres".

REFERENCES

- [1] I.J. Bateman, R. Brouwer, H. Davies, B.H. Day, A. Deflandre, S.D. Falco and T.R. Kerry, "Analysing the Agricultural Costs and Non-market Benefits of Implementing the Water Framework Directive" in *Journal of agricultural economics*, vol. 57, no. 2, 2006, pp. 221-237.
- [2] A. Bröring, J. Echterhoff, S. Jirka, I. Simonis, T. Everding, C. Stasch and R. Lemmens, "New generation sensor web enablement" in *Sensors*, vol. 11, no. 3, 2011, pp. 2652-2699.
- [3] O. Čerba, K. Charvát, J. Janečka, K. Jedlička, J. Ježek, T. Mildorf, "The Overview of Spatial Data Harmonisation Approaches and Tools" in *Proceedings of the 4th International Conference on Cartography and GIS*, 2012, vol. 1, pp. 113-124.
- [4] P.F. Donald, G. Pisano, M.D. Rayment and D.J. Pain, "The Common Agricultural Policy, EU enlargement and the conservation of Europe's farmland birds" in *Agriculture, Ecosystems & Environment*, vol. 89, no. 3, 2002, pp. 167-182.
- [5] J. Douglas, T. Usländer, G. Schimak, J.F. Esteban and R. Denzer, "An open distributed architecture for sensor networks for risk management" in *Sensors*, vol. 8, no. 3, 2008, pp. 1755-1773.
- [6] European Commission, "Council Directive 91/676/EEC of 12 December 1991 concerning the protection of waters against pollution caused by nitrates from agricultural sources" [online], 1991, available at URL <<http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:31991L0676&from=EN>>.
- [7] European Commission, "The Common Agricultural Policy: A partnership between Europe and Farmers" [online], 2012, available at URL <http://ec.europa.eu/agriculture/cap-overview/2012_en.pdf>.
- [8] European Commission, "Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy" [online], 2000, available at URL <<http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32000L0060&from=EN>>.
- [9] European Commission, "Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)" [online]. Official Journal L 108. Published 25th April 2007, p. 0001 – 0014. Cited 30th June 2011. Available at: <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:01:EN:HTML>>.
- [10] European Commission, "INSPIRE Data Specifications – Base Models – Activity Complex" [online], 2013, available at URL <http://inspire.ec.europa.eu/documents/Data_Specifications/D2.10.3_Activity_Complex_v1.0rc3.pdf>.
- [11] European Commission, "Data Specification on Agricultural and Aquaculture Facilities – Technical Guidelines" [online], 2013, available at URL <http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AF_v3.0.pdf>.
- [12] European Commission, "INSPIRE Generic Conceptual Model, version 3.4" [online], 2014, available at URL: <http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/D2.5_v3.4rc2.pdf>.
- [13] K. Feiden, F. Kruse, T. Řezník, P. Kubiček, H. Schentz, E. Eberhardt and R. Baritz, "Best Practice Network GS SOIL Promoting Access to European, Interoperable and INSPIRE Compliant Soil Information" in "ISESS 2011", vol. 359, J. Hřebíček, G. Schimak, and R. Denzer, R., Eds. Heidelberg:Springer, 2011, pp. 226–234
- [14] Food and Agriculture Organisation of the United Nations, "Chapter 1: Introduction to agricultural water pollution" [online], available at URL <<http://www.fao.org/docrep/W2598E/w2598e04.htm>>.
- [15] International Organization for Standardization, "ISO 19156:2011 Geographic information -- Observations and measurements", Geneva.
- [16] International Organization for Standardization, "ISO 19115-1:2014 Geographic information -- Metadata -- Part 1: Fundamentals", Geneva.
- [17] International Organization for Standardization, "ISO/TS 19103:2005 Geographic information -- Conceptual schema language", Geneva.
- [18] International Organization for Standardization, "ISO/IEC 10746-1:1998 Information technology -- Open Distributed Processing -- Reference model: Overview", Geneva.
- [19] International Organization for Standardization, "ISO 8601:2004 Data elements and interchange formats -- Information interchange -- Representation of dates and times", Geneva.
- [20] P. Kubiček, J. Kozel, R. Štampach, V. Lukas, "Prototyping the visualization of geographic and sensor data for agriculture" in *Computers and Electronics in Agriculture*, Elsevier, 2013, vol. 97, no. 9, pp. 83-91.
- [21] C.J.M. Ondersteijn, A.C.G. Beldman, C.H.G. Daatselaar, G.W.J. Giesen and R.B.M. Huirne, "The Dutch Mineral Accounting System and the European Nitrate Directive: implications for N and P management and farm performance" in *Agriculture, ecosystems & environment*, 2002, vol. 92, no. 2, pp. 283-296.
- [22] V. Pechanec, A. Vavra, M. Hovorkova, J. Brus and H. Kilianova, "Analyses of moisture parameters and biomass of vegetation cover in southeast Moravia", in *International Journal of Remote Sensing*, vol. 35, no. 3, pp. 967-987.
- [23] T. Řezník, "Geographic information in the age of the INSPIRE Directive: discovery, download and use for geographical information research" in *Geografie*, 2013, vol. 118, no. 1, pp. 77-93.
- [24] T. Řezník, R. Chudý, E. Mičietová, "Normalized evaluation of the performance, capacity and availability of catalogue services: a pilot study based on Infrastructure for SPatial InfoRmation in Europe (INSPIRE)", in *International Journal of Digital Earth*, Taylor & Francis, in press.
- [25] T. Řezník, V. Lukas, K.Charvát, Š. Horáková and K. Charvát jr., "Towards Farm-Oriented Open Data in Europe: the scope and pilots of the European project FOODIE" in *Agris on-line Papers in Economics and Informatics*, Faculty of Economics and Management CULS Prague2015, vol. 2015, no. 1, in press.
- [26] J. Sachs, R. Remans, S. Smukler, L. Winowiecki, S.J. Andelman, K.G. Cassman, D. Castle, R. DeFries, G. Denning, J. Fanzo, L.E. Jackson, R. Leemans, J. Lehmann, J.C. Milder, S. Naem, G. Nziguhaba, C.A. Palm, P.L. Pingali, J.P. Reganold, D.D. Richter, S.J. Scherr, J. Sircely, C. Sullivan, T.P. Tomich, and P.A. Sanchez, "Monitoring the world's agriculture" in *Nature*, 2010, vol. 466, no. 7306, pp. 558 – 560.
- [27] H. Skokanova, "Can we combine structural functionality and landscape services assessments in order to estimate the impact of landscape structure on landscape services?" in *Moravian geographical reports*, 2013, vol. 21, no. 4, pp. 2-14.

- [28] Z. Štěrba, Č. Šašinka, Z. Stachoň, "Usability testing of cartographic visualizations: principles and research methods" in Proceedings of the 5th International Conference on Cartography and GIS Proceedings, 2014, vol. 1 and vol. 2, pp. 147-256.