TheFSM | The Food Safety Market

The Food Safety Market: An SME-powered industrial data platform to boost the competitiveness of European food certification

D2.1 - Data Models & Representations

DELIVERABLE NUMBER	D 2.1
DELIVERABLE TITLE	Data Models & Representations
RESPONSIBLE AUTHOR	Svetla Boytcheva (SAI)

Co-funded by the Horizon 2020
Framework Programme of the European Union

GRANT AGREEMENT N.	871703
PROJECT ACRONYM	TheFSM
PROJECT FULL NAME	The Food Safety Market: An SME-powered industrial data platform to boost the competitiveness of European food certification
STARTING DATE (DUR.)	01/02/2020 (36 months)
ENDING DATE	31/01/2023
PROJECT WEBSITE	www.foodsafetymarket.eu
COORDINATOR	Nikos Manouselis
ADDRESS	110 Pentelis Str., Marousi, GR15126, Greece
REPLY TO	nikosm@agroknow.com
PHONE	+30 210 6897 905
EU PROJECT OFFICER	Stefano Bertolo
WORKPACKAGE N. | TITLE	WP2 | Data
WORKPACKAGE LEADER	Sirma AI
DELIVERABLE N. | TITLE	D 2.1 | Data Models & Representations
RESPONSIBLE AUTHOR	Svetla Boytcheva (SAI)
REPLY TO	svetla.boytcheva@ontotext.com
DOCUMENT URL	
DATE OF DELIVERY (CONTRACTUAL)	31 January 2021 (M12)
DATE OF DELIVERY (SUBMITTED)	31 January 2021 (M12)
VERSION | STATUS	V1.0 | Final
NATURE	Other (OT)
DISSEMINATION LEVEL	Public (P)
AUTHORS (PARTNER)	Svetla Boytcheva (SAI), Plamen Tarkalanov (SAI), Nikola Tulechki (SAI), Pavlin Gyurov (SIR), Nikola Rusinov (SAI)
CONTRIBUTORS	Branimir Rakic (PROSPEH)
REVIEWER	Danai Vergeti (UBITECH)

VERSION	MODIFICATION(S)	DATE	AUTHOR(S)
0.1	Table of contents	15/11/2020	Svetla Boytcheva (SAI)
0.2	Initial draft version	15/12/2020	Plamen Tarkalanov (SAI), Pavlin Gyurov (SIR), Nikola Rusinov (SAI)
0.3	Sections contribution, corrections, and observations	21/01/2021	Plamen Tarkalanov (SAI), Nikola Tulechki (SAI), Pavlin Gyurov (SAI), Nikola Rusinov (SAI)
0.4	Document refinement, addition of figures and descriptions	28/01/2021	Svetla Boytcheva (SAI)
0.5	Minor corrections	29/01/2021	Branimir Rakic (PROSPEH)
0.6	Internal Review	29/01/2021	Danai Vergeti (UBITECH)
1.0	Final version	31/01/2021	Svetla Boytcheva (SAI)

PARTNERS		CONTACT
Agroknow IKE (Agroknow, Greece)	Agroknow	Nikos Manouselis (Agroknow) nikosm@agroknow.com
SIRMA AI EAD (SAI, Bulgaria)	ontotext	Svetla Boytcheva (SAI) svetla.boytcheva@ontotext.com
GIOUMPITEK MELETI SCHEDIASMOS YLOPOIISI KAI POLISI ERGON PLIROFORIKIS ETAIREIA PERIORISMENIS EFTHYNIS (UBITECH, Greece)	UBITECH ubiquitous solutions	Danai Vergeti (UBITECH) vergetid@ubitech.eu
AGRIVI DOO ZA PROIZVODNJU, TRGOVINU I USLUGE (Agrivi d.o.o., Croatia)	AGRIVI	Tanja Matosevic (Agrivi d.o.o.) tanja.matosevic@agrivi.com
PROSPEH, POSLOVNE STORITVE IN DIGITALNE RESITVE DOO (PROSPEH DOO, Slovenia)	tracelabs origintrail Core Developers	Ana Bevc (PROSPEH DOO) ana.bevc@tracelabs.io
UNIVERSITAT WIEN (UNIVIE, Austria)	universität wien	Elizabeth Steindl (UNIVIE) elisabeth.steindl@univie.ac.at
STICHTING WAGENINGEN RESEARCH (WFSR, Netherlands)	WAGENINGEN UNIVERSITY & RESEARCH	Yamine Bouzembrak (WFSR) yamine.bouzembrak@wur.nl
TUV- AUSTRIA ELLAS MONOPROSOPI ETAIREIA PERIORISMENIS EUTHYNIS (TUV AU HELLAS, Greece)	TÜV AUSTRIA HELLAS	Kostas Mavropoulos (TUV AU HELLAS) konstantinos.mavropoulos@tuv.at
TUV AUSTRIA ROMANIA SRL (TUV AU ROMANIA, Romania)	TÜV AUSTRIA ROMANIA	George Gheorghiou (TUV AU Romania) george.gheorghiu@tuv.at
VALORITALIA SOCIETA PER LA CERTIFICAZIONE DELLE QUALITA'E DELLE PRODUZIONI VITIVINICOLE ITALIANE SRL (VALORITALIA, Italy)	VALORITALIA	Francesca Romero (Valoritalia) francesca.romero@valoritalia.it
TUV AUSTRIA CYPRUS (TUV AU CYPRUS, Cyprus)	TÜV AUSTRIA CYPRUS	Sousanna Charalambidou (TUV AU CYPRUS) sousanna.charalambidou@tuv.at

ACRONYMS LIST

API	Application Programming Interface
GTIN	Global Trade Item Number
JSON	JavaScript Object Notation
RDF	Resource Description Framework
RDBMS	Relational Data Base Management System
REST	Representational state transfer
SHACL	Shapes Constraint Language
SKOS	Simple Knowledge Organization System Primer
TheFSM	The Food Safety Market
XML	Extensible Markup Language

EXECUTIVE SUMMARY

This document presents the initial version of the data models developed within The Food Safety Market (TheFSM) project and their machine-readable representations. In the document we briefly present data types, data sources, and data formats. Special attention is paid to the domain-specific ontologies, standard classifications, and vocabularies. The result data models are based on the desk research and analysis, based on data inventory of publicly available datasets, as well as dataset samples provided by TheFSM consortium. The detailed analysis of data requirements of all use cases and certification schemes serves as a basis for TheFSM integrated data model. Developed machine-readable data models are in compliance with the industry-accepted standards in order to ensure interoperability of TheFSM platform. This document will be updated twice at M24 and M36 to show the progress in terms of data interoperability and coverage.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 Scope

The document describes the design of the initial version of data models developed within The Food Safety Market (TheFSM) project and their machine-readable representations. The data model design is organized in 3 iterations, aligned with the three project milestones MS3 "1st Release of TheFSM Data Platform and Applications" (M12), MS5 "2nd Release of TheFSM Data Platform and Applications" (M24) and MS8 "Final Release of TheFSM Data Platform and Applications" (M36). In the first iteration of the data model design we analyzed data from data inventory and conducted desk research of available domain-related standards and ontologies. The data inventory contains both data from public resources and data samples provided by the FSM project consortium. For the 2nd and 3rd iterations are planned activities for improvement of the data model based on additional data sources, APIs and more domain-specific ontologies.

1.2 Data

The following types of data were considered as relevant – grouped by availability (public/proprietary) and information organization in different data formats (structured/unstructured) with some examples.

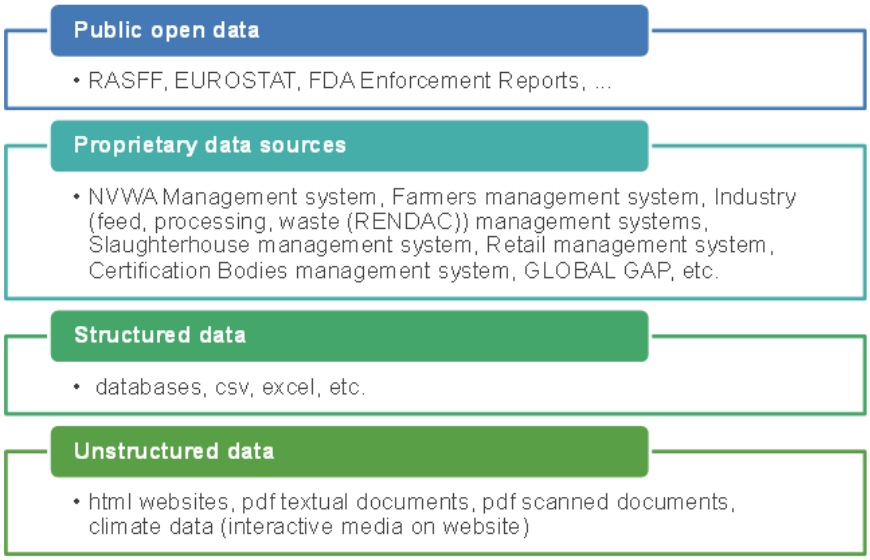

Figure 1 Types of datasets

For public open data we need to identify the scope and the particular subsets that are relevant to the project. For example, EUROSTAT is too general, and we need to specify the particular statistical data that will be analyzed.

For proprietary data we need data samples to be shared in order to analyze the data structure and to envision the necessary data services that will be developed for such data processing.

For unstructured data we need strong justification of the particular data source and lack of any other option to gather the essential data from other data providers in a structured format.

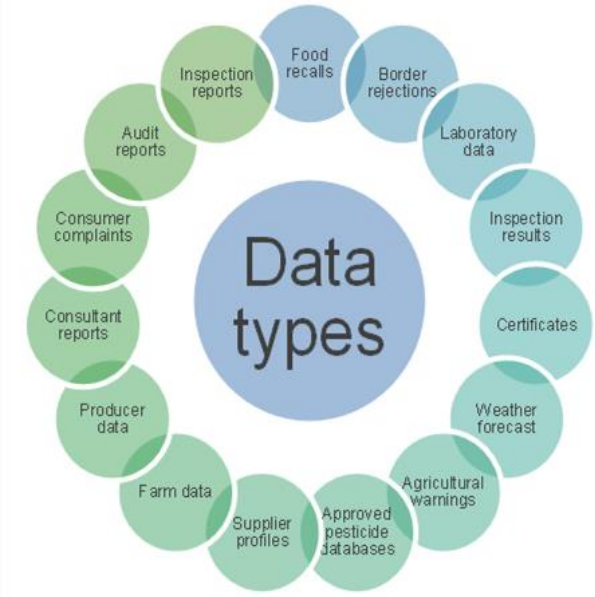

Figure 2 Data types

1.3 Relation to other deliverables

The developed data model is based on the detailed analysis of business, data and technical requirements of all use cases described in D1.1 "Report on Requirements for TheFSM" (delivered on M4), in compliance with D2.2 : "Data Services" (delivered on M12), D2.3 "Report on Data Population" (delivered on M12) and D3.1 "TheFSM Open Reference Architecture" (delivered on M12). The best practices, state-of-the-art and standards are taken in consideration as well.

1.4 Outputs of the deliverable

This deliverable provides description of the outcomes of T2.2 - the initial version of the data model. The current version of machine-readable representation of the developed data model is available at Github repository: https://github.com/OriginTrail/epcis-erm

The Github repository contains definition of heFSM data model in various formats and the supplementary documentation. Some screenshots of this repository are provided below.

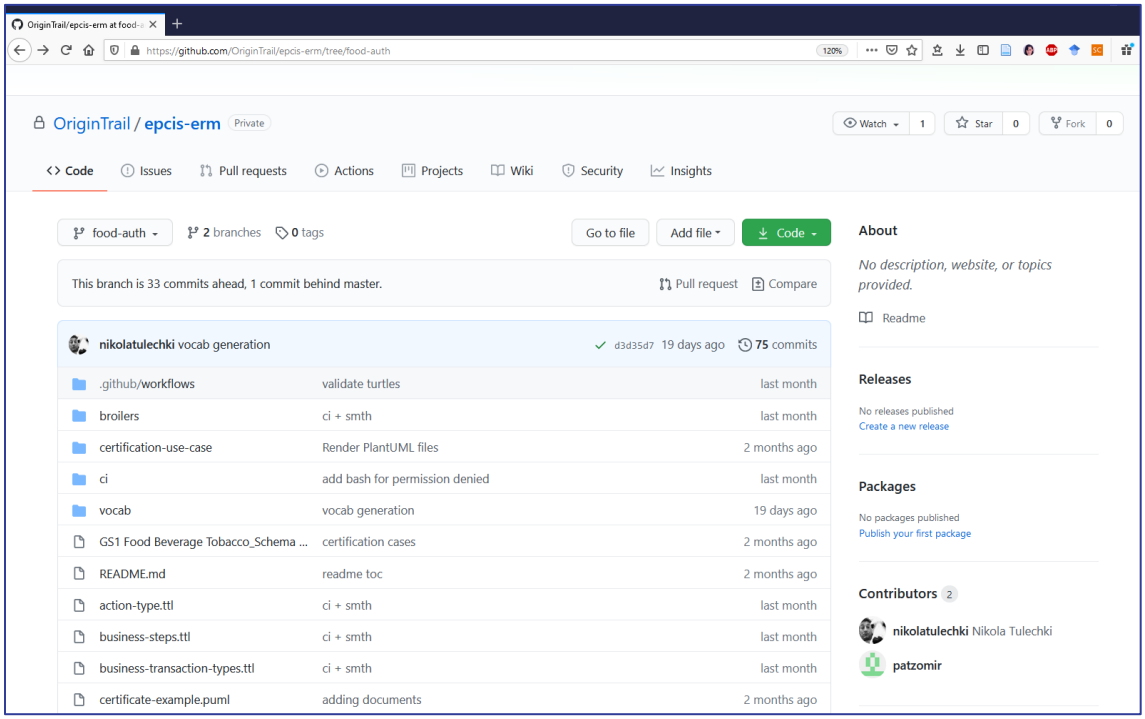

Figure 3 Github repository – food-auth branch

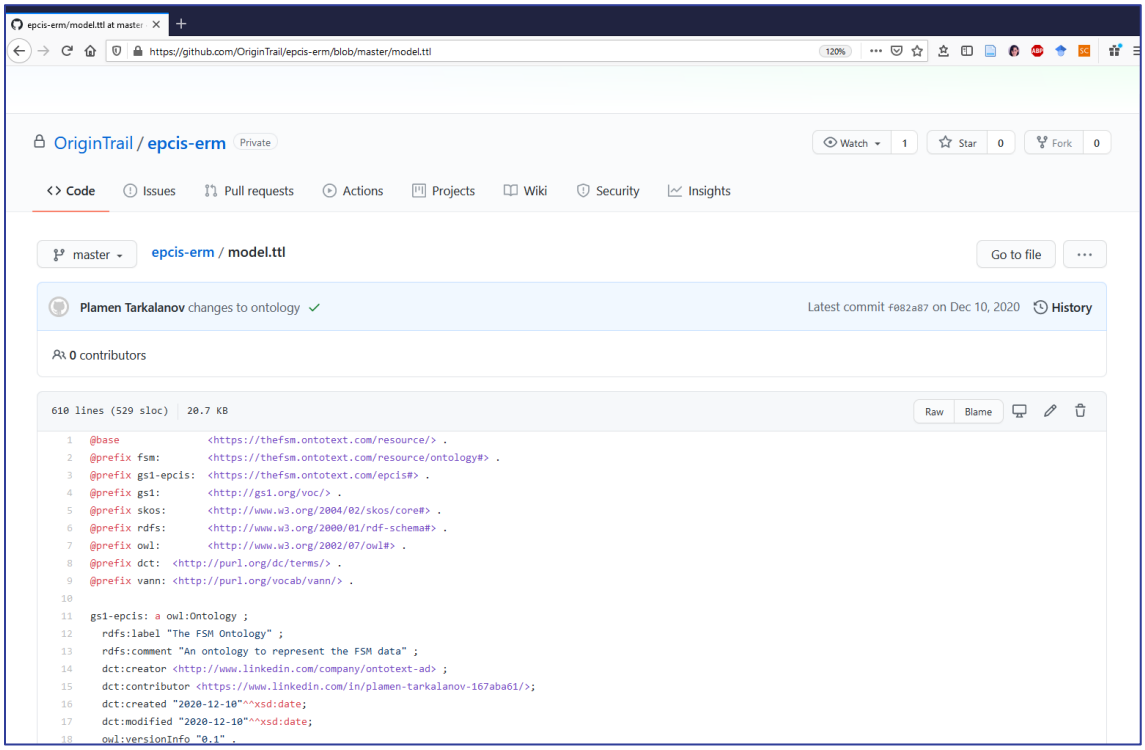

Figure 4 Github repository – epcis-erm/model.ttl

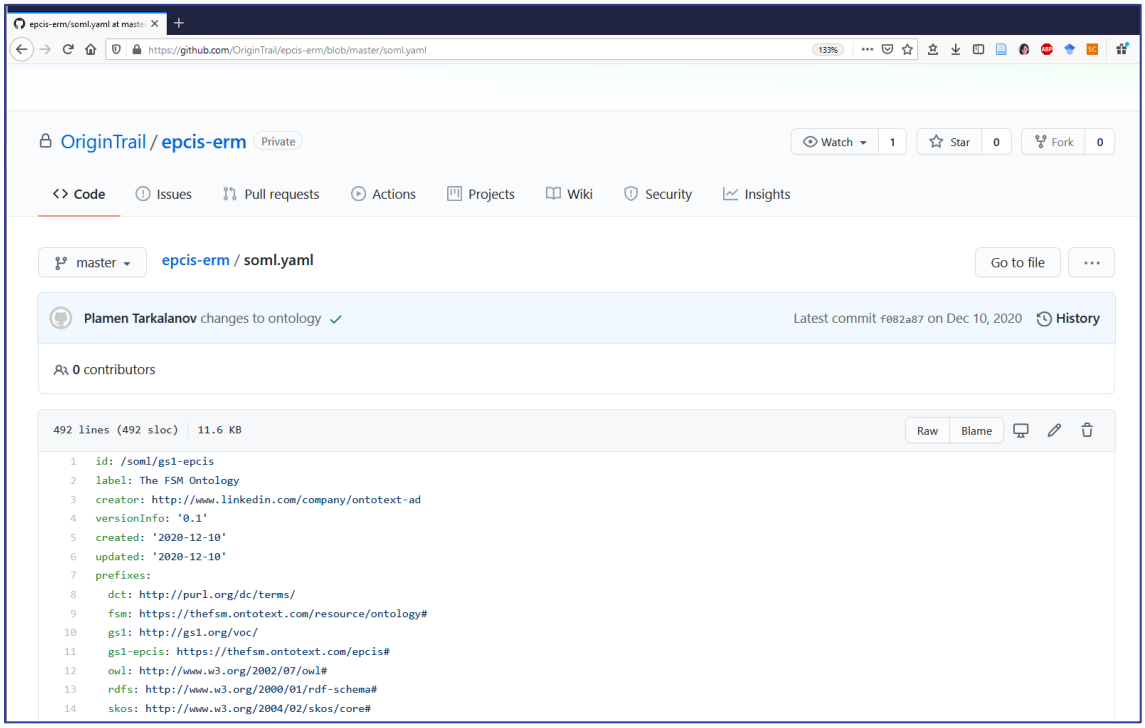

Figure 5 Github repository – epcis-erm/soml.yaml

2 DOMAIN SPECIFIC ONTOLOGIES

2.1 Standards and ontologies

Domain-specific ontologies are a special focus in data modeling. Some of the most important ones are listed in the table below.

Table 1 Data Standards and Ontologies

AgroVoc[1]	A controlled vocabulary for describing food, nutrition, agricultural, marine, forestry, environmental information. It is also part of the GACS initiative, which aims to map the core concepts of three major thesauri AgroVoc, CAB, and NAL.
CAB Thesaurus[2]	CAB Thesaurus is the essential search tool for all users of CAB Abstracts and Global Health databases and related products. It provides a controlled vocabulary resource of over 2.9 million descriptive terms, for the applied life sciences including: agriculture, forestry, horticulture, soil science, entomology, mycology, parasitology, veterinary medicine, nutrition, rural studies.
NAL Thesaurus[3]	The National Agricultural Library's Agricultural Thesaurus is now available as Linked Open Data. The Thesaurus and Glossary are online vocabulary tools of agricultural terms in English and Spanish and are cooperatively produced by the National Agricultural Library[4], USDA[5], and the Inter-American Institute for Cooperation on Agriculture, as well as other Latin American agricultural institutions belonging to the Agriculture Information and Documentation Service of the Americas[6] (SIDALC). NAL Thesaurus is used to select controlled vocabulary terms for subject indexing of AGRICOLA, PubAg and other databases.
ICASA	Data format for documenting experiments and modelling crop growth and development, facilitating exchange of information and software.

[1] http://aims.fao.org/vest-registry/vocabularies/agrovoc
[2] https://www.cabi.org/cabthesaurus/
[3] https://agclass.nal.usda.gov/
[4] http://www.nal.usda.gov/
[5] http://www.usda.gov/
[6] https://www.iica.int/en

	• The ICASA V2.0 can be found in this paper[7] • The current update version (ICASA V3.0) can be requested from the DSSAT development team[8]:
Food Safety Knowledge Markup Language[9] (FSK-ML)	In development with the participation of BfR, aiming to describe data and models relevant to risk assessment tasks. With the Food Safety Knowledge Markup Language (FSK-ML) we now extend the PMF-ML format to enable the exchange of knowledge / information that is embedded in specific script-based programming languages (e.g., "R", Matlab, Python). The FSK-ML guidance document primarily aims at harmonizing the exchange of food safety knowledge (e.g. predictive models) including the associated metadata where this knowledge is only available in a software dependent format. • In the provided link, there are tutorials and a developer guide pdf document[10] • One can also read about the FSK-Lab - an open source food safety model integration tool[11] -
FoodEX2[12]	The European Food Safety Authority (EFSA) has developed a standardized food classification and description system called FoodEx2. EFSA has established a collaboration for the use of FoodeEx2 with several institutions. Among these, the Food and Agriculture Organisation (FAO) of the United Nations and the Friedman School of Nutrition Science and Policy at Tufts University. • In the provided link, there are YouTube tutorials for the software.
GeoNames[13]	Should be used as an international standard for country codes.
eCl@ss Classification and Product Description[14]	An ISO/IEC-compliant industry reference-data standard for the classification and unambiguous description of products and

[7] https://dssat.net/wp-content/uploads/2014/02/White2013ICASA_V2_standards.pdf

[8] https://dssat.net/data/standards_v2/#:~:text=IBSNAT%20Data%20Standards&text=The%20ICASA%20standards%20have%20evolved,original%20plain%20text%20file%20format

[9] https://foodrisklabs.bfr.bund.de/fsk-ml-food-safety-knowledge-markup-language/

[10] https://foodrisklabs.bfr.bund.de/wp-content/uploads/fsk/FSK_guidance_document_V3_1.pdf.

[11] https://foodrisklabs.bfr.bund.de/fsk-lab/.

[12] https://www.efsa.europa.eu/en/data/data-standardisation

[13] https://www.geonames.org/

[14] https://www.eclass.eu/fileadmin/downloads/eCl%40ss_company_brochure.pdf

	services. The classification establishes a uniform semantic standard, enabling Internet of Things and product master data to be exchanged digitally across all borders – across sectors, countries, languages and organizations.
XACML OASIS Standard	In TheFSM, access to data assets is regulated through Attribute-Based Access Control (ABAC) policies, based on the XACML OASIS standard, that allows the data providers to protect and share their data assets, even when they do not have any prior knowledge of the potential individual data consumers in the food certification data value chain. XACML promotes common terminology and interoperability between access control implementations by multiple vendors.
GS1[15]	GS1 standards provide a common language to identify, capture and share supply chain data– ensuring important information is accessible, accurate and easy to understand. GS1 Global Data model[16] plays important role in interoperability. GS1 vocabulary has many props related to food[17] and perishable goods. Another benefit of using GS1 is the Global Location Number[18] (GLN) is used to identify locations and legal entities. This unique identifier is comprised of a GS1 Company Prefix, Location Reference, and Check Digit. Another specification provided by GS1 is Global Trade Item Number[19] (GTIN) – it can be used by a company to uniquely identify all of its trade items.
EPCIS[20]	Electronic Product Code Information Services (EPCIS) is a global GS1 Standard for creating and sharing visibility event data. The goal of EPCIS is to enable disparate applications to leverage Electronic Product Code (EPC) data via EPC-related data sharing, both within and across enterprises. Ultimately, this sharing is aimed at enabling participants in the EPC[21] global Network to gain a shared view of the disposition of EPC-bearing objects within a relevant business context.

[15] https://www.gs1.org/
[16] https://www.gs1.org/standards/gs1-global-data-model
[17] https://www.gs1.org/industries/retail/fresh-foods
[18] https://www.gs1us.org/DesktopModules/Bring2mind/DMX/Download.aspx?Command=Core_Download&EntryId=158&language=en-US&PortalId=0&TabId=134#:~:text=The%20Global%20Location%20Number%20(GLN)%20is%20used%20to%20identify%20locations,is%20a%20GLN%20used%20for%3F
[19] https://www.gs1.org/standards/id-keys/gtin
[20] https://www.gs1.org/epcis/epcis/1-1
[21] https://www.gs1.org/sites/default/files/docs/epc/EPCIS-Conformance_Requirements.pdf

2.2 Metadata

In addition to the data collection activities, TheFSM will also generate its own valuable data assets in terms of metadata that will improve the description, interlinking, normalization, unification, and quality assessment of the collected datasets. The use of W3C standards such as PROV-O for provenance and DCAT for data catalog description will be encouraged.

- PROV-O[22]: The PROV Ontology -The PROV Ontology is expected to be both directly usable in applications as well as serve as a *reference model* for creating domain-specific provenance ontologies and thereby facilitates interoperable provenance modeling.
- Data Catalog Vocabulary (DCAT) - Version 2[23] - DCAT is a vocabulary for publishing data catalogs on the Web, which was originally developed in the context of government data catalogs such as data.gov and data.gov.uk, but it is also applicable and has been used in other contexts.

2.3 W3C Cube ontology

In order to express statistical data such as laboratory measurements, meteorological data, or any other dataset resulting from a systematized observation of project-relevant entities, we require a proper representation of data. Whenever necessary, data that is commonly expressed (and exchanged) as spreadsheets will be transformed into a homogeneous semantic model and described using the **W3C Cube** ontology[24]. This process will ensure high interoperability of observation data as well as easier maintenance and backward compatibility of novel datasets.

The W3C CUBE ontology (QB) captures multidimensional observations (data cubes) using the following terminology (in bold). We map roughly these QB terms to the data items discussed in previous sections and expressed using tGS1 and EPCIS. The first 3 are called "components".

- **Dimension**: entity, quality, method
- **Attribute**: unit, context
- **Measure**: value
- **Observation** = Dimensions + Attributes + Measures

QB defines what components are expected in a specific ***qb:DataSet*** by using a ***qb:DataStructureDefinition***. QB provides some flexibility that affords data efficiencies, and avoiding combinatorial explosion:

- QB allows using several dimensions per observation, without tying them up together. E.g., you can use 3 dimensions **entity="plant", quality="height", method="whole height"**

[22] https://www.w3.org/TR/prov-o/#introduction
[23] https://www.w3.org/TR/vocab-dcat/#introduction
[24] https://www.w3.org/TR/vocab-data-cube/

- We could also use several measures per observation (e.g. as taken by a combination instrument), although this is less commonly used.
- One could split a dataset into Slices (or other kinds of ObservationGroups) by fixing some of the dimensions, so one doesn't need to repeat them for every observation.

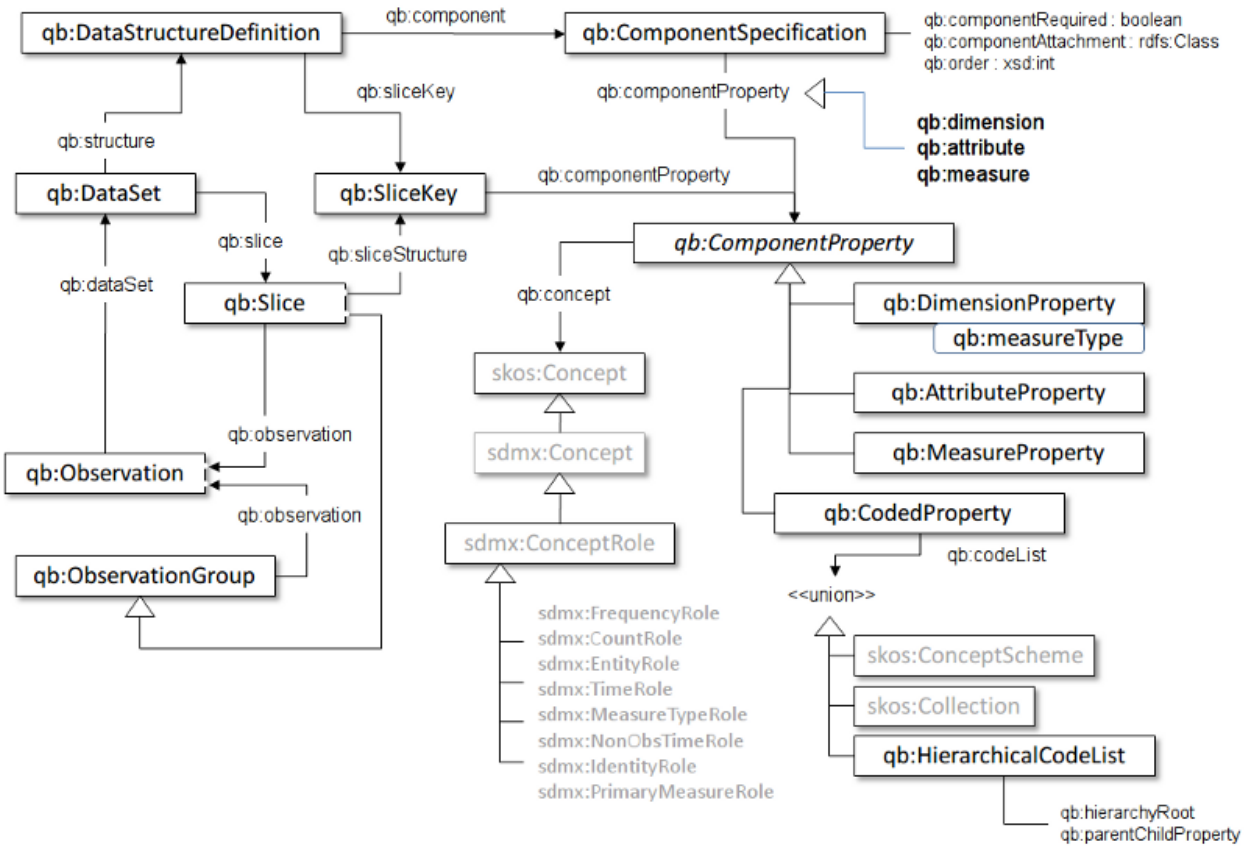

Figure 6 W3C CUBE Ontology

Adapting the model to the project-specific entities will be done by defining a large number of classes. In order to both streamline this process and ensure that it can be done collaboratively by domain experts and in a fully decentralized fashion, we employ a technique using shared Google sheets and automatic ontology generation by converting the data from the sheets into the final model. This setup is described in D2.2.

2.4 SKOS

Simple Knowledge Organization System[25] (SKOS) is a W3C recommendation designed for representation of thesauri, classification schemes, taxonomies, subject-heading systems, or any other type of structured controlled vocabulary and will be used to represent relevant entities that

[25] https://www.w3.org/TR/skos-reference/

are not covered by other schemas such as GS1 and EPCIS. The ontology generation pipeline described in D2.2 also covers creating small and medium Certificate and certificate families

2.5 Semantic Data Integration

Semantic Data Integration is one of the best ways to integrate diverse data across organizations and enterprises, and to exploit available datasets both as commercial data and in the LOD cloud. Initially adopted by Life Science and Biology researchers, semantic web techniques have by now found a wide following, notably in the Agro sector, which in many cases leverages ontologies developed in the Bio community.

Semantic Data Integration aims to harmonize data from heterogeneous sources often from various providers and convert it to a semantic form. A substantial part of the activity is to match instances about the same entity coming from different databases (aka coreferencing). The final result is an integrated Knowledge Graph of data in a domain. The following steps, which will be part of WP2 activities are part of the process:

- Sampling tabular data from project partners
- Sampling of RDF data from partners
- Data analysis
- Listing data requirements and competence questions
- State of the art research. listing existing ontologies and other related ontologies
- Selection, combination and extension of ontologies (Ontology engineering)

To date the progress of WP2 w.r.t these activities are as follows:.

- We have discussed how to represent various data aspects with partners: product workflow, data and documents, measurements/observations, equipment, experiments, etc
- Creating a semantic model with rdfpuml and text narrative
- Create application profiles and/or RDF shapes[26] (SHACL[27] and/or ShEx[28]) for validation of semantic data for conformance to the model
- Define URL design and policies
- Semantic conversion using appropriate tools depending on the source (CSV/TSV tabular, RDBMS, XML)
- Alignment of data and instance matching
- Validation and data quality management/measurement
- Sample query creation

[26] https://www.w3.org/2014/data-shapes/
[27] https://www.w3.org/TR/shacl/
[28] https://shex.io/

3 INTEROPERABILITY PER USE CASE SCENARIO

In this section are presented TheFSM integrated data model and some implementations for TheFSM use cases scenarios: food supply chain, retailer scenario, certification body scenario and food safety authority scenario. In order to ensure interoperability for the FSM platform, the data model is based on standards like GS1, EPCIS, W3C and domain-specific ontologies.

3.1 The FSM integrated model

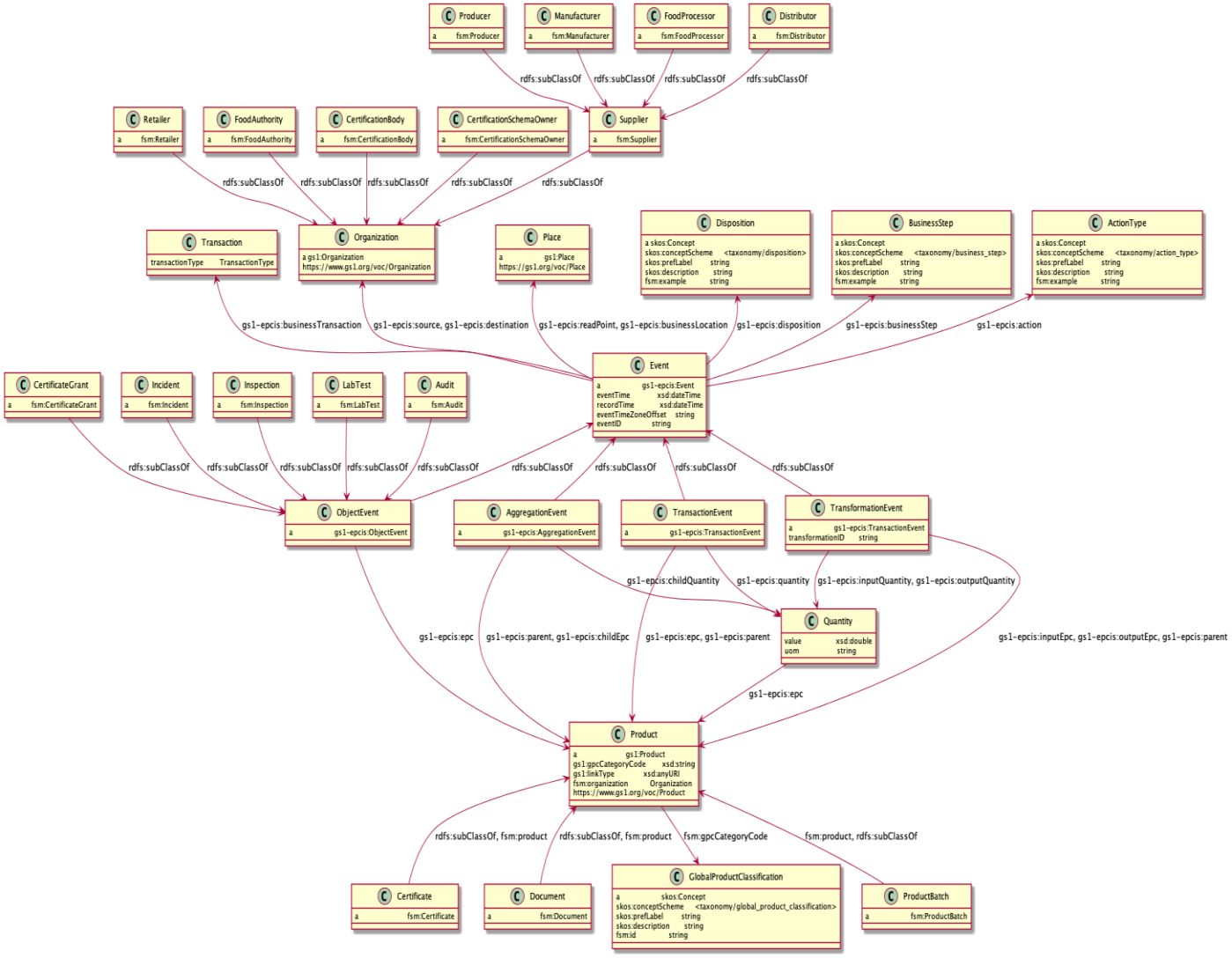

Figure 7 FSM integrated data model

The integrated model extends other ontologies. The classes which are specific to the FSM platform use cases are made subclasses of different classes from the extended ontologies. By using the

subclass property we enable all unique FSM objects and data to be read through the ontology classes thus enriching the objects and improving interoperability with other systems.

The list of subclasses:

1. The `CertificateGrant`, `Incident`, `Inspection`, `LabTest` and `Audit` are made subclasses of the EPCIS `ObjectEvent`
2. `Producer`, `Manufacturer`, `FoodProcessor`, `Distributor` are subclasses of `Supplier` which together with `Retailer`, `CertificationBody`, `CertificationSchemaOwner`, `FoodSafetyAuthority` is a subclass of GS1 Organization.
3. The `Certificate`, `Document` and `ProductBatch` classes are subclasses of GS1 Product

3.1.1 Certification

In order to express certification of producers, food processors and other parties involved in the supply chain we need to handle two generic cases.

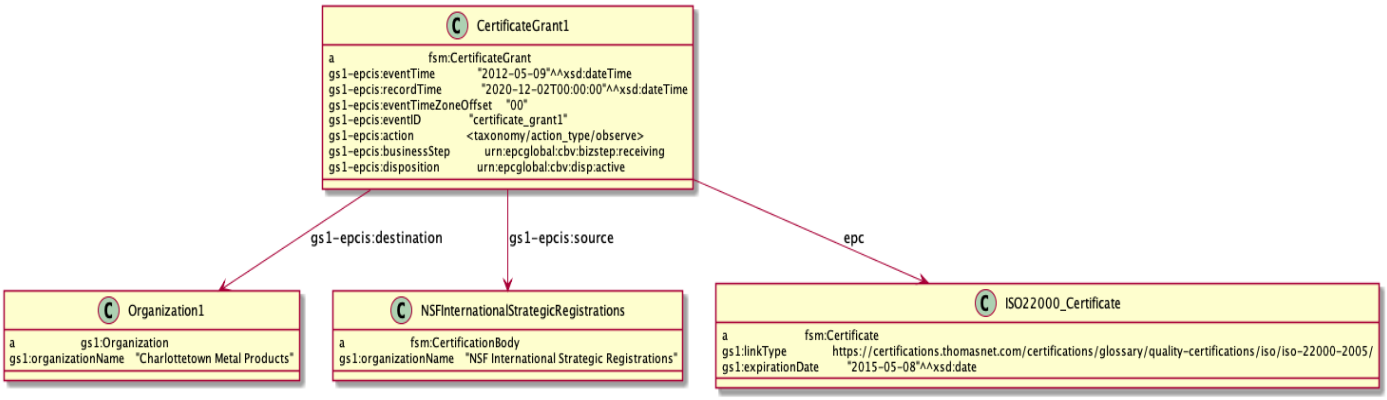

Figure 8 Organization level certification

In this example we are expressing an ISO 22000:2005[29] certificate. We treat the certificate as a product of the Certification Body. The certification body gives the certificate to the Organization.

[29] https://www.iso.org/iso-22000-food-safety-management.html

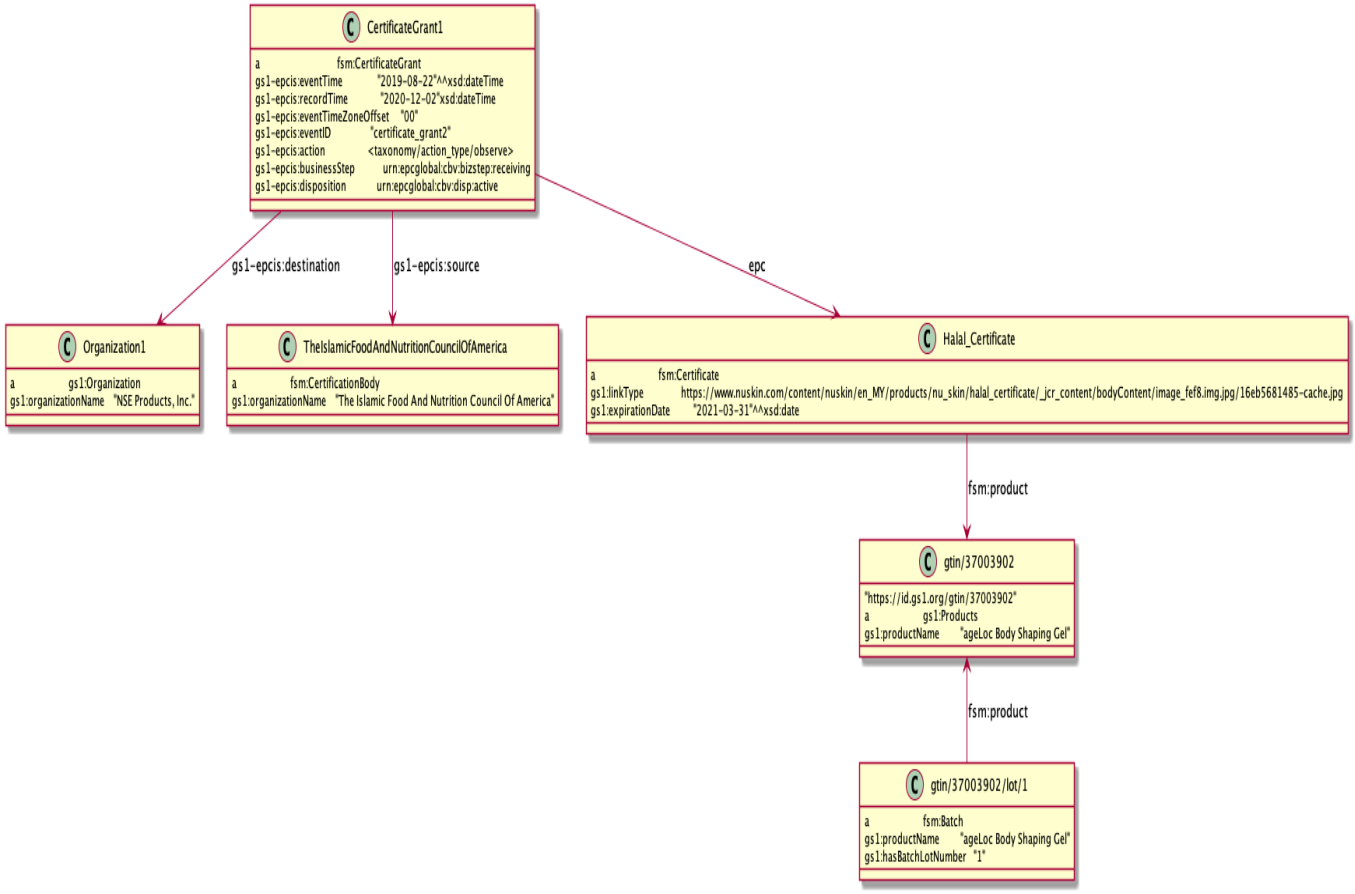

Figure 9 Product level certification

In this example, we are expressing a Halal certificate. These certificates are provided only for some products provided by a company.

We treat the certification grant in the same manner as the above example.

The new thing is that the certificate node links to a product. This product is a GTIN prefix for a company product. All batches of this product will be linked to it with the `fsm:product` link.

3.1.2 Documents

The documents are represented in a similar manner as the certifications. We treat as documents the following classes:

- Incidents
- Audits
- Inspection

- Lab tests

They can be expressed using the same logic.

In this example we will use an incident from the FOODAKAI[30] API.


```
"id":"FDK_18341544",
'                  "title":"Other hazard in ready to eat - cook meals by «  1001
France » from France",
'                  "description":"Date : 29 NOVEMBRE 2020\r\nINFORMATION
CONSOMMATEURS\r\nRAPPEL DE PRODUIT\r\nLa Société «  1001 FRANCE » procède
aujourd'hui au retrait de la vente de sa \r\nrecette « Mon mijoté de courges
butternut et bœuf »  suite à la mise en évidence de \r\npossibles morceaux de
graines de courges dans le produit.\r\nIl s'agit des lots portant les
caractéristiques suivantes : \r\nNature du Produit   :  Repas complet pour
bébé\r\nMarque : HAPPYLAL BABY\r\nGENCOD : 3770007731081\r\nFORMAT   :   220
grammes\r\nDLC   :  15/03/2021 (LOT : MEL46V11) et 09/05/2021(LOT :
MEL02B06)\r\nCode emballeur/Estampille Sanitaire   (ou numéro de lot) :  FR
29.174.020 CE\r\nL'ensemble des lots sont retirés de la
commercialisation.\r\nCertains de ces produits ont cependant été commercialisés
avant la mesure de \r\nretrait.\r\nIl est donc recommandé aux personnes qui
détiendraient des produits appartenant aux \r\nlots  décrits  ci-dessus  de ne pas
les  consommer et  de les  détruire  ou  de les \r\nrapporter au point de vente.
\r\nLa société « 1001 FRANCE » se tient à la disposition des consommateurs pour
\r\nrépondre à leurs questions au numéro de téléphone : « 06 62 26 38 63
».\r\nAfficher jusqu'au 29 décembre 2020",
'                  "entityType":"incident",
'                  "createdOn":"2020-11-29T00:00:00",
'                  "updatedOn":"2020-11-30T14:05:06.916673",
'                  "dataSource":"FOODAKAI",
'                  "tags":[
'                     "france",
'                     "ready to eat - cook meals",
'                     "other hazard",
'                     "prepared dishes and snacks",
'                     "europe"
'                  ],
'                  "published":true,
'                  "privateData":0,
'                  "linkedEntities":[
'                     {
'                        "id":"FDK_18341542",
'                        "title":"«  1001 France »",
'                        "description":"",
```


[30] https://www.foodakai.com/

```
'                      "entityType":"supplier",
'                      "createdOn":"2020-11-30T14:05:06.916704",
'                      "updatedOn":"2020-11-30T14:05:06.916707",
'                      "dataSource":"FOODAKAI",
'                      "tags":[

'                      ],
'                      "published":true,
'                      "privateData":0,
'                      "linkedEntities":[

'                      ],
'                      "internalId":null
'                   }
'                ],
'                "internalId":"18341544"
'             },
```


Figure 10 Example for an incident from FOODAKAI

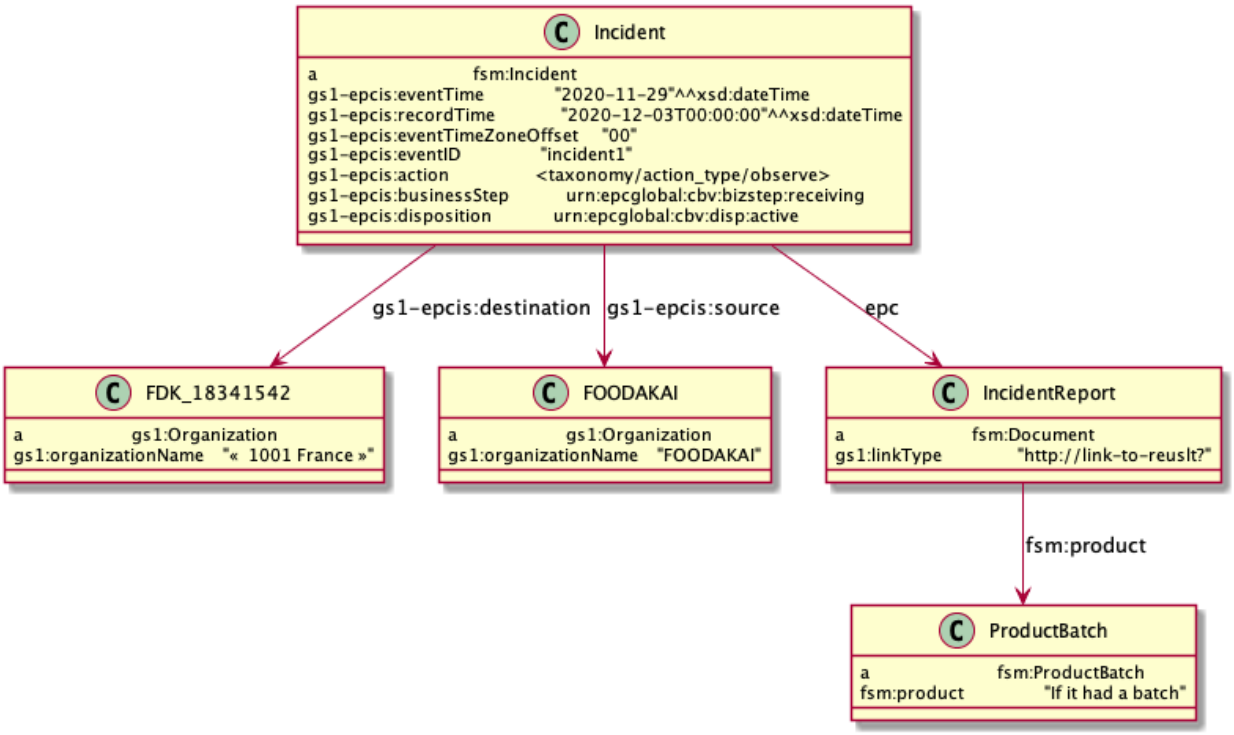

Figure 11 Document example

We treat the FOODAKAI as the source organization of the event as such information is not provided by FOODAKAI and no one outside Agroknow is able to validate it. The example incident has structured information only for the organization which supplied the food product that has led

to the incident. In the FSM Platform, we would like to be able to identify a product or a batch of products. In order to accommodate for such granularity, we use the `fsm:product` relation to a product.

3.1.3 Food supply chain example

We took the use case regarding the life cycle of broilers and their supply chain. In the diagram we can see the different organizations and their interaction in the process of making broiler meat for consumers.

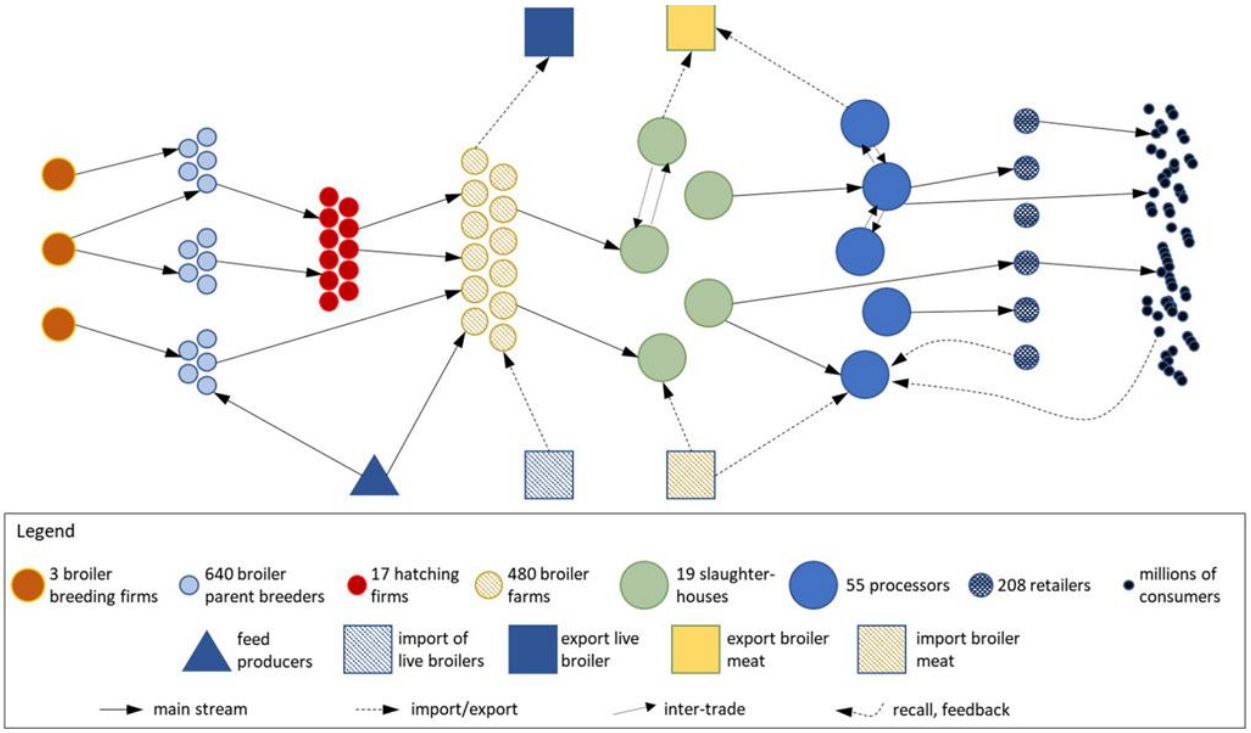

Figure 12 Food supply chain scenario

Now we are applying the GS1 EPCIS standard to the use case. We represent any movements of products and consumption of such in the case of feeding as an event.

- We have 6 Transportation events that reflect the mainstream from the presentation. These events are of type observation. We assume that with a single event we can capture the transition of the product from one organization to another. We could think of the event as the place in time which the product is received in the second organization. We omit and in fact infer the event that a product has been sent by the first organization.

- We add additional events for laying an egg, hatching an egg, slaughtering a broiler, and packaging the meat. These events are of type Transformation. We are assuming that one or more products are consumed to produce another product. In the case of laying an egg, the transformation gets as input the broiler and outputs the same broiler and the egg. For the other cases, the logic is similar.

We assume that every broiler, packaged food, broiler egg, broiler meat and broiler packaged meat is uniquely identified. The model currently is not complicated with the logic of handling aggregated packages of broilers and etc.

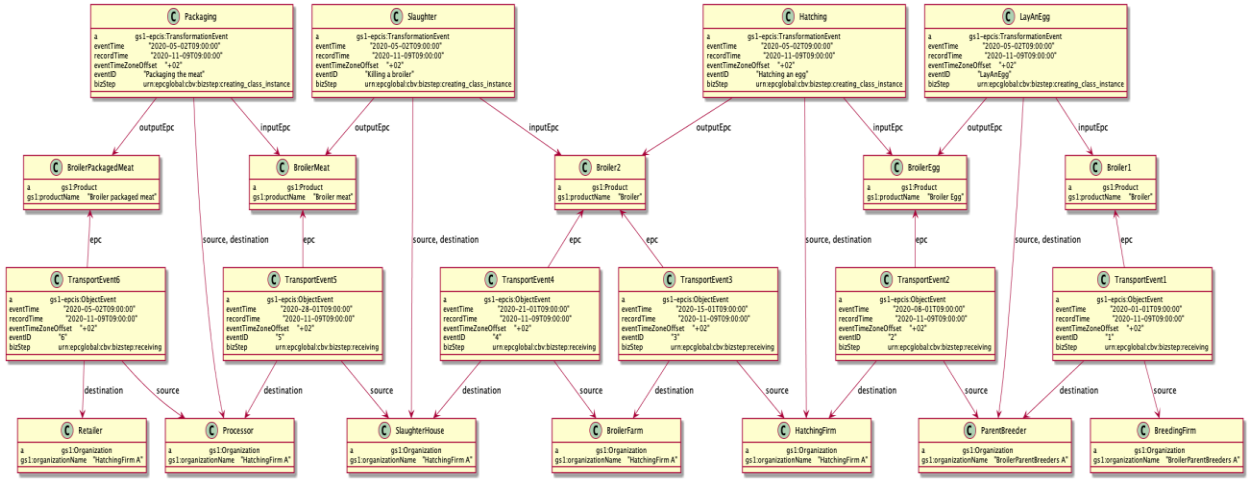

Figure 13 Scenario implementation with the FSM model

The above diagram shows the model of the main flow of the products. We will add additional diagrams displaying the secondary relations between the organizations.

The feeding of the broilers is important in order to keep track of how they have been raised and the quality of the meat. The feeding is relevant to Broiler Parent Breeders and Broiler Farms. The process of feeding is captured in two events:

- Buying of food - the point in time when a batch of food is being received in a farm or breeder organization. The food has its unique identifier signifying exactly which food package we refer to.
- The feeding itself - the point in time when a broiler is fed with a particular package of food.

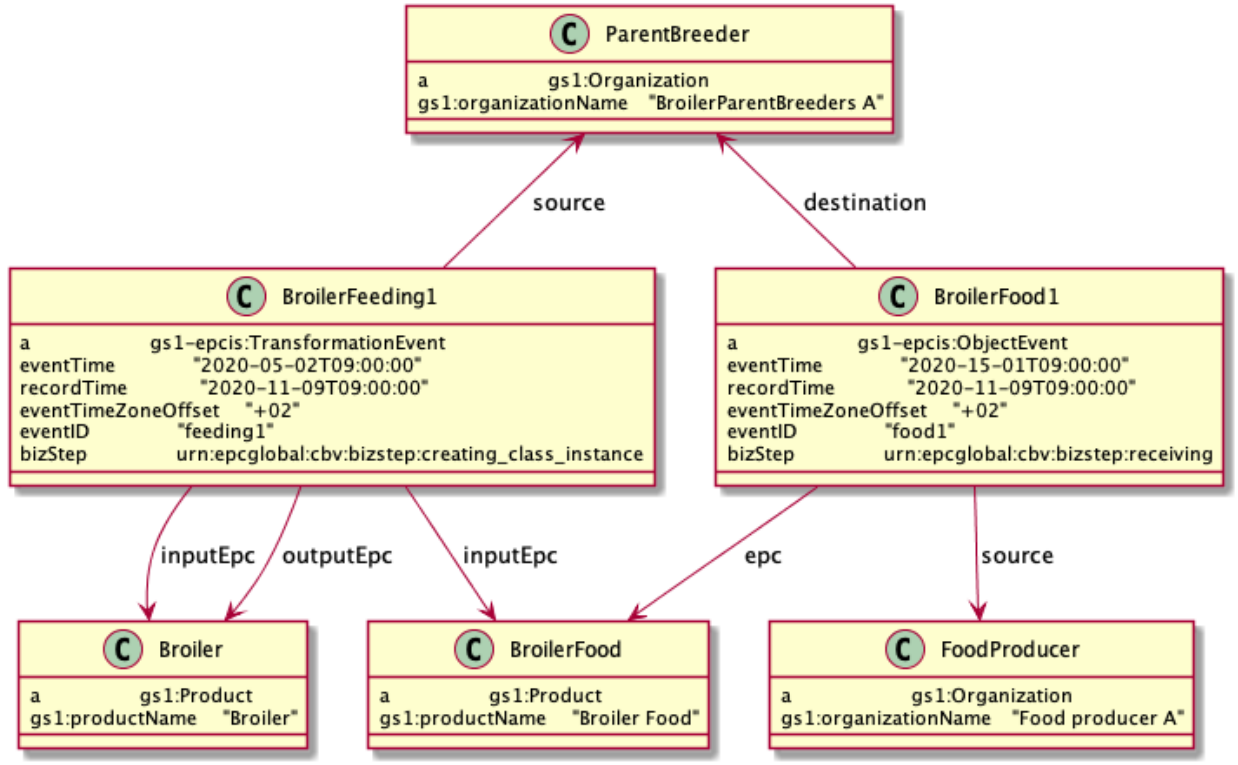

Figure 14 Scenario feeding implementation

The import and export of broilers are represented in the diagram below. An import or export is a single event. The event is the point in time when the product reaches the destination organization.

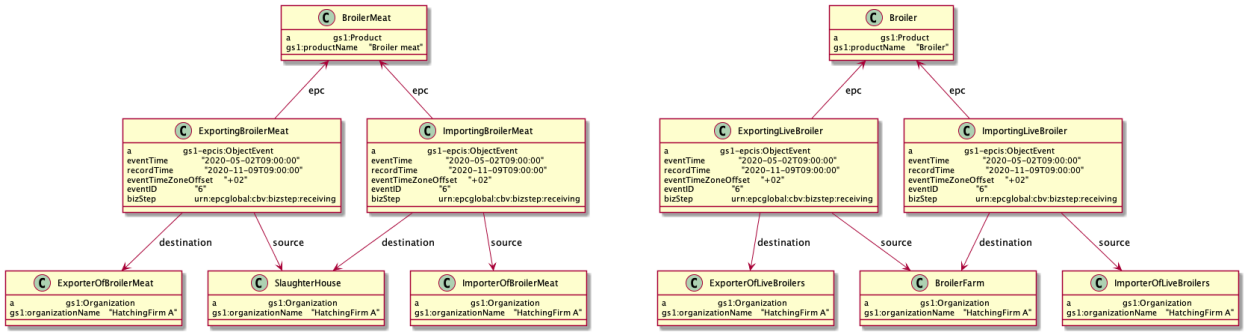

Figure 15 Scenario import/export implementation

3.2 Retailer Scenario

We use these requirements:

1. Browse information

2. list of the Certification Bodies and Certification Schema owners that the supplier (producer, processor) is working with.
3. able to perform a remote supplier verification using critical information like incidents, inspection results, certificates and lab tests
4. able to perform a supplier risk assessment in order to prioritise the audits and lab tests
5. Be able to access fully traced information
6. Be able to access information regarding findings of the inspection of suppliers in the food chain
7. Be able to access the current status of food supply actors, as far as audit results of certified organizations are concerned
8. Have access to innovative tools

How we cover the above requirements:

1. Based on the FSM model we can browse organizations, products, product batches. These organizations can be Retailers, Suppliers, Certification bodies, etc. We can query the database for Events regarding these objects. The events can be transactions between organizations, transformations of a product to a new product, certification grants, lab tests etc.
2. Certification bodies can be listed by querying for objects of type `fsm:CertificationBody`.
3. We can query directly inspection results, certificates and lab tests relating to an organization. This can be done using the event objects `fsm:Inspection`, `fsm:CertificateGrant`, `fsm:LabTest`, `fsm:Audit` respectively. The incidents are usually harder to trace back. Based on the presumption that we have the full supply chain data we can query the graph for all the events and parties involved with this product. The entry point for such an analysis should be the `fsm:Incident` object.
4. This functionality should be extracted from the raw data within the FSM model. It should be part of down-stream processing.
5. As per 3. we can fully trace a product batch through the supply chain as long as we have the full data for it.
6. This is part of 3. browsing inspection results should give you access to view the information within the system.

7. The current state should be logically aggregated based on historical event data. All the data will be within the system and will be queryable. A question like: How long results of Inspection/Audit are relevant after their release?
8. This should be part of another layer above the database.

3.3 Food Processor Scenario

Based on an overview of the responsibilities of 2 EU food authorities (Bulgarian Food Safety Agency, Netherlands Food and Consumer Product Safety Authority) we have extracted their obligations:

- Animal Health - Keep animals healthy. Animal welfare. Animal products processing secure
- Plant health - Plant diseases and prevent pests
- Food safety - Supervising the production, preparation, transport and sale of food products and sale of tobacco products.
- Product safety - Safe personal care products, indoor and outdoor games, consumer products in and around the house.
- Border inspections

The data we deem necessary for the FSM Platform are the:

- Reports of inspections - they could include information about safety risks.[31]
- News of new regulations[32] - In order to make this type of data queryable it would require to add the concepts of geo locations. As such regulations are issued for specific countries.

Currently we decided to use only the reports of inspections. For example, inspections are part of the Documents heading explained above.

[31] https://english.nvwa.nl/news/news/2017/08/03/nvwa-website-publishes-codes-of-fipronil-contaminated-eggs-from-investigated-farms
[32] https://english.nvwa.nl/news/news/2020/11/30/bird-flu-avian-influenza-now-as-well-in-belgium-and-poland-mandatory-additional-cleansing--disinfection-transport-vehicles-in-the-netherlands

3.4 Certification Body Scenario and Certification Body for Organic Wine Production Scenario

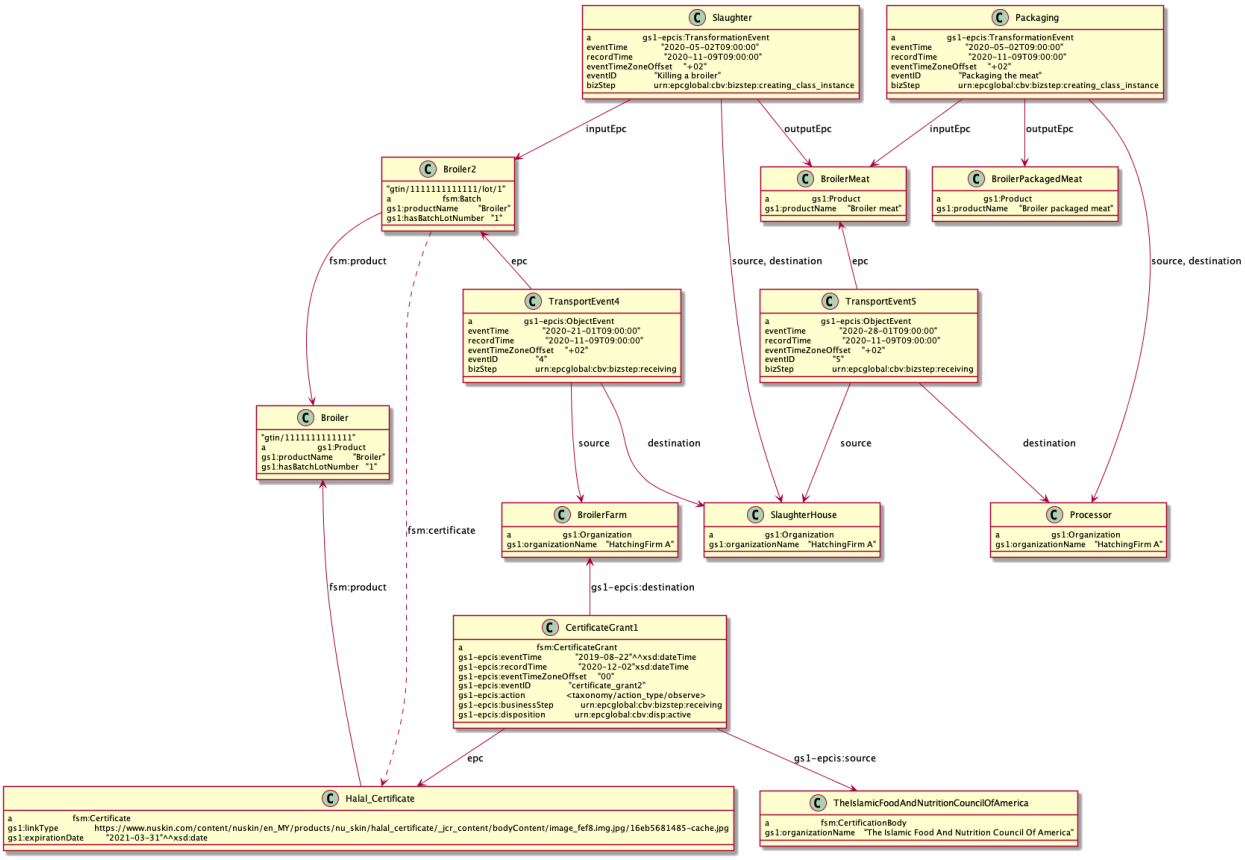

Figure 16 An example of the use cases involving certifications

The organization `BroilerFarm` has Halal certificate for broilers. The Broiler2 is a product batch for broilers. Based on the hatching of the broiler and the time of its slaughter we can deduce whether the certificate is relevant to this exact broiler batch. For the example we assume that the hatching of the broiler is after the Halal certificate grant. We can see that the transportation of the broiler to the slaughterhouse is prior to the expiration of the Halal certificate. Thus we infer that the broiler batch is under halal certification.

3.5 Food Safety Authority Scenario

The food safety authority can be identified using the `fsm:FoodSafetyAuthority`. Events can be associated with the authority. The currently supported events are of type `Incident`. They state that an incident has been reported for a certain product batch or a company product. These events fall in line with the Document scenarios explained in the FSM integrated model section Documents.

The information for incidents and the reporter is searchable and can be traced back through the supply chain to identify the companies involved in producing this batch of food.

4 CONCLUSION AND NEXT STEPS

In this document we presented TheFSM integrated data model and some implementations for the FSM use cases: food supply chain, retailer scenario, certification body scenario and food safety authority scenario. In order to ensure interoperability for the FSM platform, the data model is based on standards like GS1, EPCIS, W3C and domain-specific ontologies. The initial version of the data model (v1.00 on M12) was based on the analysis of the existing data and ontologies and some sample dataset and APIs specifications provided by TheFSM project consortium. For the next interations we consider improvements of the data model based on additional data sources, APIs and more domain-specific ontologies. According to the plans, v2.00 on M24 and v3.00 on M36 will be delivered.

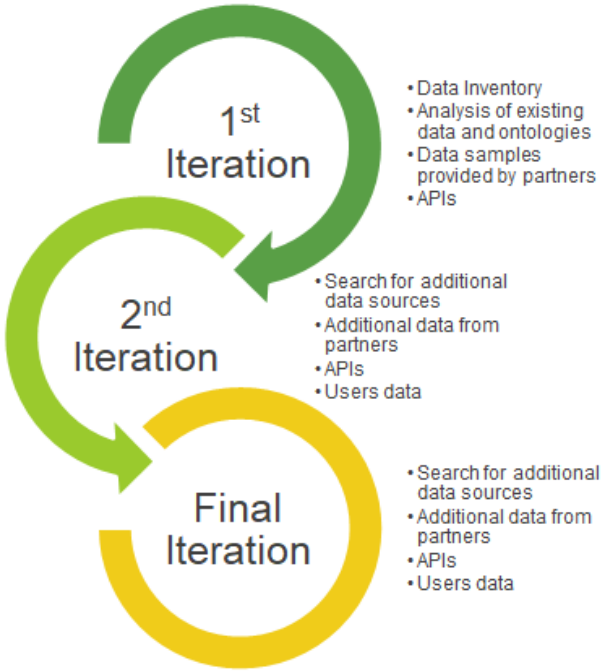

Figure 17 Next steps of data modelling