

# OntoPedigree: Modelling Pedigrees for traceability in supply chains

**Editor(s):** Name Surname, University, Country

**Solicited review(s):** Name Surname, University, Country

**Open review(s):** Name Surname, University, Country

Monika Solanki <sup>a,\*</sup>, Christopher Brewster <sup>a</sup>

<sup>a</sup> *Aston Business School, Aston University, Birmingham  
United Kingdom*

*E-mail: m.solanki, c.a..brewster@aston.ac.uk*

**Abstract.** The sharing of product and process information plays a central role in coordinating supply chains operations and is a key driver for their success. “Linked pedigrees” - linked datasets, that encapsulate event based traceability information of artifacts as they move along the supply chain, provide a scalable mechanism to record and facilitate the sharing of track and trace knowledge among supply chain partners. In this paper we present “OntoPedigree” a content ontology design pattern for the representation of linked pedigrees, that can be specialised and extended to define domain specific traceability ontologies. Events captured within the pedigrees are specified using EPCIS - a GS1 standard for the specification of traceability information within and across enterprises, while certification information is described using PROV - a vocabulary for modelling provenance of resources. We exemplify the utility of OntoPedigree in linked pedigrees generated for supply chains within the perishable goods and pharmaceuticals sectors.

**Keywords:** OntoPedigree, Linked Pedigrees, Supply chains, Design patterns, linked data, Ontologies

## 1. Introduction

One of the most important challenges in supply chains and logistics is information integration. Sharing of data and knowledge in a standardised manner along the supply chain is crucial not only to enable visibility, i.e., tracking and tracing of artifacts, but also to enable the more effective management of the supply chain.

Barcodes and more recently RFID tags have provided initial solutions to this challenge. GS1<sup>1</sup>, the standards organisation which manages barcodes, has provided various standards to facilitate end-to-end traceability and information sharing along the supply chain. The Electronic Product Code Information Service (EP-

CIS)<sup>2</sup> and the Core Business Vocabulary<sup>3</sup>, collectively provide specifications for the representation of product traceability information[1]. EPCIS provides a data model serialised as an XML schema, for capturing information artifacts that encapsulate the geographical progress and status of an item or set of items at each step of the supply chain. This is achieved by capturing data generated within the business context of scanning a barcode or RFID tag and encapsulating it within the abstraction of an “event”.

However a critical limitation of the current EPCIS specification is that though it does propose a mechanism to exchange and share data, the EPCIS XML schemas define only the structure of the data

---

\*Corresponding author. E-mail: m.solanki@aston.ac.uk

<sup>1</sup><http://www.gs1.org/>

<sup>2</sup><http://www.gs1.org/gsmv/kc/epcglobal/epcis>

<sup>3</sup><http://www.gs1.org/gsmv/kc/epcglobal/cbv>

to be recorded. The semantics of data and data curation processes are informally defined. Their interpretation is left up to the individual EPCIS specification implementing engines, thereby greatly increasing the possibility of interoperability issues arising between supporting applications, e.g., validation and discovery Web services built over the event repositories.

To overcome the limitations highlighted above, we propose the concept of event based “Linked pedigrees” [3] - interlinked datasets described in RDF, curated by consuming real time EPCIS events in the supply chain and encapsulated as linked data, that enable the capture of a variety of tracking and tracing information such as the *Chain of Custody* (CoC) and *Chain of Ownership* (CoO) about products as they move among the various trading partners. The notion of linked pedigrees is motivated by the widely prevalent use of pedigrees for tracking and tracing commodities in the pharmaceutical industry<sup>4</sup>. Pedigree or electronic pedigree (e-pedigree) is an audit trail that records the path and ownership of a drug as it moves through the supply chain, in which each stakeholder involved in the manufacture or distribution of the drug adds traceability information to the pedigree. As pedigrees are a collection of traceability information contributed by multiple parties, in order to achieve interoperability between pedigree handling systems, conformance to standards by all the parties involved is crucial. The modelling of pedigrees using GS1 specifications and representation using Semantic Web standards and linked data technologies provides a universally standardised mechanism to exchange traceability information.

Linked pedigrees specifically overcome a significant limitation prevalent in conventional pedigree exchange - that of information being available only from partners one-up or one-down in the supply chain. Dereferencing URIs make it possible to sequentially traverse the chain of pedigrees exchanged between partners and retrieve traceability information, given that adequate authentication, authorisation and access control mechanisms are in place.

In this paper we present “OntoPedigree”<sup>5</sup>, a content ontology design pattern for the data modelling of linked pedigrees, that can be specialised and extended to define domain specific or indeed product specific pedigree ontologies. We illustrate how EPCIS governing supply chain events can be exploited to generate

linked pedigrees using the OntoPedigree design pattern. We exemplify the utility of OntoPedigree driven linked pedigrees by providing motivating examples from the perishable goods and pharmaceuticals supply chain sectors.

The paper is structured as follows: Section 2 presents the requirements and competency questions we considered for the design of the pattern. Section 3 presents the intent, conceptual entities and the graphical illustration of OntoPedigree. It also highlights the relationship of OntoPedigree to other patterns and ontologies and provides details of implementation support for OntoPedigree. Section 4 presents the OWL axiomatisation of the pattern. Section 5 presents scenarios from the agri-food and pharmaceutical supply chain where OntoPedigree has been applied. Section 6 presents conclusions.

## 2. Requirements analysis

Pedigrees have been defined initially in EPCGlobal’s ratified Pedigree Standard<sup>6</sup> as follows: “A pedigree is a *certified* record that contains information about each distribution of a prescription drug. It records the sale of an item by a pharmaceutical manufacturer, any acquisitions and sales by wholesalers or repackagers, and final sale to a pharmacy or other entity administering or dispensing the drug. The pedigree contains *product information, transaction information, distributor information, recipient information, and signatures.*”

While the above description of a pedigree is given within the context of pharmaceutical supply chains, the interpretation of the definition highlights certain key requirements for the design of a generic content ontology design pattern for pedigrees, that could be reused across multiple sectors.

Recently the concept of “Event based Pedigree”<sup>7</sup> has been proposed that utilises EPCglobal’s EPCIS specification for capturing events in the supply chain and generating traceability datasets based on a relevant subset of the captured events. OntoPedigree, the design pattern for the generation of linked pedigrees as presented in this paper, builds on the event based pedigree approach.

<sup>6</sup>Version 1.0, 2007, <http://www.gs1.org/gsm/kc/epcglobal/pedigree>

<sup>7</sup>[http://www.gs1.org/docs/healthcare/Healthcare\\_Traceability\\_Pedigree\\_Background.pdf](http://www.gs1.org/docs/healthcare/Healthcare_Traceability_Pedigree_Background.pdf)

<sup>4</sup><http://www.axway.com/products-solutions/b2b/life-sciences-solutions/epedigree>

<sup>5</sup><http://purl.org/pedigree/#>

## 2.1. Requirements analysis

Below we outline some of the central requirements for pedigrees which form the basis of the entities defined in *OntoPedigree*.

- **Certification and digital signatures:** Pedigrees generated by supply chain stakeholders would need to be certified and digitally signed before they can be forwarded to other partners in the chain. Certification attributes include provenance information such as the organisational details associated with creating the pedigree and the Web service that generated it. Digital signatures would include the signature value and information needed to obtain the key associated with the signature.
- **Product information:** Master data about the products that are involved in the CoC/CoO must be interlinked with the core pedigree information.
- **Location information:** Location master data such as the business name and address details associated with the source and destination organisations, identified using for e.g., their GLN (Global Location Number) must be interlinked.
- **Consignment information:** Pedigrees must include detailed information about the consignments being shipped or received such as identities of the pallets, cases, bundles and items, timing attributes for the various business steps undertaken at the stakeholder’s facilities and the supply chain context associated with the shipment, i.e., pedigrees need to indicate their status in terms of the point (initial, intermediate or final) in the supply chain where they were generated.
- **Transaction information:** Transactional information such as identifiers of the purchase orders being fulfilled by the shipment are required to be incorporated in the pedigrees.
- **Partner pedigree information:** Apart from the pedigree initiated and created by the first partner in the supply chain, all other pedigrees must include information about the pedigrees for the stakeholders in the immediate upstream or downstream of the supply chain.

As highlighted in Section 1, a linked pedigree is a dataset, identified using an http URI, described and accessed using linked data principles and represented using the RDF data model. The definition of a pedigree must include URIs for certification, product, transaction and consignment information asserted in the trad-

ing partner’s knowledge base. It must include URIs to the pedigrees sent by the immediate upstream or downstream partners.

## 2.2. Competency questions

Given the requirements above, we define certain competency questions that govern the design of an ontology pattern for pedigrees.

- Who is the creator of the pedigree ?
- What is the supply chain creation status of a given pedigree?
- Which are the business transactions recorded against a particular consignment?
- What are the events associated with pedigrees created between dates X and Y?
- Which products have been shipped together?
- Which other pedigrees are included in the received pedigree?

## 3. Pattern description and Graphical representation

### 3.1. Intent

*OntoPedigree* provides a minimalistic abstraction and defines conceptual entities for the modelling of semantically enriched knowledge required to enhance data visibility in a supply chain. The pattern can be specialised to define domain specific pedigrees.

### 3.2. Conceptual Entities

Some of the key concepts and relationships encapsulating the data model defined by the pattern are described below. Note that several entities in *OntoPedigree* extend from or exploit concepts and relationships defined in PROV-O<sup>8</sup> and EEM(EPCIS Event Model)<sup>9</sup> as further illustrated in Figure 1.

- **Pedigree:** A concept defining the notion of a pedigree as an entity capturing certified information for a stakeholder in the supply chain.
- **PedigreeCreationStatus:** An entity defining the status of the creation a pedigree. Valid instances include `InitialPedigree`, `IntermediatePedigree` or `FinalPedigree`.

<sup>8</sup><http://www.w3.org/ns/prov-o>

<sup>9</sup><http://purl.org/eem#>

- PedigreeCreator: An entity linking the pedigree to its creating organisation.
- PedigreeCreationService: An entity representing the Web service used for the generation of the pedigree.
- pedigreeCreationTime: A datatype property linking the pedigree to the time of its creation.
- hasSerialNumber: An entity linking the pedigree to its uniquely assigned serial number.
- hasTransactionInfo: An entity defining the relationship between the pedigree and the conceptual entity capturing transaction information related to the sale, movement, return or transfer of goods.
- hasProductInfo: An entity defining the relationship between the pedigree and the conceptual entity capturing product information.
- hasConsignmentInfo: An entity defining the relationship between the pedigree and the conceptual entity capturing consignment information.
- hasReceivedPedigree: An entity defining the relationship between the pedigree and any receiving pedigrees from stakeholders in the upstream or downstream of one or more participating supply chains. Examples of such receiving pedigrees include,
  - \* A pedigree defined by a repacking trader, who combines goods from several consignments received from upstream stakeholders.
  - \* A pedigree defined by a downstream retailer who returns goods to an upstream trader.

### 3.3. Reusing related patterns and ontologies in OntoPedigree

Traceability information in OntoPedigree, regarding consignments and transactions are specified as sets of EPCIS events described using EEM [4] ontology. Certification information is defined using the PROV-O vocabulary. Extending from PROV-O serves the dual purpose of being able to attach provenance to pedigrees as well as facilitates the querying of pedigrees for certified knowledge using a dedicated vocabulary. Product master data is specified using domain specific vocabularies that in many scenarios exploit the GoodRelations ontology<sup>10</sup>.

<sup>10</sup><http://purl.org/goodrelations/v1>

Pedigrees generated using OntoPedigree are based on GS1 standards. Supply chain processes that implement these standards for the generation of linked pedigrees can declare their conformance and enforcement to these standards using the Standards Enforcer content ontology design pattern Pattern(SEP)<sup>11</sup>. and link it with the pedigrees. The added value of this interlinking is that besides the certification information included in the pedigree, it serves as an additional validation when the pedigree is received and dereferenced by other partners.

### 3.4. Graphical representation

Figure 1 illustrates the graphical representation of OntoPedigree. It depicts the entities defined for the pattern and their relationships with entities from PROV-O, EEM and GoodRelations.

### 3.5. Implementation support for OntoPedigree

In order to facilitate the creation of linked pedigrees that exploits EEM and OntoPedigree, we have implemented the LinkedEPCIS reference implementation and Java library<sup>12</sup>. An important component of the library is the “LinkedPedigreeGenerator” which automates the generation of linked pedigrees by querying the triple store for event URIs. Additionally, URIs for the product master data, made available as linked data are retrieved. Consignment information is populated with spatial data for locations from DBpedia and Geonames. The pedigree status is dynamically set, based on the point at which the pedigree is generated. LinkedEPCIS thus provides the implementation support required for the uptake of OntoPedigree by supply chain stakeholders, to be exploited in their own IT infrastructure.

## 4. OntoPedigree Axiomatisation

A pedigree is required to have certain mandatory properties and relationships as part of its definition. In this section we present an incremental axiomatisation of OntoPedigree(ped) using the OWL Manchester syntax. As noted earlier, several entities in OntoPedigree extend or exploit the definitions in PROV(prov) and EEM(eem).

<sup>11</sup><http://windermere.aston.ac.uk/~monika/ontologies/Standards-Enforcer-Pattern.owl>

<sup>12</sup><http://github.com/nimonika/LinkedEpcis/>

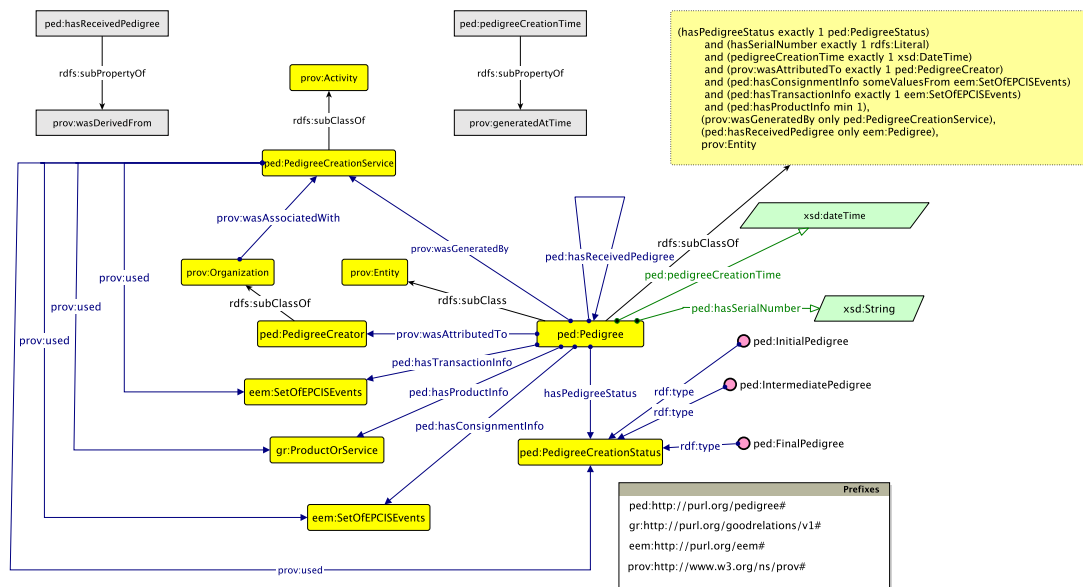


Fig. 1. Graphical Representation of OntoPedigree

- A pedigree is a PROV entity

```
Class: ped:Pedigree
SubClassOf:
  prov:Entity
```

- A pedigree must include a creation time, a serial number and describe its status.

```
Class: ped:Pedigree
SubClassOf:
  (hasPedigreeStatus exactly 1
   ped:PedigreeStatus)
  and (hasSerialNumber exactly 1
   rdfs:Literal)
  and (pedigreeCreationTime exactly 1
   xsd:DateTime)
```

- It must assert the creating organisation and may include the service used for its creation.

```
Class: ped:Pedigree
SubClassOf:
  (prov:wasAttributedTo exactly 1
   ped:PedigreeCreator),
  (prov:wasGeneratedBy only
   ped:PedigreeCreationService)
```

- It must include atleast one consignment relation, a transaction relation and atleast one relationship capturing product master data.

```
Class: ped:Pedigree
SubClassOf:
  (ped:hasConsignmentInfo someValues
   eem:SetOfEPCISEvents)
  and (ped:hasTransactionInfo exactly 1
   eem:SetOfEPCISEvents)
  and (ped:hasProductInfo min 1)
```

- Finally, it may include links to pedigrees received from the immediate upstream or downstream partners.

```
Class: ped:Pedigree
```

```
SubClassOf:
  (ped:hasReceivedPedigree only eem:Pedigree)
```

Combining the above assertions, the overall definition of Pedigree is illustrated in Figure 2.

### 5. Exemplifying OntoPedigree

In this section we present scenarios from the perishable goods and pharmaceutical supply chains that exemplify the use of OntoPedigree in the exchange of traceability information. Perishable goods is one of the scenarios considered in the FI-PPP projects SmartAgrifood<sup>13</sup> and FISpace<sup>14</sup>.

#### 5.1. Perishable Goods supply chains

The lifecycle of perishable goods e.g., tomatoes in the agri-food sector, is a complex process until they reach the end consumer because of the number of involved stakeholders and the diverse set of data that is produced. The tomato supply chain involves thousands of farmers, hundreds of traders and few retail groups. Figure 3 shows a generalised food chain scenario with a reduced level of complexity. This scenario covers 90% of the supply scenarios for fresh food products. The general workflow involving the captur-

<sup>13</sup><http://www.smartagrifood.eu/>

<sup>14</sup><http://www.fispace.eu/>

```

Class: ped:Pedigree
SubClassOf:
  (hasPedigreeStatus exactly 1 ped:PedigreeStatus)
  and (hasSerialNumber exactly 1 rdfs:Literal)
  and (pedigreeCreationTime exactly 1 xsd:DateTime)
  and (prov:wasAttributedTo exactly 1 ped:PedigreeCreator)
  and (ped:hasConsignmentInfo someValuesFrom eem:SetOfEPCISEvents)
  and (ped:hasTransactionInfo exactly 1 eem:SetOfEPCISEvents)
  and (ped:hasProductInfo min 1),
  (prov:wasGeneratedBy only ped:PedigreeCreationService),
  (ped:hasReceivedPedigree only eem:Pedigree),
  prov:Entity

```

Fig. 2. Manchester syntax serialisation of OntoPedigree

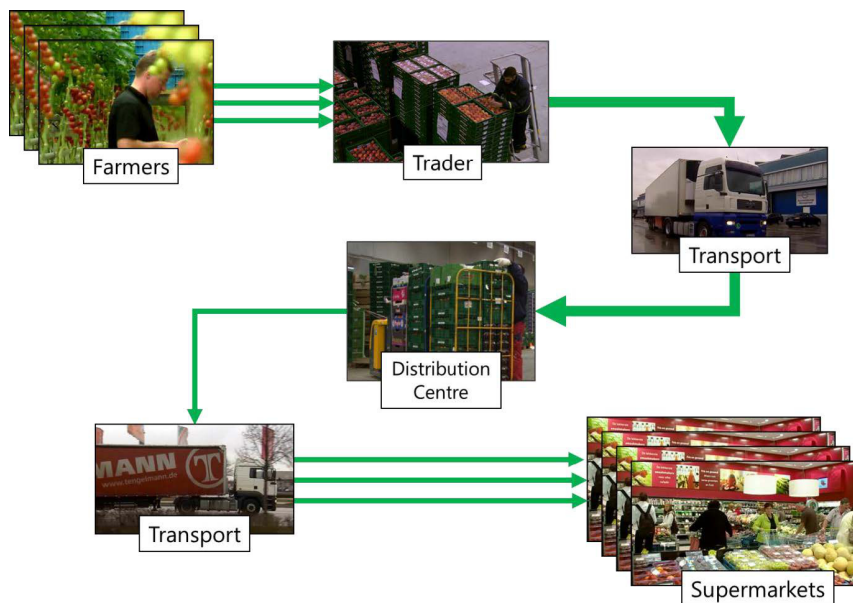


Fig. 3. Generalised agri-food chain scenario for tomatoes

ing of events, generation of linked pedigrees and exchange of pedigrees related to the sale of tomatoes between stakeholders such as the farmer (Franz), the trader (Joe), the distribution centre (FreshFoods Inc) and the supermarket (Orchard) is outlined below. We assume that all supply chain trading partners have a supporting EPCIS implementation infrastructure installed that can be exploited for capturing and querying event and pedigree datasets.

- Franz farmer specialises in growing tomatoes. The packaging of tomatoes is done in punnets, each of which are tagged with RFID labels. Shipment of tomatoes to downstream partners is done

in cardboard boxes each of which is tagged with a RFID tag.

- Joe trader bundles tomatoes procured from multiple farmers to larger product batches before dispatching them to distribution centres.
- Freshfoods Inc. sources tomatoes from multiple traders and splits up large product batches into smaller batches for distribution to retail supermarkets.
- Orchards is a supermarket that receives fresh produce from distribution centres such as Freshfoods Inc.

Joe trader requests pedigree information on an identified tomato batch that has been delivered to him by Franz farmer. The request is made by RESTfully invoking Franz farmer's agent which is part of the FMS (farm management system) installed at Franz farmer's end[2]. Joe trader receives an authenticated and certified message containing the pedigree URI, where the pedigree has been generated using the OntoPedigree design pattern. Joe trader's agent dereferences the URI and receives the pedigree dataset. Object property value resources in the pedigrees are asserted using EPCIS event data URIs.

Joe trader combines the tomato produce received from Franz with those received from other farmers (e.g., Bob) into shipments which are then forwarded to Freshfood Inc. On receiving a pedigree request from Freshfood Inc, Joe trader's agent sends the pedigree which includes URIs to the pedigree provided by Franz farmer and Bob farmer.

The pedigrees generated and exchanged in the workflow defined above are illustrated in Figure 4.

The significant advantage of exchanging traceability information using linked pedigrees over conventional mechanisms is that the pedigree received by FreshFood Inc. from Joe trader includes URIs to pedigree datasets provided by Franz farmer and Bob farmer, even though they are not FreshFood Inc's one-up or one-down partners. Consuming EPCIS event data curated as linked data to generate and exchange linked pedigrees as outlined above can help derive implicit knowledge that can expose inefficiencies such as shipment delay, inventory shrinkage and out-of-stock situations.

### 5.2. *Pharmaceutical Supply chains*

We outline the scenario of a pharmaceutical supply chain, where trading partners exchange product track and trace data using linked pedigrees. Figure 5 illustrates the flow of data for four of the key partners in the chain.

The *Manufacturer* commissions<sup>15</sup>, i.e, assigns an EPC(Electronic Product Code) to the items, cases and pallets. The items are packed in cases, cases are loaded onto pallets and pallets are shipped. At the *Warehouse*, the pallets are received and the cases are unloaded. The cases are then shipped to the various *Distribution centers*. From the Distribution centers the cases are sent

to retail *Dispenser* outlets, where they are received and unpacked. Finally, the items are stacked on shelves for dispensing, thereby reaching their end-of-life in the product lifecycle.

As the serialised items, cases and pallets move through the various phases of the supply chain at a trading partner's premises, EPCIS events are generated and recorded at several RFID reader nodes.

When the pallets with the cases are shipped from the manufacturer's premises to the warehouse, pedigrees encapsulating the minimum set of EPCIS events are published at an IRI based on a predefined IRI scheme. At the warehouse, when the shipment is received, the IRI of the pedigree is dereferenced to retrieve the manufacturer's pedigree. When the warehouse ships the cases to the distribution center, it incorporates the IRI of the manufacturer's pedigree in its own pedigree definition. As the product moves, pedigrees are generated with receiving pedigrees being dereferenced and incorporated, till the product reaches its end-of-life stage.

## 6. Conclusions

Data visibility in supply chains has received considerable attention in recent years. Information systems are now being designed to facilitate the process of making data available in real time to stakeholders in the supply chain.

In this paper, we have proposed a design pattern "OntoPedigree" that provides a minimalistic abstraction for designing domain specific pedigree ontologies. OntoPedigree builds upon EPCIS, a GS1 standard for facilitating event based traceability and PROV-O, a vocabulary for specifying provenance of linked datasets. OntoPedigree can be mapped with other content ontology patterns such as SEP to declare the conformance of supply chain processes to these standards when they generate pedigrees using OntoPedigree.

The pattern design is supported by a robust set of requirements and a rigorous set of competency questions. An OWL axiomatisation of OntoPedigree has been provided. Finally, an implementation framework that facilitates the seamless uptake of the generation of linked pedigrees in supply chains has been presented. It is worth noting that OntoPedigree is domain independent and can be widely applied to most scenarios of traceability. We strongly believe that linked pedigrees generated using OntoPedigree can make a significant difference to current visibility approaches in supply chains. We are in the process of applying the pat-

---

<sup>15</sup>associating the serial number with the physical product

```

fsc:FranzTomatoFarmerPedigree123 rdf:type ped:Pedigree;
prov:wasAttributedTo <http://www.joetomatotrader/foaf>;
ped:pedigreeCreationTime "2014-06-05T07:53:14.189+01:00"^^xsd:dateTime;
ped:hasSerialNumber "tomPed123"^^xsd:String;
ped:hasStatus ped:Initial;
ped:hasConsignmentInfo fci:FranzFarmerObjectEvent10,
                        fci:FranzFarmerAggregationEvent6;
ped:hasTransactionInfo fti:FranzFarmerShippingEvent12;
ped:hasProductInfo ftp:FranzTomatoesMay2013Data.

jsc:JoeTomatoTraderPedigree456 rdf:type ped:Pedigree;
prov:wasAttributedTo <http://www.joetomatotrader/foaf>;
ped:pedigreeCreationTime "2014-06-08T03:67:09.175+01:00"^^xsd:dateTime;
ped:hasSerialNumber "joeTradePed456"^^xsd:String;
ped:hasStatus ped:Intermediate;
ped:hasConsignmentInfo jci:JoeTraderObjectEvent20,
                        jci:JoeTraderObjectEvent30;
ped:hasTransactionInfo jti:JoeTraderTransactionEvent40;
ped:hasProductInfo jpi:JoeTradesMay2013Info.
ped:hasReceivedPedigree fsc:FranzTomatoFarmerPedigree123,
                        bsc:BobTomatoFarmerPedigree123.

```

Fig. 4. Linked pedigrees created using OntoPedigree and exchanged between the supply chain partners

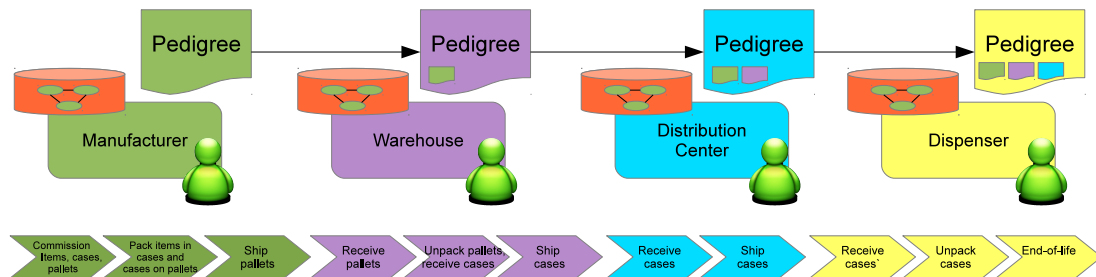


Fig. 5. Trading partners in a pharmaceutical supply chain and the flow of information

tern to the generation of pedigrees in the Wine supply chain where traceability information is required to be integrated with datasets for the environmental conditions prevailing in the vineyards.

## References

- [1] Kary Främling, Sagar Parmar, Ville Hinkka, Jaakko Tätilä, and Dirk Rodgers. Assessment of epcis standard for interoperable tracking in the supply chain. In Theodor Borangiu, Andre Thomas, and Damien Trentesaux, editors, *Service Orientation in Holonic and Multi Agent Manufacturing and Robotics*, volume 472 of *Studies in Computational Intelligence*, pages 119–134. Springer Berlin Heidelberg, 2013. ISBN 978-3-642-35851-7.
- [2] Alexandros Kaloxylas, Robert Eigenmann, Frederick Teye, Zoi Politopoulou, Sjaak Wolfert, Claudia Shrank, Markus Dillinger, Ioanna Lampropoulou, Eleni Antoniou, Liisa Pesonen, Huether Nicole, Floerchinger Thomas, Nancy Alonistioti, and George Kormentzas. Farm management systems and the future internet era. *Computers and Electronics in Agriculture*, 89(0):130 – 144, 2012. ISSN 0168-1699. .
- [3] Monika Solanki and Christopher Brewster. Consuming Linked data in Supply Chains: Enabling data visibility via Linked Pedigrees. In *Fourth International Workshop on Consuming Linked Data (COLD2013) at ISWC*, volume Vol-1034. CEUR-WS.org proceedings, 2013.
- [4] Monika Solanki and Christopher Brewster. Representing Supply Chain Events on the Web of Data. In *Workshop on Detection, Representation, and Exploitation of Events in the Semantic Web (DeRiVE) at ISWC*. CEUR-WS.org proceedings, 2013.