Enabling the business-based
Internet of Things and Services

(FP7 257852)

D3.2 Vertical and horizontal business vocabularies

Published by the ebbits Consortium

Dissemination Level: Public

Project co-funded by the European Commission within the 7th Framework Programme
Objective ICT-2009.1.3: Internet of Things and Enterprise environments

Document version: 1.0 Submission date: 31.08.2011

Document control page

Document file: D3.2 Vertical and horizontal business vocabularies
Document version: 1.0
Document owner: Martin Knechtel (SAP)

Work package: WP3 Enterprise Frameworks for Life-cycle Management
Task: Task T3.2 –Semantic business decision models
Deliverable type: R = Report

Document status: ☒ approved by the document owner for internal review
☒ approved for submission to the EC

Document history:

Version	Author(s)	Date	Summary of changes made
0.0	Tomáš Sabol (TUK)	2011-03-09	Initial ToC
0.4	All Contributors • TUK: Tomáš Sabol, Jozef Glova • SAP: Premchand Nutakki, Martin Knechtel • COMAU: Pietro Cultrona • TNM: Thomas Nejsum Madsen, Michael Jacobsen, Sigurjón Björnsson	2011-06-30	First draft version of all contributions, mainly (in the order of the authors) • 3.1 • 3.2, 4.1, • 5, 4.3 • 6, 4.2
0.5	All Contributors (see above)	2011-08-03	Contributions ready
0.6	Martin Knechtel (SAP)	2011-08-04	Document sent to reviewers
0.9	Reviewers	2011-08-11	Internal review results
1.0	Jozef Glova (TUK)	2011-08-31	Final version submitted to the European Commission

Internal review history:

Reviewed by	Date	Summary of comments
Helene Udsen (In-JeT)	2011-08-11	Accepted with comments.
Riccardo Tomasi, Claudio Pastrone (ISMB)	2011-08-05	Accepted with comments.

Index:

1. Executive Summary

The ebbits project aims at providing semantic interoperability of services, things and people. The basis for this interoperability is a common vocabulary used in communication both between systems and between humans but also between systems and humans. This deliverable provides the basis for this vocabulary by discussing vertical and horizontal business vocabularies.

Chapter 2 introduces the Deliverable and brings it into context with the ebbits project.

After a general introduction with definitions of the terms “controlled vocabulary”, “thesaurus”, “taxonomy”, and “ontology” Chapter 3 discusses vocabularies and taxonomies available in literature from research projects, research groups, but also organizations and enterprises.

Chapters 5 and 6 provide the identification, discussion, evaluation, selection and summary of relevant terms for the ebbits vocabularies for the Automotive Manufacturing and the Food Traceability domains, respectively.

Chapter 4 starts with an introduction of standardization organizations and the standards relevant to ebbits. Furthermore, Chapter 4 provides a selection of terms that are not domain-specific and thus should be considered in defining a vocabulary that is usable across both domains. For instance, terms related to ordering and transporting of goods (e.g. price, quantity, weight, destination, location) are included. This Chapter is a result of discussions of both end-user partners based on the results of Chapter 5 and 6. Although created later, it has been put in front since this order is more logical from the reader’s point of view.

Overall the goal of the deliverable is to provide a justified selection of terms from the two ebbits application domains. These terms form an initial set of suggested terms for the ebbits vocabulary. This deliverable is the basis for capturing the ebbits vocabulary formally, e.g. in an OWL or RDF ontology. This deliverable will support the creation of, e.g., the Deliverables “D4.7.1 Use-case driven semantic models (M24)” and “D4.7.2 Use-case driven semantic models (M36)”.

2. Introduction

2.1 Purpose, context and scope of this deliverable

The purpose of this Deliverable is to identify vocabularies that will be used in the ebbits project. The identification, discussion and evaluation, and distillation of relevant vocabularies are considered within the scope of this deliverable. On the other hand, the transformation of this vocabulary into a formal format, e.g. an OWL ontology, is considered out of scope, as this is treated in other deliverables.

2.2 Background

This Deliverable is one of the outcomes of Task T3.2 Semantic business decision models in Work Package 3.

Selected WP3 objectives that are relevant for this Deliverable are

- Creation of vocabularies and description of business logic and business processes across various application domains
- Creation of an open ontology framework to address the issue of semantic interoperability

A main objective (and the one relevant for this Deliverable) of T3.2 is "to develop a common business vocabulary (vertical & horizontal) for business concepts and processes to be supported by the ebbits platform." The ebbits DoW further elaborates:

> *Implementation of semantic interoperability requires a well defined, generic business vocabulary and a semantic ontology framework which can be used in a wide range of application domains. The business vocabulary will allow us to describe business logic and business processes across various application domains.*
>
> *The vertical domain-specific vocabularies are specific to a business application domain. Examples of vertical artefacts could be production setup instructions, workflow scheduling, maintenance procedures, etc.*
>
> *At the same time, a horizontal vocabulary will be developed for various types of business concepts and data types that are common across multiple domains to be used for dynamic interoperability between different business partners. The artefacts in the cross-domain vocabulary consist of common business concepts, data and documents like purchase orders, shipping notices, joint process and workflow planning, bilateral contracts, trust, roles, authentication, etc.*
>
> *Dynamic SOA based business processes operate on the "publish-find-bind" principle, where business processes may dynamically involve business partners and associated applications. The ebbits platform provides semantic interoperability to handle such dynamic situations involving service brokers to bind enterprises (including economic relationships) that have no prior business relationships between them.*

3. Vocabularies and Taxonomies

3.1 General characterization of vocabularies and taxonomies

Modelling a domain is a difficult intellectual exercise that requires in-depth domain knowledge. In general a domain can be modelled in many ways. In the following chapter we discuss three of the four prevalent domains modelling schemes, namely controlled vocabularies, thesauri, and taxonomy; the fourth scheme i.e. ontology will be processed in other deliverables.

Controlled Vocabularies

A controlled vocabulary is a restricted list of words or terms used for labelling, indexing or categorizing. It is defined as controlled because only terms from the list may be used for the subject area covered by the controlled vocabulary. The definition also implies that, if it is used by more than one person, there is control over who adds terms to the list, when modification are authorized, and how the list is updated. As a result, the list can grow over time, but only under defined policies.

Controlled vocabularies do not necessarily have any structure or relationships between terms within the list. Controlled vocabularies are often used for name authorities (proper nouns), such as persons, organization names, company names, etc.

Online controlled vocabularies often have synonyms or cross-references directing users or search engines from an incorrect (unfavoured) variant to the equivalent preferred term in the controlled vocabulary. However, this is not a required attribute of a controlled vocabulary.

Controlled vocabularies are the broadest category, which also includes thesauri and taxonomies. In other words, thesauri and taxonomies are specific kinds of controlled vocabularies, but not all controlled vocabularies are thesauri or taxonomies.

Business vocabulary and semantics

A business vocabulary contains all the specialized terms like rules, fact types and concepts expressed that a community uses in their talking and writing in the course of doing business. '*Semantics*' is "the meaning or relationship of meanings of a sign or set of signs" (MWCD, 2011). The signs can be of any form: words, phrases, codes, numbers, icons, sounds, etc. The *business semantic* presents the contextual meaning of key business assets for the firm by adding business facts and rules to define the business context.

Semantics of Business Vocabulary and business Rules (SBVR)

SBVR is the result of the integration of ISO standards, formal logics, linguistics, and practical experience from foremost practitioners in the field of business vocabulary for business rules. SBVR also includes two specialized vocabularies of business semantics. On the one hand, the SBVR "Vocabulary for Describing Business Vocabularies" deals with all kinds of terms and meanings (other than meanings of Business Rules). On the other hand, the SBVR "Vocabulary for Describing Business Rules" deals with the specification of the meaning of business rules, and builds on the "Vocabulary for Describing Business Vocabularies." The two have been separated so that the "Vocabulary for Describing Business Vocabularies" could be used independently - for example, as a basis for vocabularies for business processes or organizational roles. The next two sub-clauses deal with the semantics of business vocabularies and the semantics of business rules.

Thesaurus

Thesauri (or Thesaurus)

More structured kinds of controlled vocabularies are arranged in a known order and built so that the various relationships among terms are clearly shown and identified by standardized relationship indicators. Among these vocabularies, an even more structured approach is followed by thesauri. A Thesaurus, in fact, defines each term only by three types of relationships; *hierarchical*, *associative* and *equivalent*.

Thesauri provide information about each term and its relationships to other terms within the same thesaurus. In addition to clearly specifying which terms can be used as synonyms (called "used from"), a thesaurus also indicates which terms are more specific (narrower terms), which are broader, and which are related terms. National and international standards have been developed to provide guidance on creating thesauri, including standards like ISO 2788, ISO 5964, ANSI/NISO Z39.19. The standards explain in great detail the types of relationships that fall into the three types mentioned above: hierarchical (Broader Term/Narrower Term), associative (Related Term), and equivalence (Use/Used from).

A literature retrieval thesaurus, like a dictionary-thesaurus lists similar terms at each controlled vocabulary term entry. The difference is that in a dictionary-thesaurus all the associated terms might be used in place of the term entry depending upon the specific context, which the user needs to consider in each case. Nevertheless in certain contexts some of these terms are not appropriate. The literature retrieval thesaurus, on the other hand, is designed to be used for all contexts, regardless of a specific term usage or document. The synonyms or near synonyms must therefore be suitably equivalent in all circumstances.

The following standards from International Organization for Standardization (ISO)[1] provide guidance for creating thesauri:

- ISO 2788 (1986): Guidelines for the Establishment and Development of Monolingual Thesauri;

- ISO 5964 (1985): Guidelines for the Establishment and Development of Multilingual Thesauri;

- ISO 2788 and 5964 are to be replaced in 2011 by ISO 25964: Thesauri and Interoperability With Other Vocabularies.

Taxonomy

Taxonomy

Originally, taxonomy (from Ancient Greek: *taxis* "arrangement" and *nomia* "method") was the practice and science of classification. In general, a taxonomy is a system for naming and organizing objects into categories (e.g. topic hierarchies, table of contents, advanced search forms) that share similar characteristics.

The word taxonomy means a collection of controlled vocabulary terms organized into hierarchical structure to make them easier to identify, study, or locate. A taxonomy models concepts within a domain from the more abstract to the more specific. Each term in taxonomy is in one or more parent/child (broader/narrower) relationships to other terms in the taxonomy, but not necessarily the related-term relationships and other requirements of a standard thesaurus. Unlike a thesaurus, where a given term may or may not have broader or narrower terms, in a taxonomy all terms belong to a single, large hierarchy that encompasses all concepts of a certain class, category, or aspect. The structure is sometimes referred to as a "tree" and the terms as "nodes" in the tree. A node may be repeated at more than one place within the taxonomy if it has multiple broader terms. This is referred to as a polyhierarchy.

The term taxonomy tends to be used to refer to two different things (ASI, 2011):

- A tree-hierarchical controlled vocabulary lacking more complex relationships found in thesauri or ontologies, or
- Any kind of controlled vocabulary, especially when applied to the world of enterprise content management and website information architecture, rather than library science literature retrieval.

[1] www.iso.org/iso/iso_catalogue.htm

According to (MIR, 2001), taxonomies generally consist of two parts, i.e., *structures* and *applications*. More specifically, the structures consist of the categories (or terms) themselves and the relationships that link the different categories together. Applications refer to the navigation tools available to help users in finding the required information.

```
,Industries
    Health care
    Financial services
    Manufacturing
    Wholesale and retail trade
    Transportation (see also “Employment [transportation]”)
        Air
        Ground transport
        Maritime
        Space
(Ocean transportation see “Maritime transportation”)
```

Figure 1: Example of taxonomy structure.
Source: own, according to Montague Institute Review's Managing taxonomies strategically; available on http://www.montague.com/abstracts/taxonomy3.html.

Taxonomy structures typically consist of the following elements:

- *List of standard terms* (for instance in the example above we have used “maritime” instead of “ocean” as the standard term);
- *Hierarchical relationships* (for instance “Manufacturing” is subordinate to “Industries”).
- *Cross references*.

Taxonomy structures can also make automated processes more efficient. Taxonomy terms can in fact be used in a search engine query to help users find information more easily, and in a filtering program to personalized e-mail alerts or websites.

Taxonomy structures and applications working together can help perform the following kinds of tasks:

- Identification
- Discovery
- Delivery

Although it is possible for one taxonomy structure to serve multiple applications, in practice most taxonomy are *eccentric* (suitable for only one environment and application). Most business taxonomies are highly customized, and it is not unusual for a single company to use multiple taxonomies for different functions or applications (e.g. one for marketing, another for product development).

In business, the taxonomy must respond to rapid changes in three areas:

- *Business processes* – geographic taxonomies often conform to sales territories. Product taxonomies originate in manufacturing processes.
- *Budgeting and managing* – organization categories reflect deployment of human and physical resources..
- *Strategic planning.*

The business taxonomies are often designed for a single task or process. Different organizations require different taxonomies because each organization has unique processes, organizational configuration, core competencies, and histories.

General instructions about creating a business taxonomy can be found in (MIR, 2001):

1. A taxonomy usually starts with *a list of standard terms, a "vocabulary"*, that describes the content to be organized.
2. *Adding relationships among the terms*, where relationships include cross-references from nonstandard terms (e.g. ASI) to standard terms (e.g. the American Society for Indexing), from narrower terms to broader terms (e.g. "manufacturing" see also "industry"), and from one term to a related term (e.g. "indexing" see also "taxonomy"). Many people also incorporate definitions and notes (e.g. "data mining" refers to a computer data extraction technique, not a method of extracting ore from the earth). This part of taxonomy is often called the "thesaurus". It contains synonyms, but it also does a lot more.
3. The vocabulary and thesaurus constitute the taxonomy structure. The next step is to *connect the terms with sources* (websites, documents, people, etc.). Typically this happens in the taxonomy application, along with sorting and formatting the terms (for instance the Montague Institute Index is formatted in five ways – alphabetical order, chronological order, subjects, organizations, and people).

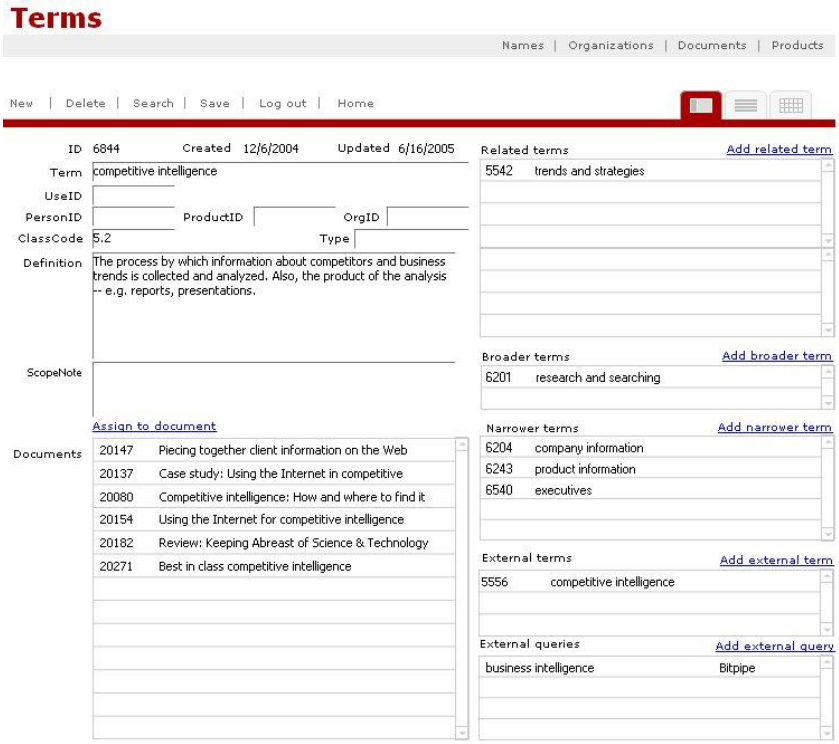
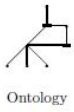


Figure 2: Example of Montague Institute teaching Lab showing the terms (vocabulary) section. Source: available on http://www.montague.com/images/terms.jpg.

Recently the term taxonomy has also become popular as the term for any kind of controlled vocabulary, whether a structured thesaurus, a simple synonym ring, or anything in-between. This is especially the case in the corporate world, where one might speak of "enterprise taxonomies." Thus, these taxonomies may or may not have the hierarchical structure that is associated with traditional taxonomies. It is simpler to have a one-word term for the concept of controlled vocabularies, especially when speaking of the people involved, such as "taxonomists" instead of "controlled vocabulary creators/editors."



Ontologies

Ontology is the most formal model. It defines the meaning of concepts with attributes and relationships between the various concepts that contain various meanings, all to define a domain of

knowledge expressed in a format that is machine-readable. Certain applications of ontologies, as used in artificial intelligence or biomedical informatics, may define a domain of knowledge through terms and relationships as the end goal. In the area of taxonomies and information science, however, an ontology can be seen as a more complex type of thesaurus, in which instead of having simply "related term" relationships, there are various customized relationship pairs that contain specific meaning, such as "owns" and a reciprocal "is owned by." For more details see (Schwarz, 2005) or other sources.

3.2 State-of-the-art analysis

There are many enterprises, organizations and researchers working to develop methods and tools for automated vocabulary generation and taxonomy creation. Additionally, the existence of communities and groups which provide help to develop and manage taxonomies has made the process of building a new taxonomy simple. Intel developed a cloud computing taxonomy to use within Intel (Intel 2010). It is based on several externally developed taxonomies as well as Intel IT capability frameworks that describe its primary applications and infrastructure services. Recently released IDC's security products taxonomy provides a standardized and detailed framework of security marketplace (Christiansen et al., 2011).

Products and Services Ontology (Hepp 2006) provides core classes and properties derived from the products and services categorization standard *eCl@ss 5.1.4 en*[2]. Below, we provide a state-of-the-art analysis of the work being done in this area. We focus mainly on vocabularies and taxonomies related to business products and services.

3.2.1 Enterprises and organizations

A lot of organizations and companies started developing free (or commercial) tools and products to automatically generate vocabularies/glossaries and their taxonomies. Some of them are listed below. The last two entries (Gene Ontology Project and BRITE) are related to other fields. They have been included for their general relevance for ontologies and for intra-EU interaction.

Collibra[3]

The Business Semantics Glossary [4] is a product of Collibra and it is a web-based, collaborative Data Governance platform where business and IT get aligned in defining the business vocabularies, facts and rules and manage data stewardship across the organization. It allows organizations to launch or mature their Data Governance programmes by helping them to define what their data means from a business perspective, involving all stakeholders, while integrating with their existing IT infrastructure. Its core capabilities:

- Manage business definitions, relations and rules,
- Manage taxonomies, classifications, hierarchies, reference data codes, …
- Involve all stakeholders through different roles and responsibilities,
- Capture any type of business metadata: Security, Policy, Audit, etc.
- Integrate with existing applications through an open API,
- Import and Export various formats: Excel, PDF, UML, XML Schema, RDF/OWL, …

Autonomy

Autonomy provides Automatic Classifier and Taxonomy Generator[5] and takes a holistic approach to taxonomy generation, balancing the use of automating processes and the support for manual refinement. By forming a conceptual understanding of the information in the enterprise, IDOL[6]

[2] http://www.heppnetz.de/projects/eclassowl/

[3] http://ww.slideshare.net/collibra/business-semantics-for-application-integration-and-soa

[4] http://www.collibra.com/products-and-solutions/products/business-semantics-glossary

[5] http://www.autonomy.com/content/Functionality/idol-functionality-categorization/index.en.html

[6] http://www.autonomy.com/content/Functionality/idol-functionality-idol-eduction/index.en.html

automatically generates taxonomies and instantly organizes data into a familiar child/parent taxonomic structure. Autonomy's data agnostic classification capabilities enable organizations to build taxonomies based on any type of data: unstructured, semi-structured, structured, audio and visual. The Autonomy Collaborative Classifier is a management interface for subject-matter experts: the people who know and use taxonomies, and knowledge engineers: the people responsible for the controlled vocabularies used within the enterprise.

Earley and Associates[7]

Earley & Associates is an information management consulting company experienced in taxonomy development and metadata management along with few other things. It also provides Webinars on Taxonomy and Metadata.

SchemaLogic

The SchemaLogic Industry Solutions Team[8] provides guidance and expertise for Information Management assessments worldwide. It has

1. designed, built, and evaluated taxonomic systems for business and government clients,
2. performed usability and user studies of metadata usage across a variety of systems,
3. published in technical and scholarly literature on metadata management.

SchemaLogic Enterprise Suite[9] enables business subject matter experts and IT professionals to define and manage a semantic standard.

Sandpiper Software™

Sandpiper Software[10] provides business semantics infrastructure solutions for context-driven search, collaborative applications, and cross-organizational content interoperability. The company develops semantically aware, knowledge-based products and provides context development services that facilitate business information interoperability, terminology normalization and context resolution across web-based and enterprise information systems. Sandpiper's context modelling and transformation products work in concert with enterprise application and data integration tools and services to facilitate collaboration and information-sharing among multiple databases, applications and users in multi-vendor environments.

WAND Product and Service Taxonomy [11]

The WAND Product and Service Taxonomy is a structured vocabulary of over 82,000 product and service terms of which approximately 42,000 are preferred terms and approximately 40,000 are synonyms. Adhering to the *ANSI/NISO Z39.19*[12] standard for thesaurus development, this taxonomy is one of the most extensive horizontal industry taxonomy in existence. Moreover, WAND has several other kinds of taxonomy packages[13] like General Business Taxonomy, etc.

eOTD

eOTD is the acronym for the ECCMA Open Technical Dictionary. It is the first dictionary to be compliant with ISO 22745 and it contains terms, definitions and images linked to concept identifiers. eOTD concept identifiers are used to create unambiguous language independent descriptions of individuals, organizations, locations, goods, services, processes, rules and regulations. The process

[7] http://www.earley.com/consulting-services/taxonomy
[8] http://www.schemalogic.com/services/the-industry-solutions-team/#team
[9] http://www.schemalogic.com/products/schemalogic-enterprise-suite/
[10] http://www.sandsoft.com/products.html
[11] http://www.sharepartxxl.com/products/taxonomy/partners/taxonomy/default.aspx
[12] www.niso.org/kst/reports/standards?step=2&gid=&project_key=7cc9b583cb5a62e8c15d3099e0bb46bbae9cf38a
[13] http://www.datafacet.com/downloads/DataFacet_Taxonomy_Library.pdf

of using concept identifiers from an external open technical dictionary is a form of semantic encoding compliant with the requirements of ISO 8000110:2008 the international standard for the exchange of quality master data. The eOTD is maintained by the Electronic Commerce Code Management Association (ECCMA). Vocabularies and taxonomies can be downloaded from http://www.eccma.org/resources/downloads.php

Drupal's Vocabulary[14] and Taxonomy Manager[15]

Drupal is an open source content management platform powering websites and applications. It is built, used, and supported by an active and diverse community of people around the world.

Common Procurement Vocabulary (CPV)[16]

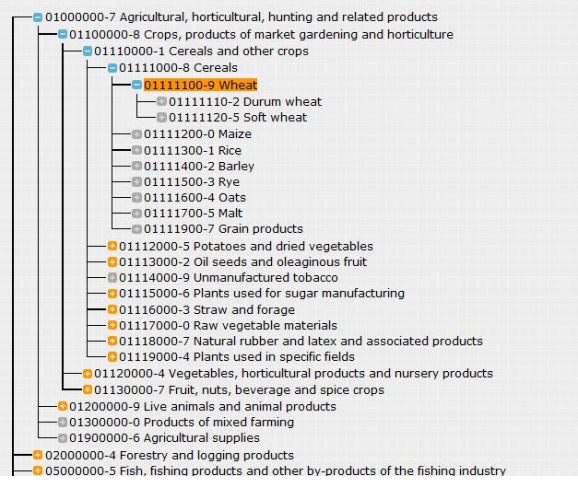


Figure 3 Sample classification obtained from CPV

The CPV establishes a single classification system for public procurement aimed at standardising the references used by contracting authorities and entities to describe the subject of procurement contracts. The CPV consists of a main vocabulary for defining the subject of a contract, and a supplementary vocabulary for adding further qualitative information. The main vocabulary is based on a tree structure comprising codes of up to 9 digits (8 digit code plus a check digit) associated with a wording that describes the type of supplies, works or services forming the subject of the contract.

The Contivo Vocabulary Management Solution[17]

Contivo's products help enterprises reduce the time, effort and cost of integration projects by leveraging the semantics and structure of data. Contivo's approach focuses on providing automation, leverage, and repeatability to integration projects.

[14] http://drupal.org/project/vocabulary
[15] http://drupal.org/project/taxonomy_manager
[16] http://www.cpvclassification.com/
[17] http://convergence.liaisonconnect.com/docs/Liaison-Contivo-VMS-WP.pdf

NICTA – Vocabulary Management[18]

NICTA has worked with partners in the financial services industry and in government to understand how and why multiple organizations arrive at a shared common understanding of their business. They also investigate the use and extension of semantic technologies such as SKOS and OWL in the design of a system to formally define vocabularies and their relationships, and to automatically check for logical consistency.

Medical Education Taxonomy Research Organization (METRO)[19]

With the ever-expanding use of information technologies in medical education and medical education research, there are growing needs for robust and appropriate semantic and ontological systems to facilitate cataloguing, storage and retrieval activities. METRO maps and creates terms to describe the processes, procedures and concepts of medical education. The terms will be used in a variety of purposes, such as describing content for retrieval from e-learning environments, indexing literature in databases, and extracting and exchanging data for research, such as for BEME[20] systematic reviews.

Scope[21]

Scope has developed its own taxonomy, thesaurus, lexicon and ontology for Medical and Life Science domains. It develops thesauri as per the international standards for thesaurus constructions such as ANSI/ NISO Z39.19-2005 Guidelines, to facilitate both pre-coordinate and post-coordinate indexing of scholarly articles.

3.2.2 Projects and Research papers

Most of the research projects in the recent years have focused mainly on developing ontologies. Deliverable "*A Business Data Ontology*" [22] from the project DIP[23] provides high-level reference ontology of business data.

A few other projects related to developing ontologies or vocabularies and taxonomies for business products and services are:

(project) ONTORULE: ONTOlogies meet business RULEs[24]

ONTORULE is a large-scale integrating project (IP) partially funded by the European Union's 7th Framework Programme under the Information and Communication Technologies (ICT) Call 3 (ICT-231875). The objective of ONTORULE is to integrate all the required pieces of knowledge and technology, including some that will need to be researched and developed within the project, to allow exactly that: the acquisition of ontologies and rules from the most appropriate sources, including natural language documents; their separate management and maintenance; and their transparent operationalisation in IT applications.

(Project) TOVE[25]

The goal of the TOVE project is to develop a set of integrated ontologies for the modelling of both commercial and public enterprises. (Fox *et al.* 1993) discusses how TOVE tackles these needs by

[18] http://www.nicta.com.au/research/projects/business_adaptation_and_interoperation/vocabulary_management
[19] http://onlinelibrary.wiley.com/doi/10.1111/j.1471-1842.2004.00534.x/pdf
[20] http://www.amee.org/index.asp?lm=104&cookies=True
[21] http://www.scopeknowledge.com/ContentEnhancement-KnowledgeServices/Taxonomy-Thesaurus-Ontology.aspx
[22] http://dip.semanticweb.org/documents/D3.3-Business-data-ontology.pdf
[23] http://dip.semanticweb.org/
[24] http://ontorule-project.eu/
[25] http://www.eil.utoronto.ca/enterprise-modelling/tove/index.html

defining a framework for modelling generic level representations such as activities, time, and resources. Since there has never been a well-defined set of criteria to evaluate such models, this paper also introduces a set of evaluation criteria which may be used to evaluate modelling efforts.

(Project) OBELIX: Ontology-Based ELectronic Integration of compleX products and value chains.[26]

The OBELIX project is an e-business ontology project. OBELIX focuses on researching and providing smart scalable integration and interoperability capabilities needed in the coming e-business stage of dynamic value constellations, characterised by much more complex products and services, supply chains and value networks, and associated electronic market transactions. It developed the OBELIX ontology tool suite for smart collaborative e-business.

(Project) ORBI: Ontology-based Reconciliation for Business Integration[27]

The objective of the project is to develop methods and application components for semantic e-business integration. In (Rebstock *et al.* 2008) novel methods and frameworks for ontology management and e-business integration are developed and the results are applied to e-business practice. (Rebstock *et al.* 2008) shows what ontology management can do for process, information and application integration under dynamic e-business conditions. The authors also published a series of articles[28] on some of the subtopics covered in the project.

(Project) MDT/SBVR[29]

This is an open source component of the MDT project, providing a meta-model implementation and sample tools based on the adopted *Semantics of Business Vocabulary and Business Rules (SBVR)*[30] OMG specification. In addition, this component will include sample tools that aid in testing and validating the meta-model implementation, and that help SBVR newcomers to better understand its use and potential.

The Enterprise Ontology

The Enterprise Ontology is a collection of terms and definitions relevant to business enterprises. The ontology was developed in the Enterprise Project[31] by the Artificial Intelligence Applications Institute at the University of Edinburgh with its partners. The project focused on management innovation and the strategic use of IT to help manage change. It supports the use of enterprise modeling methods which capture various aspects of how a business works and how it is organized.

(Project) unspscOWL[32]

An OWL ontology based on the United Nations Standard Products and Services Classification (UNSPSC[33]) was developed. By applying the GenTax[34] algorithm, a consistent OWL ontology has been derived from the UNSPSC.

(Project) eClassOWL[35]

eClassOWL is an OWL ontology for describing the types and properties of products and services on the Semantic Web (also known as the "Web of Linked Data"). eClassOWL is meant to be used in

26 http://www.e3value.com/projects/ourprojects/obelix/
27 http://www.orbi-project.org/
28 http://www.orbi-project.org/articles.htm
29 http://www.eclipse.org/modeling/mdt/?project=sbvr
30 http://www.omg.org/spec/SBVR/1.0/
31 http://www.aiai.ed.ac.uk/project/enterprise/
32 http://www.heppnetz.de/projects/unspscowl/
33 http://www.unspsc.org/
34 http://www.heppnetz.de/projects/gentax/
35 http://www.heppnetz.de/projects/eclassowl/

combination with the GoodRelations ontology[36] for e-commerce, which covers the commercial aspects of offers and demand, e.g. prices, payment, or delivery options.

(Project) MyOntology[37]

The aim of the research project "myOntology" is to produce the theoretical foundations and deployable technology for the Wiki-based, collaborative and community-driven development and maintenance of ontologies, instance data and mappings, in order to make the vision of a Semantic Web for E-Commerce and other application domains a reality.

(Project) MyClassify[38]

myClassify is an open-source API that implements standard machine-learning algorithms for the classification of product descriptions. It helps to generate proper classification of individual products according to a given hierarchy. Numerous other projects are also being investigated under Prof. Dr. Martin Hepp. Various other research projects in which he participated are listed here: http://www.heppnetz.de/projects/

An Analysis of Existing Ontological Systems for Applications in Manufacturing and Healthcare

This paper (Schlenoff et al. 1999) is one of the first attempts to move closer to the ultimate goal of seamless system integration using the principle behind ontological engineering to unambiguously define domain-specific concepts. The output of the entire work documented in this paper will be a taxonomy of terms and concepts, formal definitions of exactly what those terms and concepts mean, and how they interrelate.

A Methodology for Creating Ontologies for Engineering Design

This paper (Ahmed et al. 2007) describes a methodology for developing ontologies for engineering design. The methodology combines a number of methods from social science and computer science, together with taxonomies developed in the field of engineering design. The methodology is based upon empirical research and hence, focuses upon understanding a user's domain models as opposed to extracting ontology from documentation.

Development and Utilization of Ontologies in Design for Manufacturing

The primary goal of this paper (Chang *et al.* 2010) is to put forward the process of ontology development and utilization for (Design For Manufacturing) DFM and to study the most important phases in the process, including: the concept categorization and class hierarchy development, slot categorization and development, identification and realization of relations among slots, and methods to support knowledge capture and reuse.

InfoAnalyzer: a computer-aided tool for building enterprise taxonomies

(Zhang *et al.* 2004) studied the problem of collecting training samples for building enterprise taxonomies. The InfoAnalyzer tool can effectively assist the enterprise in preparing a large set of samples to be used for machine learning in text categorization.

Trying on Taxonomies: Research Developing Taxonomies for Analyzing and Developing Icon-Based Visual Language Systems

This paper describes research analyzing the visual language of Isotype using various taxonomic schemes and the application of those taxonomic schemes to current symbol systems (Meylor *et al.*,

[36] http://www.heppnetz.de/projects/goodrelations/
[37] http://www.myontology.org/index.html
[38] http://www.heppnetz.de/projects/myclassify/

2009). The aim of this research is to develop a shared understanding and controlled vocabulary for a collaborative, multi-centre symbol development project, *Hablamos Juntos*[39].

Domain modelling tools

This paper (Bennett 2007) gives an overview of the modelling formats and technologies available for modelling of the business problem domain, with particular reference to the financial services industry. It is intended primarily to group information on the leading open source tools and technologies available to the practitioner.

Gene Ontology Project[40]

The Gene Ontology project is a bioinformatics initiative with the aim of standardizing the representation of gene and gene product attributes across species and databases. Though it is not related to business product and services, it is one of the popular ontologies which showcase the importance and advantages of ontologies.

Ontology-based process mediation in the European project BRITE

The paper (Mondorf and Herborn 2008) elaborates the usage of a process ontology within the context of the integrated European project BRITE[41]. In an enlarged EU the Business Registers (BRs) need to interact across borders in a seamless and coherent way and thereby BRs have to dissolve administrative, technical, cultural and language barriers. The vision is that each BR is able to exchange and interpret company registration information coming from BRs of other countries electronically.

3.2.3 Communities and Research groups

Various communities and research groups are also actively working to provide up-to-date information about the various research activities going on around the world thus providing a good platform to get familiar with the various ways to develop vocabularies, taxonomies and ontologies. Some of them include:

Taxonomy Warehouse[42]

Taxonomy Warehouse was created in 2001 as a valuable community resource, available free to users and vocabulary publishers to help organizations maximize their information assets and break through today's information overload. Taxonomy Warehouse provides the information needed to effectively categorize internal and external data and ensures that the users find the information easily.

It also contains various taxonomies like Gale Business Thesaurus, IPSV taxonomy etc. The Gale Products and Services Thesaurus[43] is a subset of the master Gale Business Thesaurus[44] in the narrower domain of all types of marketable products and services. Products include manufactured goods, processed goods, and raw materials (including plants, animals, and minerals if they are sold commercially). Services comprise all those that are available commercially. This is the largest category of subject descriptors, since it contains thousands of highly specific terms.

[39] http://www.hablamosjuntos.org/
[40] http://www.geneontology.org/

[41] http://www.ecgi.org/brite/index.php
[42] http://www.taxonomywarehouse.com/include_resources.asp
[43] http://www.taxonomywarehouse.com/vocabdetails_include.asp?vVocID=16
[44] http://www.taxonomywarehouse.com/vocabdetails_include.asp?vVocID=3

E-Business and Web Science Research Group

They work on the theoretical and practical challenges of using the World Wide Web for facilitating the exchange of tangible and intangible goods between economic actors, and economic activity in general. Their core research interest is employing Semantic Web technology for E-Procurement and Business Process Management.

SLA Taxonomy Division[45]

It contains information on Controlled Vocabularies including Taxonomies, Thesauri, Ontologies, Terminologies, and other Knowledge Organization and Classification Systems.

Taxonomies & Controlled Vocabularies SIG[46]

This is a Special Interest Group of the American Society for Indexing. It organizes events consisting of presentations or workshops on taxonomies and Controlled Vocabularies. It also provides online courses for developing the above.

Taxonomy Watch[47]

This is a weblog about taxonomies and their applications in organizing digital content, also including related topics such as controlled vocabulary, thesauri, topic maps, ontologies and semantic technologies.

Taxonomy Research & Information Network[48]

Taxonomy Research & Information Network (TRIN) addresses critical gaps in taxonomic knowledge of key Australian animal and plant groups. Research within the current projects of TRIN covers a range of taxonomic and systematic investigators using a diversity of approaches and methodologies. TRIN provides the basis for identifying and monitoring Australia's biodiversity which in turn provides the knowledge needed for effective environmental management.

The MONTAGUE INSTITUTE[49] also provides a course on Creating and Using Business Taxonomies.

[45] http://wiki.sla.org/display/SLATAX/SLA+Taxonomy+Division

[46] http://www.taxonomies-sig.org/resources.htm and http://www.taxonomies-sig.org/links.htm

[47] http://taxonomy2watch.blogspot.com/

[48] http://wiki.trin.org.au/

[49] http://www.montague.com/five-virtual.html

4. Horizontal Cross-Domain Vocabularies

There are many standards available to define the vocabulary of any kind of system. Using these standards many vocabularies are also developed which could be used either in various domains (cross-domain) or for specific purposes only (industry-specific). This section introduces some of the cross-domain vocabularies and standards which are relevant to (and useful for) the ebbits project.

4.1 Relevant standardization organizations

Below are some of the cross-industry standards which are widely used today and based on which vocabularies are also developed.

- **ISO**[50]

National standards institutes of countries around the globe together form the International Organization for Standardization. ISO is recognized as one of the four global de jure (force of law) standards development organizations. ISO standards encompass a broad spectrum of technical and business focus areas and have an impact on virtually every business and information technology community.

ISO 13584-511 Fastener Taxonomy (has replaced North American Industry Classification System NAICS at DLIS request). ISO 13584-511:2006 specifies a reference dictionary for representing fasteners with their properties and domains of values, as they are described in the various ISO mechanical fastener standards. These fasteners include bolts, screws, nuts, rivets, pins, washers etc. The reference dictionaries series of parts of ISO 13584 specify ontologies for representing the entities of an application domain, together with their descriptive properties and domains of values.

ISO 10303-1:1994, Industrial automation systems and integration – Product data representation and exchange – Part 1: Overview and fundamental principles. It provides a representation of product information along with the necessary mechanisms and definitions to enable product data to be exchanged. Applies to the representation of product information, including components and assemblies; the exchange of product data, including storing, transferring, accessing, and archiving. It defines the basic principles of product information representation and exchange used in ISO 10303.

- **OMG**[51]

In the field of business vocabulary and semantics of business vocabulary there is a very intensive standardization initiative provided by the Object Management Group (OMG), which resulted in the Semantics of Business Vocabulary and Business Rules' (SBVR) standards "Vocabulary for Describing Business Vocabularies" based on the ISO terminology standards ISO 1087-1 (2000) "Terminology work — Vocabulary — Theory and application" [ISO1087-1]; ISO 704 (2000) "Terminology work — Principles and methods" [ISO704]; ISO 860 (1996) "Terminology work – Harmonization of concepts and terms" [ISO860]. These standards have been used for many decades for multilingual correlation of vocabularies in support of language translation work. SBVR is the result of the integration of these ISO standards, formal logics, linguistics, and practical experience from foremost practitioners in the field of business vocabulary for business rules.

- **United Nations Standard Products and Services Code (UNSPSC)**

UNSPSC is the Universal Standard Products and Services Classification. The UNSPSC Code is a coding system to classify both products and services for use throughout the global marketplace. The management and development of the UNSPSC Code is coordinated by ECCMA, the *Electronic Commerce Code Management Association*. The current version consists of more than 18 000 terms.

[50] http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html
[51] http://www.omg.org/

The public version of the current code set can be downloaded for free from http://www.unspsc.org/, http://www.eccma.org/resources/downloads.php

- **eCl@ss**[52]

  Leading German companies developed their own classification called eCl@ss. eCl@ss is characterized by a 4-level, hierarchical classification-code with over 21 000 subgroups and a keyword registry with 75 000 search options. It should display the source-market for purchasers and assist engineers with development, planning and maintenance. eCl@ss depicts all of the products across all of the product lines and describes the articles explicitly with up to 1000 features.

- **OAGi – Cross Industry**[53]

  The Open Applications Group (OAGi) is a not-for-profit open standards organization that defines process-based XML standards called Business Objects Documents (BODs) that span both B2B and A2A integration scenarios. The adoption of OAGi specifications is primarily in North America, with emerging adoption in the European Union and Asia. The OAGi specifications are used in a wide variety of integration scenarios for A2A integration. Conversely, implementations for B2B are typically in manufacturing with a particular emphasis in the North American automotive market. The OAGi specifications are used in business solutions packages (e.g. from SAP), which are designed to streamline supply chain management processes, reducing costs and improving time-to-value.

- **ANSI X12**[54]

  The American National Standards Institute (ANSI) Accredited Standards Committee X.12 is responsible for creating cross-industry business standards using ANSI ASC X12 and XML syntax. ANSI cross-industry standards include more than 315 individual EDI transaction sets that cover five markets including finance, government, transportation, supply chain and insurance (both medical and hazard).

  A few other, less popular cross-domain standards also exist.

- **UN/CEFACT**[55]
  The United Nations/Centre for Trade Facilitation and Electronic Business (UN/CEFACT) is a chartered activity of the UN Economic Commission for Europe (UN/ECE). The UN/CEFACT mission is to support, enhance, and promote trade facilitation between developed, developing and transitional economies. UN/CEFACT standards consist of EDI Messages and supporting data elements, technical specifications for the UN/CEFACT EDI syntax (EDIFACT), technical specifications for methodologies, and syntax specific implementation specifications such as XML Naming and Design Rules.

- **IEC62264**[56]

  It is a multi-part standard that defines the interfaces between enterprise activities and control activities. This standard is based upon ANSI/ISA-95.00.01-2000 (Systems and Automation Society - ISA). This standard provides standard terminology and models for describing the interfaces between the business systems and production control systems of an enterprise. It consists of the following parts: IEC 62264-1, Enterprise control systems and integration – Part 1: Models and terminology; IEC 62264-2, Enterprise control systems and integration – Part 2: Object model attributes; IEC 62264-3, Enterprise control systems and integration – Part 3: Models of manufacturing operations.

- **EPCglobal**[57]

  EPCglobal is a non-profit consortium that drives the collaborative development of Electronic Product Code (EPC) related RFID standards. EPCglobal was formed in 2003 as a member of the GS1 Global

[52] http://www.eclass.de/template/intro_en.html
[53] http://www.oagi.org/dnn2/
[54] http://www.disa.org/bookstore/public/index.cfm
[55] http://www.unece.org/cefact/ebxml/CCTS_V2-01_Final.pdf
[56] http://en.wikipedia.org/wiki/IEC_62264
[57] http://www.gs1.org/gsmp/kc/epcglobal

family of standards organizations and is based on RFID research performed at the AutoID Center at the Massachusetts Institute of Technology (MIT). The EPCglobal specifications and the EPCglobal Network Architecture define mechanisms by which companies can efficiently communicate information about their business processes in real time. Companies are enabling entirely new business processes with RFID event data, yielding significantly increased efficiency in supply chain operations.

- **Federal Catalog System (FCS)**[58]

  The Defence Logistics Agency (DLA) administers the Federal Catalog System (FCS) under the direction of the Assistant Secretary of Defence (Installations and Logistics). The FCS encompasses the naming, description, classification, and numbering of all items carried under centralized inventory control by the Department of Defence (DOD) and the civilian agencies of the US Federal Government, as well as the publication of related identification data. The FCS is also used by North Atlantic Treaty Organization (NATO) countries.

4.2 Horizontal vocabulary distilled from agricultural domain

Ebbits focuses on two domains, manufacturing and traceability. Vocabularies specific for these domains will be described in the following chapters. In this subsection focus is on the part of the horizontal vocabulary that is relevant for the agricultural domain.

The world is an international marketplace where products and commodities are traded between various businesses. The goods need to be identifiable, measurable and categorized for this marketplace to function. The horizontal vocabulary for ebbits is just a minor subset of what will be needed but at this stage it focuses on goods and their attributes, orders and transportation of goods. The list below is based in input from the feed production industry (DLG), the Farmers Organisation in Denmark (Danish Agriculture & Food Council), slaughtering industry (Danish Crown). It is extracted to only contain terms that are considered to be general and used outside the agricultural domain as well.

1. Goods
   a. Price
   b. Weight
   c. Storage Conditions
   d. Shape/Form
   e. Pack type
   f. GTIN [Global Trade Item Number]
   g. Batch ID
   h. Location
2. Orders
   a. Order
      i. Date
      ii. Goods
         1. GTIN
         2. Quantity
         3. Price
      iii. Total price
      iv. Taxes
   b. Order confirmation
      i. Date
      ii. Order status
3. Transportation
   a. Goods
      i. GTIN
      ii. GLN [Global location Numbers]
      iii. Quantity
      iv. Weight
      v. Size

[58] http://www.dlis.dla.mil/PDFs/Procedures/vol01.pdf

b. Locations
    i. From
    ii. To
c. Processes
    i. Transporting
    ii. Storing
    iii. Importing
    iv. Exporting

4.3 Horizontal vocabulary distilled from manufacturing domain

As already described in chapter 4 the aim of this section is to examine the manufacturing and traceability domains in order to identify a vocabulary common for both environments.

Like the traceability scenario, manufacturing is also a very large and heterogeneous environment. A large number of devices, data and sub-processes is involved in the production processes.

Unlike the agricultural domain, it is more difficult to organize the vocabulary in a hierarchic form in manufacturing. For this reason the "manufacturing terms" are listed in a table and a brief description for each of them is provided.

Terms collected from the manufacturing environment that could be considered as common in Manufacturing and traceability environments (the first list is based on the initial glossary produced at the beginning of the project)[59]:

Manufacturing environment	Traceability environment	A	
		Accuracy	The degree of exactness of a model or simulation, high accuracy implying low error. Accuracy equates to the quality of a result, and is distinguished from precision, which relates to the quality of the operation by which the result is obtained and can be repeated.
		Activity	In modelling and simulation, a task that consumes time and resources and whose performance is necessary for a system to move from one event to the next.
		Activity-based simulation	A discrete simulation that represents the components of a system as they proceed from activity to activity; for example, a simulation in which a manufactured product moves from station to station in an assembly line.
		Activity models	Models of the processes that make up the functional activity showing inputs, outputs, controls, and mechanisms through which the processes of the functional activity are (or will be) conducted.

[59] Found on http://www.glossaryofmanufacturing.com/ and http://www.glossaryofmanufacturing.com/

		ADLs	Architecture Description Languages are used to define software architectures consisting of components and their connections. In case of technical architecture, the architecture must be communicated to software developers to provide the required characteristics. With functional architecture, the architecture is communicated to stakeholders and enterprise engineers.
		Algorithm	A prescribed set of well defined unambiguous rules or processes for the solution of a problem in a finite number of steps.
		Analytical model	A model consisting of a set of solvable equations
X		**ANSI**	American National Standards Institute
X		**Architecture**	The structure of components in a program/system, their interrelationships, and the principles and guidelines governing their design and evolution over time.
		B	
		Bandwidth	A structure or machine which is put together (or "built") by fitting together two or many sub-assemblies. By contrast, a sub-assembly usually has a defined, single function within the overall assembly in which it is incorporated. Sub-assemblies themselves are likely to be built from individual product components.
		Barcode	One Dimensional (1D) bar codes are the familiar (*) representation of the individual characters and digits making up an item's code by a succession of vertical lines of varying thickness, the lines capable of being read and interpreted by a computer scanning device. There are numerous alternative 1D systems for representing an item's code. Examples are Interleaved 2 of 5 (having the advantage of high physical density of the code); Code 39 (popular in manufacturing industry); EAN (European Numbering System); and UPC (Universal Product System). See the individual entries for these four systems. See also RFID tags. Two Dimensional (2D) bar codes, also known as stacked bar codes, comprise small postage stamp size complex patterns that appear to the naked eye as square dots and

			other tiny geometric shapes. 2D bar codes convey vastly more data than their 1D cousins and have been referred to as portable databases. Technologies include PDF417, MaxiCode and DataMatrix. They require to be read by special scanners.
		B2B	Business-to-Business, also known as e-biz, is the exchange of products, services, or information between businesses rather than between businesses and consumers.
		Benchmark	A standard, used for comparison or reference.
		C	
		CACSD	Computer-Aided Control System Design. Software tools used to support the control engineer during the description, specification, implementation and validation of automation systems
		CASE	Computer-Aided Software Engineering. Software tools used to support the control engineer during the description, specification, implementation and validation of automation systems
X		**CAD**	Computer-Aided Design. Computer-based systems for product design that may incorporate analytical and "what if" capabilities to optimise product designs. Many CAD systems capture geometric and other product characteristics for engineering-data-management systems, producibility and cost analysis, and performance analysis.
		CIM	Computer-Integrated Manufacturing. A variety of approaches in which computer systems communicate or interoperate over a local-area network. Typically, CIM systems link management functions with engineering, manufacturing, and support operations. In the factory, CIM systems may control the sequencing of production operations, control operation of automated equipment and conveyor systems, transmit manufacturing instructions, capture data at various stages of the manufacturing or assembly process, facilitate tracking and analysis of test results and operating parameters, or a combination of these.

X		**Computer simulation**	Dynamic representation of a model, often involving some combination of executing code, control/display interface hardware, and interfaces to real-world equipment
		Conceptual model	A statement of the content and internal representations which are the user's and developer's combined concept of the model. It includes logic and algorithms and explicitly recognizes assumptions and limitations
		Constrained simulation	A simulation where time advances are paced to have a specific relationship to wallclock time. These are commonly referred to as real-time or scaled-real-time simulations. Here, the terms constrained simulation and (scaled) real-time simulation are used synonymously. Human-in-the-loop (e.g., training exercises) and hardware-in-the-loop (e.g., test and evaluation simulations) are examples of constrained simulations
X		**Control Logic**	Logic which controls a *device* in context of its use
		Co-simulation	See [*Solver, Coupling*]
		Coupling, Leak	See [*Solver, Coupling*]
		Coupling, Strong	See [*Solver, Coupling*]
		Coupling, Pseudo-strong	See [*Solver, Coupling*]
X		**Clamp**	A stationary work-holding device having one or more moveable jaws that is used to clamp a workpiece.
		Class Diagram	Class diagrams represent a set of classes, interfaces and collaborations with their relationships. Class diagrams describe the static structure of a system.
		Collaboration Diagram	A collaboration diagram describes interactions among objects in terms of sequenced messages. Collaboration diagrams represent a combination of information taken from class, sequence, and use case diagrams describing both the static structure and dynamic behaviour of a system.
		Conveyor	a machine used to move parts and materials along an assembly line at a constant rate of speed
		D	

		Data	A representation of facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, or processing by humans or by automatic means
		Data Flow Diagram	A diagram, typically automated using various tools, which depict data flowing from one process to another. This model type is used within the companies to depict major data flows and related activities within a system or across the enterprise. These techniques are defined in the General Systems Architecture Assessment.
		DCM	Design Cycle Model. Is the reference model used to structure all the different phases needed for an automation system design. The different phases are: description of the process to be controlled, specification of the behaviour of the controlled process, definition of the control system architecture and functionalities, software code implementation, software code validation
		Descriptive model	A model used to depict the behaviour or properties of an existing system or type of system
		Deterministic	Pertaining to a process, model, simulation or variable whose outcome, result, or value does not depend upon chance
X		**Device**	*Electrical, Mechatronic* component
		Disaggregation	The ability to represent the behaviour of an aggregated unit in terms of its component entities
		Distributed simulation	A set of disparate models or simulations operating in a common synthetic environment
		Domain	The physical or abstract space in which the entities and processes operate
		E	
		Encapsulation	The process of hiding the details of an object that do not contribute to its essential characteristics
		Event-driven	The *control* of a system is *event-driven* when it takes account each event
		Event Modelling	A series of events, depicted in a

			diagrammatic format, that mirror real life activity cycles and can be interpreted by certain development tools in order to create operational systems.
		Expert system	An expert system is a knowledge collection combined with an inference engine capable of interpreting queries and chaining together separate items of knowledge to develop new inferences. The knowledge is typically causally represented as a system of rules. In some cases, expert systems can retrace their paths of inference on demand, thus explaining their conclusions and reasoning
		F	
		Fidelity	The accuracy of the representation when compared to the real world
		FLP **Functional Logical Physical**	Three different levels of abstraction when considering system modeling and simulation. See also [*Model, Functional]*, [*Model, Logical]*, [*Model, Functional*] and [*Model, Multilevel*]
X		**Fixture**	a device used to locate or support material, workpiece(s), or tool(s) during machining operations. Usually custom designed to hold a specific workpiece.
		G	
		Graphical model	A symbolic model whose properties are expressed in diagrams
		H	
		Heterarchical	A heterarchy is a network of elements sharing common goals in which each element shares the same "horizontal" position of power and authority, each having an equal vote. Heterarchy can be defined as an organizational form somewhere between hierarchy and network that provides horizontal links that permit different elements of an organization to cooperate whilst individually optimizing different success criteria.
		Heuristic	Relating to or using a problem-solving technique in which the most appropriate

			solution of several found by alternative methods is selected at successive stages of a program for use in the next step of the program
		Higher level architecture	Defines the functional elements, interfaces, and design rules for simulation interoperability. Provides a common framework within which specific system architectures can be defined
		Hierarchical	The hierarchical data model organizes data in a tree structure. There is a hierarchy of parent and child data segments. This structure implies that a record can have repeating information, generally in the child data segments.
		Homogenous, system	The parts of which cannot be mechanically separated and which has uniform physical properties throughout its mass or volume.
		Heterogeneous, system	The system such that contains various distinct and mechanically separable parts or phases, such as a suspension.
		Hybrid simulation	A simulation that combines constructive, live, and/or virtual simulations, typically in a distributed environment. Such simulations typically combine simulators with actual operational equipment, prototypes of future systems, and realistic representations of operational environments
		I	
		Implementation	The means by which a synthetic environment, or portions of a synthetic environment, is realized
		Internal Logic (of device)	A logic that models the behavior of a *device.*
		Interoperability	The ability of a model or simulation to provide services to and accept services from other models and simulations, and to use the services so exchanged to enable them to operate effectively together
		J	

		K	
		Knowledge-based system	A system in which the domain knowledge is explicit and separate from the system's operational instructions/information
X		**Kanban**	A Japanese word with no equivalent in English (except, perhaps, "signage"), used by The Toyota Car Company since 1953 to mean a card about 3" x 5" in a plastic vinyl pouch holding brief, simple information relating to the manufacture to hand. A formal definition of kanban is: a card or signal conveying information and indicating the need to take action within a Just-in-Time or lean manufacturing system. Toyota use kanban to control the production of parts (the production kanban) and the movement of parts (the movement kanban). The number of kanban cards in the system is critical, since this determines the amount of stock present on the factory floor: every effort is made to operate with fewer and fewer cards so as to reduce such stock. To calculate the number of cards, see the Toyota Equation. Kanban in Japanese is pronounced "con bon". Note that a "two bin" system of replenishment, whereby a new lot is called for from supplies when a previous lot has been used up, is often (wrongly) called a kanban system - people who insist on calling their simple two-bin system "kanban" also like to call the two-bin replenishment ticket a kanban ticket. (God bless the British amateur.)
		L	
		M	
X		**Mechatronic**	Approach to the of intelligent *dynamic systems* integrated across engineering disciplines
		Metadata	Descriptive information about data in a resource includes domain assignment, ownership, access restrictions, and database/model.
		Metaknowledge	Descriptive information about knowledge in a resource includes ontology, domain coverage, ownership, access restrictions, and representation.
		Methodology	A detailed and structured approach,

			containing generic and tool related step-by-step guidelines, to developing, upgrading, improving or replacing application systems.
		Modelling	Transition from a *dynamic system* to a *model description* of a certain *level of modelling abstraction*, *idealization* and *granularity*
		Modelling (3D)	Mechanical *modelling* in 3D context
		Modelling, interval	Interval of time, of *ambient parameters* and of *system variables* as well as of their time rates determined by *model validation*
		Model, conceptual	Characterizes causes-and-effects in a *dynamic system* or *system module* respecting algebraic rules
		Model, description	Equations, block diagram, bond graph, multipole diagram, etc.
		Model, discrete	*Functional model* the model variables of which are approximated by discrete-time and/or discrete-level functions
		Model, functional	Characterizes interrelations of variables a *dynamic system* or *system module* respecting algebraic rules
		Model, hybrid	*Discrete model* combined with a *physical model* or a continuous-time continuous-level *functional model*
		Model, identification	Process of determination of *model parameters* based on physical hypotheses and/or *experiments*
		Model, multidomain (Model, multiphysics)	*Physical model* the parts of which belong to different *energy domains (= multiphysics)*
		Model, multilevel	Model the parts of which are considered on a different *level of modelling abstraction*
		Model, object	A specification of the objects intrinsic to a given system, including a description of the object characteristics (attributes) and a description of the static and dynamic relationships that exist between objects
		Model, parameters	Characterize a model regardless of its current dynamic state
		Model, physical	Characterizes continuous-time continuous-level *energetic interactions* of *system modules* respecting physical laws
		Model, predictive	A model in which the values of future states

			can be predicted or are hypothesized
		Model, representation	geometric (showing geometric dimensions and positions), structural (displaying the topology of internal structure) or behavioural (concerned only about the outer effects)
		Model, technological	Model of a technological process required to produce a *dynamic system*
		Model, validation	Process of relating the *system model* or *module model* to a class of *experiments* to be performed on a *dynamic system*
		Model, variables	Characterize the current dynamic state of a model
		Module, model	Abstract and idealized representation of a *system module* in the form of a *model description* related to a class of *experiments*
		N	
		Notation	A system of signs or symbols used to represent information; a graphical or textual set of rules for representing a model
		O	
		Open system	A system in which the components and their composition are specified in a non-proprietary environment, enabling competing organizations to use these standard components to build competitive systems. There are three perspectives on open systems: portability - the degree to which a system component can be used in various environments, interoperability - the ability of individual components to exchange information, and integration - the consistency of the various human-machine interfaces between an individual and all hardware and software in the system
		P	
		Parallel simulation	See [*Solver, Coupling*]
		Prototype, real	*Dynamic system* built to verify correctness of a *system design* by *experiments*
		Prototype, virtual	*System model* set up to verify correctness of a *system design* by *simulation experiments*

X		**Programmable Logic Control (PLC)**	*Digital electronic stored program device used in sequencing, timing, counting, and arithmetic to control – through digital or analogue input/output modules – various types of machines or processes.*
		Q	
		R	
		Range	Usually, the difference between the smallest and largest measurement in a sample. The range is used in Statistical Process Control (SPC) and is plotted on a variable control chart.
		Rapid Prototyping	Involves iterative, rapid refinement of software designs by quickly generating working prototypes and using feedback of prototype results to improve design specifications.
		Replication	Technique used by legacy designers and programmers to rapidly deploy replacement systems or build new programs by copying and modifying existing programs. It differs from reusability in that functionality of the originating component is cloned and modified one or more times, forcing departure from the original baseline and replication of maintenance effort over the long term.
		Resolution	The degree of detail and precision used in the representation of real world aspects in a model or simulation
		Reusability	The characteristic of a component that allows it to be used in more than the application for which it was created, with or without modification.
X	**X**	**RFID**	Radio Frequency Identification. The attachment to a product of a transponder bearing the product's code and other data, to enable the product's identity to be recorded on an electronic scanner positioned some way from it
		Runtime engine	The device and/or system software that runs a software application (e.g. a PLC, an embedded system with its Operating System, a UNIX- or MS Windows-based

			workstation, ...).
		Robot	a computerized or electromechanical automatic, general-purpose device, whose primary function is to produce motion in order to accomplish some task. Computerized robots usually consist of three main components: the machinery or mechanical parts, the controller/computer system, and the software. An industrial robot is designed especially for industrial use such as materials handling, tool changing, assembly, welding, and measuring. An industrial robot can be programmed or taught by using a digitizing system which translates movements into commands for the robot to understand. The five basic types of industrial robots are: Anthropomorphic, or Articulated Robot: has movements such as the human arm, with rotating shoulder, bending elbow, and rotating wrist. Cylindrical Robot: similar to the Cartesian model except that it rotates about a stationary base instead of moving from side to side. Overhead, or Gantry Robot: moves on a crane or bridge-type support; the arm can have a number of axial movements. Polar, or Spherical Envelope Robot: rotates about a perpendicular axis; its arm is capable of moving in and out and through an up-and-down arc. Rectilinear, or Cartesian Robot: moves in straight lines, the X, Y, and Z axes
		S	
X		**Scalability**	The ability of a distributed simulation to maintain time and spatial consistency as the number of entities and accompanying interactions increase
		Sequence diagram	A sequence diagram is a means to illustrate a use case by representing the collaborations between entities from a temporal point of view.
		Simulation	The exercising of a model over time
		Simulator	A device, computer program, or system that performs simulation
X		**Servo Control**	Industrial robot control system in which sensing devices monitor movement, and report any deviation between commands as issued and movement as monitored.

			Deviations will automatically cause corrective action to be taken
X		**Smart Device**	*Mechatronic* component with an *internal logic*
		SOA	A Service-Oriented Architecture is a collection of services that communicate with each other. The services are self-contained and do not depend on the context or state of the other service. They work within a distributed system architecture.
		Software Re-engineering	The use of tools and techniques to facilitate the analysis, improvement, redesign and reuse of existing software systems to support changing business and technical information requirements.
		Solver	In simulation, a component that computes the model's behaviour. Usually considered as the simulator's core component. See also [*Simulator*]
		Solver, coupling	Two solvers are coupled when one need information from another to make its computations. The particular case of no coupling is named co-simulation or parallel simulation. Eventually, the time may be the only information shared by both solvers. Coupling is weak when a solver A needs information from a solver B, but B doesn't need any information from A. Weak coupling induce a processing order over the various solvers involved in a simulation. Coupling is strong when both solvers need information from the other. It becomes pseudo-strong if information is delayed by a time step between producer and consumer.
		Stakeholder	All persons or organisations affected by the system. This includes the user, the customer, the employer, developers, regulatory bodies, maintenance staff, support desk, etc.
		STD	State Transition Diagram. A diagram consisting of circles to represent states and directed line segments to represent transitions between the states. One or more actions (outputs) may be associated with each transition. The diagram represents a finite state machine.
		System, analysis	A subset of *system simulation* related to an

			analysis mode
		System, design	Process of determining the structure and parameters of a *dynamic system* respecting specified criteria
X		**System, diagnostics**	Investigation of the cause of a (possible) *dynamic system* failure
		System, dynamic	Region in space filled with quantities of matter interrelated by *energy interactions*
X		**System, mechanical**	Collection of bodies in which some or all of the bodies can move relative to one another
X		**System, mechatronic**	*Multidisciplinary system* with intelligent control
		System, model	Abstract and idealized representation of a *dynamic system* in the form of a *model description* related to a class of *experiments*
		System, modularization	Process of imaginary decomposition of a real *dynamic system* into real *system modules*
		System, module	Part of a *dynamic system* separated from the remainder of the system by a *module boundary*
		System, multidisciplinary	*Dynamic system* the investigation of which requires experts from different traditional engineering disciplines
		System, optimization	Process of modification of system *model parameters* to achieve specified criteria
		System, surrounding	Mass or region outside a *dynamic system*
		System, synthesis	Systematic and unambiguous *system design* procedure
		T	
		Takt Time	Total available production time divided by the customer & requirement. Note: Include all planned activities such as clean-up, safety meetings, etc. Example (1) 8 Hour Shift=480 Minutes-(2) 10 Minute Breaks=460 --- 1840 Pieces/Day Customer Requirements TAKT Time = .25 minute or 15 seconds
		Task analysis	The study of tasks, their sequences of action, and their information requirements.

		Task (of device)	Function performed by a device. E.g MOVE A to B for a robot
		Three(3)-D	Three-dimensional, refers to the visual display that exhibits breadth, height and thickness or depth.
		Time driven	Considering a real-time system control, consists in observing periodically the global state of the system, instead of reacting to single event. Whatever the number of event s occurred during a period, what is detected is their global effect.
		Time, CPU	In a simulation context, the amount of time spent by solvers to compute model behaviour. In a virtual time simulation, CPU time is considered orthogonal to virtual time axis (virtual time is suspended during solver's computation). But in real-time context, CPU time is partially or completely projected on time axis, making real-time simulation more difficult to get. See also [*Time, Virtual*] and [*Time, Real*]
		Time, Virtual	In a simulation context, theoretical idea of the time that goes by. Used as a mathematical entity and a logical model. May differ from real time since it can be suspended, restarted, reset accordingly to simulation requirements. May also go faster or slower than real time. But anyway, must be perceived identically by all solvers involved in the simulation. See also [*Time, CPU*] , [*Time, Real*] and [*Virtual Time Manager*]
		Time, Real	In a simulation context, simulated time advances at the same rate as actual time. See also [*Time, CPU*] and [*Time, Virtual*]
		U	
		Usability	The extent to which a system or product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in specified context of use.
		UML	Unified Modelling Language. A standardized visual notation for communication about system specifications and design details. UML is a non-proprietary, object modelling and specification language used in software engineering. It includes a standardized graphical notation that may be used to create an abstract model of a system.

		Use case	A use case is a set of scenarios tied together by a common user goal. A scenario is a sequence of steps describing an interaction between a user and the system.
		V	
		Validation	The process of determining the degree to which a model or simulation is an accurate representation of the real-world from the perspective of the intended uses of the model or simulation
		Verification	The process of determining that a model or simulation implementation accurately represents the developer's conceptual description and specification. Verification also evaluates the extent to which the model or simulation has been developed using sound and established software engineering techniques
		Virtual Validation	*Validation* of the *digital model* of the production system
		Virtual Time Manager	In a simulation context, the component in charge of making the virtual time going by and telling the solvers about it.
		Visual Controls	Graphical images that display the status of an activity.
		VRML	Virtual Reality Modelling Language (VRML) is an Internet standard for the rendering of 3D graphics. To view these files, requires a VRML browser or a VRML plug-in to a Web browser.
		W	
X		**Work cell**	Set of *devices* which produce one or more products.
		X	
		Y	
		Z	

5. The Automotive Manufacturing Domain

This Chapter contains a description of data (which are currently stored in databases) on Body welding and assembly, Energy reduction management process etc.

In an automotive manufacturing plant, during the production process, the system accesses data stored in a database. The main information exchanged among the different production lines are regarding the production data, but the new requirements fixed by the customers are evolving in the "green" direction with a focus on consumption and environmental impact.

Standards and Specifications specific for manufacturing environment

During the design and the realization of a new machine or manufacturing plant many directives must be considered in order to avoid safety risks. The fields of application are summarized in Figure 4:

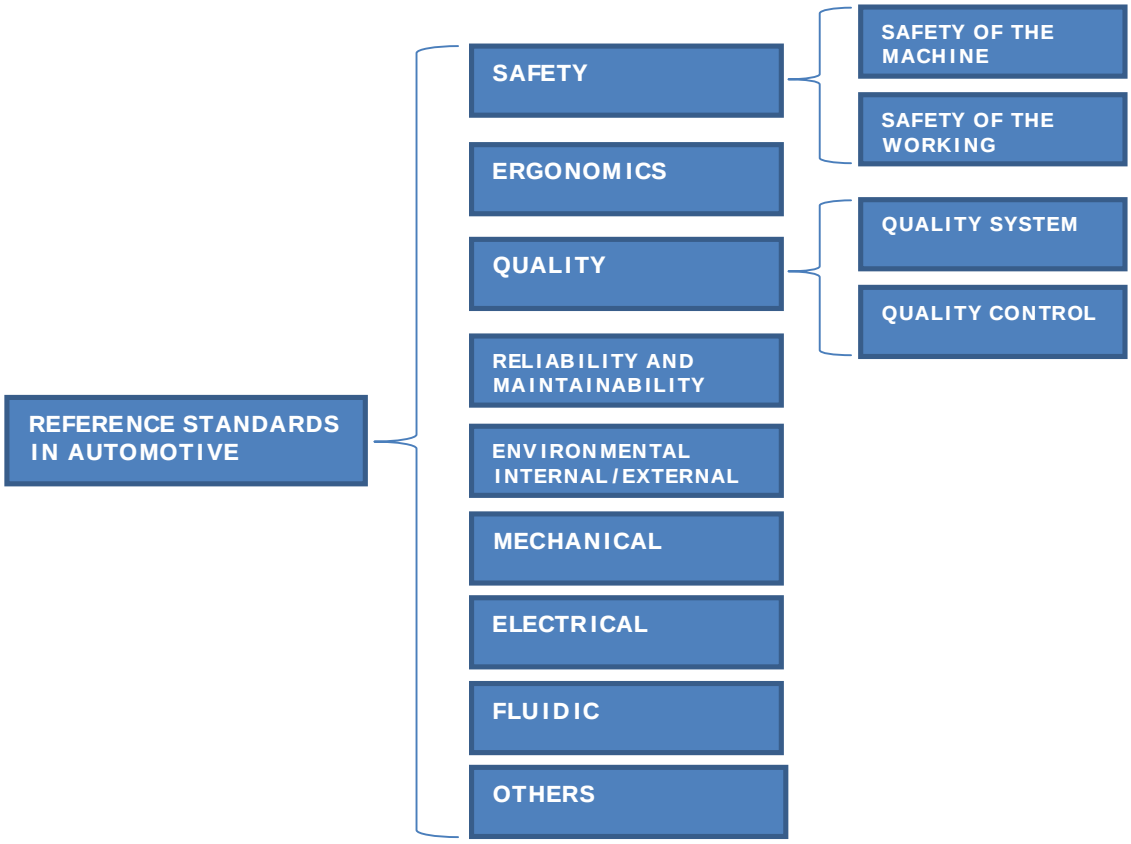


Figure 4 Fields of application in the manufacturing sector

In manufacturing several standards are used in order to certify the product and to assume that it complies with the laws and norms applicable to it, in particular:

- **EN:** the European Norm
- **ISO:** for general, mechanic and fluidic norms (used all over the World)
- **EC:** for electric parts and reliability (used all over the World)
- **UNI and CEI:** Italian norms
- **DIN and VDE:** German norms
- **BS:** British Standards
- **CNOMO and AFNOR:** French
- **MIL, SAE, NIOSH, UL**: USA

The most important directives applied in manufacturing are:

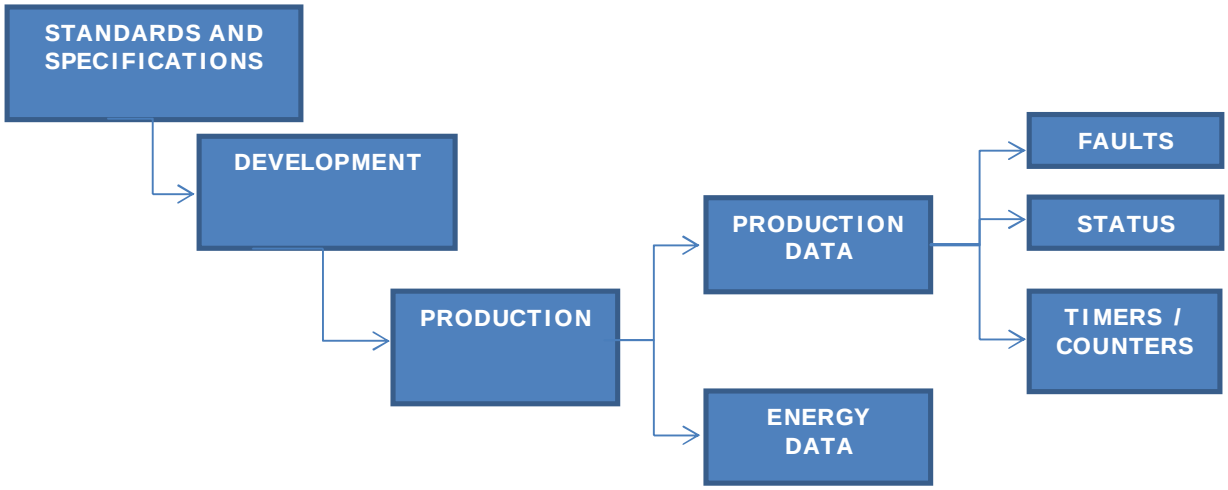
- **Directive 2006/42/CE**: machinery directive
- **Directive 2004/108/CE**: that considers "EMC"
- **Directive 2006/95/CE:** for the low voltage circuits
- **Directive 97/23/CE:** for the equipments under pressure
- **ATEX:** directive that considers explosive environments

In particular in Italy a document called "Testo Unico 81/08" includes all the Italian laws on the safety.

Other norms applied are:

- **EN/ISO 12100:** general for all the safety
- **UNI/EN/ISO/TR 14121:** norm on the safety of the machinery used for the risk assessments
- **EN 12417:** specific on the machining centres
- **EN 14070:** safety of machine tools
- **EN/ISO 10218-1** and **EN/ISO 10218-2**: for robots and robotized systems
- **EN60204-**1: electric equipment of the machines
- **EN349:** that considers the risk of crushing
- **UNI 9910:** that considers the terminology on the trustworthiness and quality of service
- **ISO 14000 and 14001:** regarding environmental management
- **EN16001 (as the ISO 50001):** considering the Energy Management System

The list of specific terms used in the manufacturing environment is virtually endless, but in order to simplify this description it has been strictly correlated to the demonstrator described in "D5.4.1 Multi-sensory fusion and context awareness prototype 1", which will be shown at the M12 review meeting. It can be summarized as follows:



Production data

All information regarding the production process is stored in this section of the database. In particular the system is able to log any kind of faults, status of the machine and timers/counters.

- **Status:** each condition of the machine is classified as a status. In particular it is possible to define:
    - **Fault** that includes all those condition not expected to be caused by a defect or a break of the machine interrupting the normal production process.

  - **Stop,** which is a voluntary interruption of the production process not necessarily related to a fault of the machine, stop, tip change, timeout, missing load/unload, Manual/automatic mode, Cycle not started, Production
  - **Tip change:** the life of the electrodes of a welding gun is given by the number of welded points. In order to keep the tool in the proper working order, a counter rises an alarm when the end life of the electrodes has come. This operation stops the robot in a specific position and interrupts the flow of the water inside the cooling circuit in order to let maintenance replace the electrodes.
  - **Timeout:** for each operation performed during the production cycle a specific counter verifies if the time spent is more than expected. In this case a timeout message informs the operator of this particular condition that could evolve in a fault status.
  - **Missing load/unload:** this status informs that the machine is not able to work for a reason not depending on it.
    In the missing load a problem referred to the upstream line does not enable it to provide the elements to be worked/machined.
    In the missing unload the first station of the downstream line is full, so a fault that stopped the machine may have occurred.
  - **Manual/automatic mode:** during the production process, the line normally works in automatic mode. Sometimes in order to restore a fault status it is necessary to switch it in manual mode, fix manually the problem and then restart the process in automatic mode.
  - **Cycle not started**: switching from manual to automatic mode it is important to confirm this choice starting the production cycle.
  - **Production**: this condition is verified when the line is working properly in automatic mode without any faults.
- **Faults:** as described above those conditions not expected to be caused by a defect or a break of the machine interrupting the normal production process are included in this group. They can be related to a part of the line, a station or a single device. They can also be classified in faults caused by the production process or by the safety section. In this case a danger can compromise the safety of the worker.
  The difference with the general status of fault is that while the status highlights a problem on the machine, here a specific message clarifies what and where is really the malfunction.
- **Timers/counters:** during the production process it is extremely important to monitor the production trend. In order to achieve this objective several timers and counters are configured in the production system. The most important parameters monitored are the quantity elements produced, the number of scrap parts, and the cycle times, also known as takt times (term coming from the German language).

Energy data

At the moment in the existing production plants no data is available on the real-time monitoring of energy consumption, but the direction being taken by the market requires increasing information on this aspect. The opinion within the manufacturing suppliers is that it will become a key point in the future requirements of the customers.

Normally power consumption is monitored on a higher level without details for the single device or machine. In order to determine the real consumption of each device, these parameters could be considered:

- Compressed air used for producing one single element or a car
- Water used for cooling a welding gun and its transformer during the welding process of a single element, and the sum to calculate the energy used to produce a car
- Energy consumption of a robot or each device inside the plant
- Energy consumption of conditioning and lighting

In order to optimize the energy efficiency of the machines involved in the production process a document that is classified as Reference Document on Best Available Techniques for Energy Efficiency[60] has been developed by the European Commission.

On the "Energy data" side of the manufacturing environment the typical used terms [61] are:

- **Air conditioning:** Cooling and dehumidifying the air in an enclosed space by use of a refrigeration unit powered by electricity or natural gas. Note: Fans, blowers, and evaporative cooling systems ("swamp coolers") that are not connected to a refrigeration unit are excluded.
- **Central chiller:** Any centrally located air-conditioning system that produces chilled water in order to cool air. The chilled water or cold air is then distributed throughout the building, using pipes or air ducts or both. These systems are also commonly known as "chillers," "centrifugal chillers," "reciprocating chillers," or "absorption chillers." Chillers are generally located in or just outside the building they serve. Buildings receiving district chilled water are served by chillers located at central physical plants.
- **Circuit:** A conductor or a system of conductors through which electric current flows. Often a cooling system is also, improperly, called a circuit.
- **CO:** Carbon Monoxide
- **CO2:** Carbon Dioxide
- **Compressor station:** Any combination of equipment that supplies energy to move gas in transmission or distribution lines or into storage by increasing the pressure. Typically in the manufacturing plant compressed air is used in order to move clamps and fixtures inside the plants.
- **Connection:** The physical connection (e.g., transmission lines, transformers, gear switches, etc.) between two electric systems permitting the transfer of electric energy in one or both directions.
- **Cooling:** Conditioning of room air for human comfort by a refrigeration unit (such as an air conditioner or heat pump) or by circulating chilled water through a central cooling or district cooling system. Use of fans or blowers by themselves, without chilled air or water, is not included in this definition of cooling.
- **Current (electric):** A flow of electrons in an electrical conductor. The strength or rate of movement of the electricity is measured in amperes.
- **Electric current:** The flow of electric charge. The preferred unit of measure is the ampere.
- **Energy:** The capacity for doing work as measured by the capability of doing work (potential energy) or the conversion of this capability to motion (kinetic energy). Energy has several forms, some of which are easily convertible and can be changed to another form useful for work. Most of the world's convertible energy comes from fossil fuels that are burned to produce heat that is then used as a transfer medium to mechanical or other means in order to accomplish tasks. Electrical energy is usually measured in kilowatthours, while heat energy is usually measured in British thermal units (Btu).
- **Energy consumption:** The use of energy as a source of heat or power or as a raw material input to a manufacturing process.
- **Energy Efficiency:** A ratio of service provided to energy input (e.g., lumens to watts in the case of light bulbs). Services provided can include building sector end uses such as lighting, refrigeration, and heating: industrial processes; or vehicle transportation. Unlike conservation, which involves some reduction of service, energy efficiency provides energy reductions without sacrifice of service. May also

[60] http://eippcb.jrc.es/reference/ene.html
[61] http://www.eia.gov/tools/glossary

refer to the use of technology to reduce the energy needed for a given purpose or service.

- **Heating equipment:** Any equipment designed and/or specifically used for heating ambient air in an enclosed space. Common types of heating equipment include: central warm air furnace, heat pump, plug-in or built-in room heater, boiler for steam or hot water heating system, heating stove, and fireplace. Note: A cooking stove in a housing unit is sometimes reported as heating equipment, even though it was built for preparing food.
- **Humidity:** The moisture content of air. Relative humidity is the ratio of the amount of water vapour actually present in the air to the greatest amount possible at the same temperature.
- **Lamp:** A term generally used to describe artificial light. The term is often used when referring to a "bulb" or "tube."
- **Lighting equipment**: These are light bulbs used to light the building's interior, such as incandescent light bulbs, fluorescent light bulbs, compact fluorescent light bulbs, and high-intensity discharge (HID) lights.
- **Maintenance expenses:** That portion of operating expenses consisting of labour, materials, and other direct and indirect expenses incurred for preserving the operating efficiency and/or physical condition of utility plants used for power production, transmission, and distribution of energy.
- **Process cooling and refrigeration**: The direct process end use in which energy is used to lower the temperature of substances involved in the manufacturing process.

6. Agricultural Domain – Food Traceability

This chapter contains a description of data (which are currently stored in databases) on the Field, Feed Production, Farm, Slaughterhouse etc.

In the food production chain, data are stored in different databases according to the business structure. At farm level data are typically stored in local databases, and they are managed by pc based management systems. In feed production, at the slaughterhouse and in retail data are managed by centralized ERP systems.

Most data in the food traceability scenario are expected to be generated at farm level whereas the processing and distribution will add minor details to the history of the food products. A full agricultural vocabulary would of course be out of the scope for this deliverable. The following subsections will introduce data structures at the different levels in the food production chain. The items listed are the most common terms which are expected to be used in the data exchange model. Some of the terms are already defined by ISO, under Technical Committee 23[62] (TC 23) which handles standards for agriculture and forestry.

6.1 Farming

At farm level data are managed by locally installed applications which to some extent follow the standard ISO 11788. Management systems like AgroVision[63], PigChamp[64] and AgroCom[65] offer specially designed applications for pig, cattle and crop production, respectively. These management systems are typically used by the staff on the farm for registration of production and legislation data in the agricultural sector.

Besides serving as a vital tool for daily operations, these management systems also provide the farmer and his advisors with production reports containing production specific KPIs. Traditionally KPIs are calculated on a monthly or quarterly basis in animal production and on a yearly basis in crop production.

In sow and cattle production, data are normally related to an *animal number*. In piglet, finisher and chicken production, data are linked to *batch id*, hence in these groups an id on individual animals is not yet common. In crop production data are traditionally related to *field id*.

Relevant data in farm production is:

Farming production type [Organic farming, Traditional farming]

a) Crop production
   i) Working the soil [subset: ISO 8910:1993 Machinery and equipment for working the soil - vocabulary]
      (a) Plowing
      (b) Harrowing
      (c) Seeding
   ii) Crop protection [subset: ISO 5681:1992 Equipment for crop protection - vocabulary]
      (1) Spraying
         (a) Spraying product
         (b) Dose
   iii) Fertilizers [subset: ISO 8157:1984 Fertilizers and conditioners – vocabulary]
      (1) Nitrogen
      (2) Phosphorous

[62] http://www.iso.org/iso/iso_catalogue/catalogue_tc/
[63] www.agrovision.com
[64] www.pigchamp.com
[65] www.agrocom.com

(3) Potassium

iv) Harvesting [subset: ISO 11450:1999 Equipment for harvesting and conserving - Terminology and commercial specifications]

v) Crops [e.g. Barley, Corn, Wheat]

b) Animal production

i) Poultry

(1) Capons: [Castrated male birds]
(2) Chicks: [Young birds with down feathers]
(3) Hen: [Mature female chicken]
(4) Pullets: [Young immature birds]
(5) Rooster: [Mature male chicken]

ii) Cattle

(1) Bull: [Fertile male cattle over one year old]
(2) Calf: [Any young cattle less than one year old]
(3) Cow: [Mature female cattle that has given birth]
(4) Heifer: [A young beef or dairy female that has not given birth]
(5) Ox: [Mature, castrated bull]
(6) Steer: [An ox less than four years old]

iii) Pigs

(1) Barrow: [Young castrated male]
(2) Boar: [Fertile male]
(3) Gilt: [Young female pig]
(4) Hog, Finisher: [Full grown swine]
(5) Pig: [Young, immature swine]
(6) Piglet: [Baby pig]
(7) Sow: [Female]

iv) Events

(1) Medicine [Medication given to an animal or a group of animals]
(2) Mating, Service of animal [Making the animal pregnant]
(3) Calve, Farrowing [Birth of calf or piglet]
(4) Weaning [Moving the piglets away from the mother]
(5) Feeding [Feed given to an animal or a group of animals]
(6) Transfer [Moving an animal or a group of animals from one location to another]

6.2 Feed production

Feed is produced either on the farm or at a feed production plant. A given feed mix and its components are identified by a *batch number*, as illustrated in Figure 6.

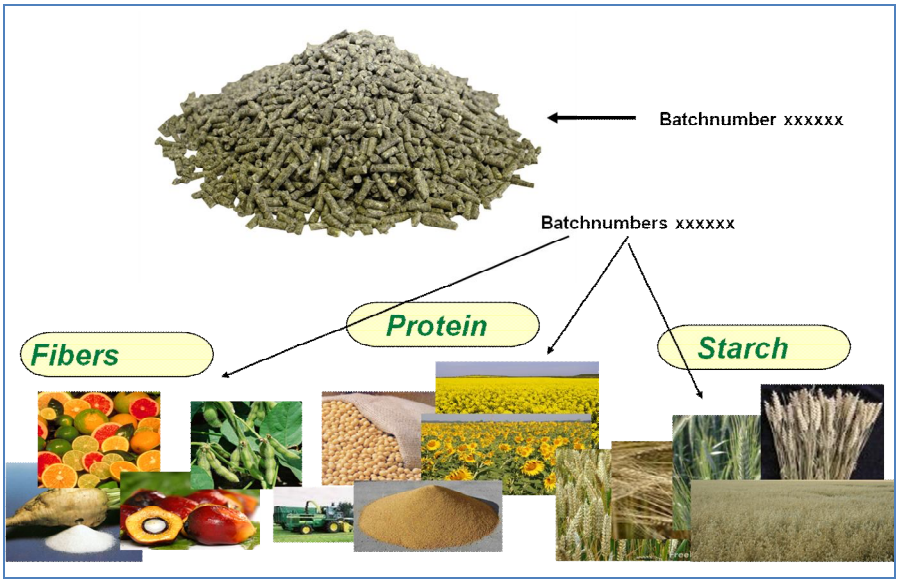


Figure 6: Feed is identified by batch number as well as all components

Feed producing companies are obliged to store information (at batch level) on all produced feed mixes. All feed mixes can be described by the batch ids of the components of the given mix.

Data relevant for ebbits is limited to the following list:

Feed production

a. Feed component [Name of a given feed component – could be an ingredient or a complete feed mix]
   i. Feed component number
   ii. Feed lot number
   iii. Feed type number
   iv. Feed component name
   v. Dry matter content % of feed component
   vi. Energy Type
   vii. Energy per kilogram
   viii. Crude protein per kilogram g
   ix. Phosphorus per kilogram g
   x. Calcium per kilogram g
   xi. Lysine per kilogram g
   xii. Methionine per kilogram g
   xiii. Threonine per kilogram g
   xiv. Tryptophan per kilogram g
   xv. Nitrogen per kilogram g
   xvi. Starch per kilogram g
   xvii. Sugar per kilogram g
   xviii. Copper per kilogram mg
   xix. Zink per kilogram mg
   xx. pH

xxi. Feed density kg/liter

b. Batch [A portion of a given feed component – e.g. a ship load of soya beans or one production of a feed mix]

c. Batch ID [A unique identifier of a batch]

d. Weight [weight of a given batch]

e. Location [Location of a given batch]

6.3 Slaughterhouse

At most slaughterhouses data are handled by a central ERP system (e.g. SAP). Slaughter pigs are delivered batch-wise, typically with a truck load of approx. 100 pigs coming from the same farm. Each pig is tattooed with a *supplier id* (id of the farm from where the animals come).

When a pig is entering the slaughter line, it is anesthetized with CO2 and put onto a hanging iron. This hanging iron is equipped with an RFID chip, and the supplier id is linked to the chip id. All data recorded during the first part of the slaughtering process (as long as the animal is on the hanging iron) is then stored with relation to the supplier id. Relevant data are:

Slaughterhouse products and machinery

a. Animals
    i. Type
        1. Pigs
        2. Cattle
    ii. Breed
    iii. Sex
    iv. Weight
    v. Meat Percentage
    vi. Price
    vii. Color
    viii. Age
    ix. Organic
b. Meat Products
    i. Weight
    ii. Meat Percentage
    iii. Price
    iv. Best before date
    v. Raw meat
        1. Whole carcass [Body of a slaughtered animal]
        2. Half carcass [Body of a slaughtered animal cut down lengthwise]
        3. Cut
            a. Head
            b. Spare Rib Roast
            c. Blade
            d. Hand
            e. Hock
            f. Trotters
            g. Loin
            h. Belly
            i. Leg/Ham
            j. Blade shoulder
            k. Arm shoulder
            l. Spare rib
            m. Side
    vi. Refined (examples…)
        1. Sausages
            a. Frankfurter
            b. Wiener
            c. Bratwurst
        2. Pâté
            a. French pâté
            b. Bacon pâté
c. Devices [Machines and tools used in slaughterhouses]

i. Storage holders: [Tool for storing and handling meat]
    1. Carcass gambrel: [Hook for large pieces]
    2. Crate: [Plastic boxes for smaller pieces]
ii. Machines [Machines used in slaughterhouses]
    1. CO2 anesthesia [Lift used to lower pigs into CO2]
    2. Cutting: [Machine specialized in cutting carcasses]
        a. Bung Dropper Machine: [Starts the incision of the carcass]
        b. Carcass Opener Machine: [Cuts open the belly]
        c. Evisceration Machine: [Removes the intestines]
iii. Cooler [Temperature]
    1. Cooling tunnel: [Starts the cooling of the fresh carcass]
    2. Equalizing cooler: [Obtains an even cool temperature in the carcass]

All data are stored in a database in the slaughterhouse, and is accessible for the suppliers (the farmers). The common interchange of these data is based on either traditional mail or on text files that can be downloaded from a homepage.

During the second part of the slaughtering process (the cutting), the pig is cut into pieces which are put onto an array of hooks, as illustrated in Figure 7.

Figure 7 Cutting and placing onto hooks

These hooks with meat go directly into trucks for distribution. One hook contains meat from 5-15 pigs and from that step in the process it is no longer possible to identify meat from individual animals.

Note: In relation to traceability, changes in processes are necessary to maintain animal specific identification of the meat.

6.4 Retailing

Data structures seem to be much more structured in retail compared to farming, feed production and slaughtering. As mentioned in chapter 3, several vocabularies describe the retail domain. One of the more widespread ones is the UNSPSC[66] standard. The vocabulary listed below which is mainly based on elements from UNSPSC v13.1201, is organized in a way that should satisfy the needs in the ebbits traceability scenario.

Food and Beverage Products

a. Meat and poultry products
    i. Minimally processed meat and poultry products
        1. Pork, minimally processed without additions
            a. Refrigeration state
            b. Shape/Form
            c. Boneless claim
            d. Source
            e. Quality
            f. Pork Cut
            g. Organic
            h. Religious law suitability claim
            i. Pack type
        2. Pork, minimally processed with additions
            a. Refrigeration state
            b. Shape/Form
            c. Boneless claim
            d. Source
            e. Quality
            f. Pork Cut
            g. Primary additive
            h. Organic
            i. Religious law suitability claim
            j. Pack type
    ii. Processed meat and poultry products
        1. Pork, processed without additions
            a. Refrigeration state
            b. Shape/Form
            c. Boneless Claim
            d. Source
            e. Level of Cooking
            f. Non-thermal Preservation
            g. Mechanical Processing
            h. Pork Cut
            i. Quality level
            j. Organic
            k. Religious law suitability claim
            l. Pack type
        2. Pork, processed with additions
            a. Refrigeration state
            b. Shape/Form
            c. Boneless Claim
            d. Source
            e. Level of Cooking
            f. Non-thermal Preservation
            g. Mechanical Processing
            h. Primary additive
            i. Pork Cut
            j. Quality level
            k. Organic
            l. Religious law suitability claim
            m. Pack type
b. Dairy products and eggs

[66] http://www.unspsc.org/

i. Milk and butter products
    1. Fresh milk or butter products
    2. Shelf stable milk or butter products
    3. Frozen milk and butter products
    4. Powdered milk
    5. Whey
    6. Lactose
ii. Cheese
    1. Natural cheese
    2. Processed cheese
    3. Imitation cheese
    4. Live plant and Animal Material and Accessories and supplies

c. Animal feed
    i. Livestock feed (examples...)
        1. Pure wheat bran
        2. Feed oats
        3. Feed corn
        4. Hay
        5. Compound feed

6.5 Business processes

To the knowledge of the authors of this deliverable, vocabularies covering business processes in the complete traceability chain do not exist. It is considered to be unmanageable to develop such a vocabulary from scratch at this stage of the project. Instead it is proposed to develop and extend this vocabulary in parallel with identification of the relevant processes during the project.

Table 1 illustrates the description of a few business processes related to delivery of pigs to the slaughterhouse.

Table 1: The delivery process [From farm to slaughterhouse]

Process	Purpose	Action
Notify next delivery	The farmer informs the slaughterhouse about the number of pigs he expects to deliver during the next 5 days	Staff at the farm identifies the approximate number of finishers above a given weight limit Farmer logs on to the home page of the slaughterhouse and enter the number of finishers he expects to deliver.
Selection for delivery	Selection of animals that have the size to be delivered.	Pigs are marked with color, based og weighing or the caretakers judgment Pigs that are marked are taken out of the pen and gathered in a delivery pen Pigs from the delivery pen are loaded to a truck Pigs are manually counted during loading
Delivery at the slaughterhouse	Transfer from truck to slaughter	Pigs are unloaded into a waiting area at the

			slaughterhouse Pigs are counted during unloading Pigs are anesthetized and put onto the hanging iron. Id of hanging iron is related to supplier id
	Information on delivered pigs	Slaughterhouse returns detailed information about the delivered pigs	Farmer downloads text file from slaughterhouse homepage Farmer import text file to his management system Farmer calculates KPIs on his production

7. References

(ASI 2011) American Society for Indexing. Retrieved from About Taxonomies and Controlled Vocabularies, from http://www.taxonomies-sig.org/about.htm.

(Ahmed et al. 2007) Ahmed, S., Kim, S., and Wallace, K. M.: A Methodology for Creating Ontologies in Engineering Design, Transactions of the ASME of Journal of Computing in Information Science in Engineering, 2007, Vol. 7(June), pp.132-140.

(Bennett 2007) Bennett, M.: Domain modelling tools, Hzpercube Limited, March 2007.

(Chang et al. 2010) Chang, X., Rai, R., Terpenny, J.: Development and utilization of ontologies in design for manufacturing. Journal of Mechanical Design 132(2) (2010).

(Christiansen et al., 2011) Christiansen, C. A., Kolodgy, C. J., Grady, J., Hudson, S., Domage, E., Chung, D., Chee, W., Avyar, D., Hanaoka, H., Hegde, N., Hochmuth, P., Xiao, A.: IDC's Worldwide Security Products Taxonomy, 2011, IDC.

(Fox *et al*., 1993) Fox, M., Chionglo, J.F., and Fadel, F.G.: A Common Sense Model of the Enterprise, in Proceedings of the 2nd Industrial Engineering Research Conference , pp. 425-429, Norcross GA: Institute for Industrial Engineers.

(Hepp 2006) Hepp, M.: Products and Services Ontologies: A Methodology for Deriving OWL Ontologies from Industrial Categorization Standards, in: International Journal on Semantic Web & Information Systems (IJSWIS), Vol. 2, No. 1, pp. 72-99, January-March 2006

(Intel 2010) IT@Intel Breif, Intel Cloud Computing Taxonomy and Ecosystem Analysis, Intel Information Technology, Cloud Computing, Feb., 2010.

(Meylor *et al.* 2009) Meylor, M., Zender, M., Fernandez, O.: Trying on Taxonomies: Research Developing Taxonomies for Analyzing and Developing Icon-Based Visual Language Systems, Rigor and Relevance in Design, Seoul, 2009.

(MIR 2001) Montague Institute Review. Retrieved from Manageing taxonomies strategically: http://www.montague.com/abstracts/taxonomy3.html.

(Mondorf and Herborn 2008) Mondorf, A. & Herborn, T. Bichler, M.; Hess, T.; Krcmar, H.; Lechner, U.; Matthes, F.; Picot, A.; Speitkamp, B. & Wolf, P. *(Eds.):* Ontology-based process mediation in the European project BRITE *Mulitkonferenz Wirtschaftsinformatik (MKWI 2008).*

(MWCD 2011) Merriam-Webster Collegiate Dictionary. Retrieved June 2011, from http://www.merriam-webster.com.

(Rebstock *et al.* 2008) Rebstock, M., Fengel, J., Paulheim, H.: *Ontologies-based Business Integration.* Springer Publishing company, Incorparated, 2008.

(Schlenoff et al. 1999) Schlenoff, C., Ivester, R., Libes, D., Denno, P., Szykman, S., Libes, D.: An Analysis of Existing Ontological Systems for Applications

in Manufacturing, Proc. ASME Design Engineering Technical Conference/Computers in Engineering.

(Schwarz 2005) Domain model enhanced search - A comparison of taxonomy, thesaurus and ontology. Thesis. Utrecht, Netherland.

(Zhang *et al.* 2004) Zhang, L., Liu, S., Pan, Y., and Yang, L.: InfoAnalyzer: A Computer-Aided Tool for Building Enterprise Taxonomies, Proc. Conference on Information and Knowledge Management, Washington, U.S.A., pp 477-483.