D14.3

Extended Virtual Factory Lessons Learned and Evaluations (1st issue)
<table>
<thead>
<tr>
<th>VERSION</th>
<th>DATE</th>
<th>NOTES AND COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>15/04/2015</td>
<td>First Draft</td>
</tr>
<tr>
<td>0.2</td>
<td>16/04/2015</td>
<td>Add Information about MoVA</td>
</tr>
<tr>
<td>0.3</td>
<td>20/04/2015</td>
<td>MoVA info refined</td>
</tr>
</tbody>
</table>
Table of Contents

EXECUTIVE SUMMARY ................................................................. 4

1  INTRODUCTION ................................................................. 5
  1.1 OBJECTIVES OF DELIVERABLE D14.3 .................................. 5
  1.2 STRUCTURE OF THE DELIVERABLE .................................. 5
  1.3 RELATIONS WITH OTHER WP’S ........................................ 5

2  GENERATION AND TRANSFORMATION OF VIRTUALIZED ASSETS (GETOVA) ............... 6
  2.1 SHORT OVERVIEW OF GETOVA SE .................................... 6
  2.2 EXPERIMENTS AND RESULTS ........................................... 7
    2.2.1 TANET ..................................................................... 7
    2.2.2 COMPlus .................................................................. 9
  2.3 LESSONS LEARNED ......................................................... 10

3  ADVANCED MANAGEMENT OF VIRTUALIZED ASSETS (MOVA) ........................................... 11
  3.1 SHORT OVERVIEW OF MOVA SE ....................................... 11
  3.2 EXPERIMENTS AND RESULTS .......................................... 11
    3.2.1 Data Modelling ......................................................... 11
    3.2.2 Cluster modelling ..................................................... 12
    3.2.3 Importing ............................................................... 13
    3.2.4 Cluster Search ......................................................... 14
  3.3 LESSONS LEARNED ......................................................... 16

4  CONCLUSION ........................................................................... 18
Executive Summary

This document reports on the lessons learned and evaluation results of the two FITMAN Specific Enablers (SE) generated in the context of FITMAN Task T14.1 namely:

- **advanced Management of Virtualized Assets (MoVA)** aiming to support Virtual Factories (VF) in intuitively generating, composing, and transforming virtual representations of in-/tangible assets (VAaaS) within Manufacturing Ecosystems. Design and implementation of an intuitive user-centric graphical interface for dynamic discovery and flexible composition of Virtualized in-/tangible Assets (as a Service) targeting at team building applications as well as advances in production networks;

- **Generation and Transformation of Virtualized Assets (GeToVA)** aiming to support Virtual Factories (VF) in semi-automatic generation and clustering of Virtualized intangible Assets (VAaaS) from real-world semi-structured enterprise and network resources. GeToVA enables as well multi-format ontology transformation between various representations of Virtualized in-/tangible Assets.

In the following, there is one chapter devoted to each of the SE implemented including:

- general information
- experiments done in the context of FITMAN Trials and their results
- lessons learned
1 Introduction

1.1 Objectives of Deliverable D14.3

The objective of this document is to report on the lessons learned and evaluation results of the FITMAN Specific Enablers (SE) generated in the context of FITMAN Task T14.1 specifically focusing extended virtual platform. The Description of T14.3 is cited here below for the reader convenience [FITMAN DOW]:

Task 14.3 EVF lessons learned and evaluations (resp. DiTF)

Task 14.3 will analyse and evaluate the new experimentations developed in T14.2 in the light of identifying lessons learned (bottlenecks and opportunities) for a further expansion of the selected VF trials in Phase III and in Trials exploitation actions. Feedback to the whole Fi PPP will be also provided both from a technical and from a business viewpoint.

1.2 Structure of the deliverable

Being an accompaniment document to the technical prototype there is one dedicated chapter for each of the SE implemented; each of which includes:
- general information
- experiments done in the context of FITMAN Trials and their results
- lessons learned

The conclusions provide an overview about how the SEs are implemented

1.3 Relations with other WP’s

Inputs
- Specific Enablers requirements come from WP1 and open call specification.

Implementation
- This deliverable has been created in cooperation with WP6.

Output
- The delivered Specific Enablers will be used in:
  - WP 6.6 Virtual Factory for Trials final edition
  - WP 14 piloting
2 Generation and Transformation of Virtualized Assets (GeToVa)

2.1 Short overview of GeToVA SE

As today manufacturing ecosystems deal with increasing quantities of unstructured and semi-structured information in webpages, e-mails, text documents, spreadsheets, news articles, collaborative posts, patents to name but a few, there is a real need to extract this information, to represent it in a meaningful, structured way, to cluster and transform it in multiple formats in order to support interoperability. The FITMAN Specific Enabler for Generation and Transformation of Virtualized Assets is aiming at providing a state-of-the-art Information Extraction-driven semantic tool for (semi-)automatic Virtualized intangible Assets in order to heavily reduce manual data entry for the population of the FITMAN-CAM Specific Enabler. The GeToVA Specific Enabler provides the following core functionalities:

1. Extraction of Virtualized Assets information from real-world semi-structured enterprise and network resource;
2. Generation of semantic representation of Virtualized intangible Assets according to ontological models;
3. Clustering of Virtualized intangible Assets enabling better search of such assets
4. Multi-format ontology transformation between various formats, mapping and exchanging Future Internet (FI) data e.g. USDL

The GeToVA Specific Enabler is provided as a set of RESTful services being implemented on top of the FITMAN baseline VF Platform. The GeToVA services APIs have been designed as fully compatible with FITMAN Platform components, namely the Data.SemanticsSupport for the GeToVA multi-formation ontology transformation and the Apps.Repository for registration of the assets generated by GeToVA.

The high-level architectural of FITMAN-GeToVA is depicted in Figure 1.

![Figure 1](image-url)
FITMAN-GeToVA Specific Enabler includes seven components which are briefly recapped below (a detailed description is available in D14.1 and D14.2).

GeToVA includes several components:

1. **Europass Format Handler** - used to manage structured data i.e. XML according to Europass Format
2. **Converter** - responsible for transforming the Base RDF formatted generated by the FormatHandler into various formats, according to various ontologies
3. **Clustering** - which provides clustering of Virtualized intangible Assets
4. **Knowledge Extractor** - responsible for extracting information from real-world semi-structured enterprise and network resource
5. **Ontology Manager** - used to create RDF data that is valid to the used ontologies within our platform.
6. **Tagging System** - allows the user to define annotations that can reused to automatically spot properties within semi-structured data
7. **Search** - provide Full Text search among our data
8. **Database and Search Engine** – for storing the raw and processed information
9. **RESTful API** - exposed in a unified RESTful API all the functionalities / components mentioned above.

### 2.2 Experiments and Results

GeToVA has been deployed and used in the context of two FITMAN Virtual Factory Trials, namely TANET and ComPlus as follows.

#### 2.2.1 TANET

In the context of the TANET trial GeToVA is used to import suppliers and tenders from unstructured and semi-structured sources (e.g. The Welsh Automotive Forum and Sell2Wales). The integration and usage of GeToVA SE in the TANET trial is illustrated in Figure 2.
Being integrated in the TANET trial, GeToVA provides the following functionalities. Given a set of suppliers and tenders GeToVA is able to semi-automatically extract information about these companies from unstructured and semi-structure data sources such as raw documents and web sites. The information extraction is performed using the Knowledge Extractor GeToVA component and then represented internally in GeToVA as RDF using the Ontology Manager component. In addition the information is transformed in other formats using the Converter component. The information in RDF is than imported into our sister SE, MoVA which offers additional functionalities for the SME Cluster in TANET (see Section 3 for more details).

GeToVA was also used to cluster, i.e. create groups of suppliers with similar profiles. A cluster created by the GeToVA Clustering component using TANET trial data is shown in Figure 3.
For the TANET trial we generated 143 tender opportunities from the Sell2Wales-website. An example of a generated virtual asset is listed below.

```
_:g2157029680 <http://fitman.sti2.at/company/hasLegalName> "CastAlum\n" .
_:g2157029680 <http://fitman.sti2.at/company/hasDescription> "Diecast and machined aluminium components, Design for manufacture\n" .
_:g2157029680 <http://fitman.sti2.at/company/hasWebsite> "www.Castalum.com\n" .
_:g2157029680 <http://fitman.sti2.at/company/hasLegalAddress> "Buttington Cross Enterprise Park\nWelshpool\n, Powys, SY21 8SL\n" .
_:g2157029680 <http://fitman.sti2.at/company/hasHQAдрес> "Powys, SY21 8SL\n" .
Benjamins-MacBook-Air:ditf benjaminhiltpolt$ cat 'Magor Designs .rdf'
_:g2157293240 <http://fitman.sti2.at/company/hasLegalName> "Magor Designs\n" .
_:g2157293240 <http://fitman.sti2.at/company/hasDescription> "Design Engineering, Precision Engineering\n" .
_:g2157293240 <http://fitman.sti2.at/company/hasWebsite> "www.magordesigns.co.uk\n" .
_:g2157293240 <http://fitman.sti2.at/company/hasLegalAddress> "Neath Vale Business Park\nResolven\n, Neath, SA11 4SR\n" .
_:g2157293240 <http://fitman.sti2.at/company/hasHQAдрес> "Neath, SA11 4SR\n" .
```

In addition we have generated 95 assets out of companies description in unstructured format. An example is given below.

```
_:g2181915040 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://fitman.sti2.at/company/Company> .
_:g2181254880 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://fitman.sti2.at/company/Company> .
_:g2181254880 <http://fitman.sti2.at/company/name> "Rumm Ltd" .
_:g2181254880 <http://fitman.sti2.at/company/hasLocality> " Mid Glamorgan" .
_:g2181254880 <http://fitman.sti2.at/company/country> "Wales" .
```
2.2.2 COMPlus

In the context of the COMPlus trial GeToVA is used for enrichment of the knowledge base use by the network manager. By having a richer knowledge base, the network transparency is improved and becomes easier for her/him to be aware of all possible choices of business partners and chose the most appropriate ones for their business network. The integration and usage of GeToVA SE in the COMPlus trial is illustrated in Figure 4.

![GeToVA in COMPlus trial](image)

**Figure 4: GeToVA in COMPlus trial**

Being integrated in the COMPlus trial, GeToVA provides the following functionalities. Given a set of company profiles GeToVA is able to semi-automatically extract information about these companies from unstructured and semi-structure data sources such as raw documents and web sites. The information extracted includes the company name, type, location, web site address, industry branch, etc. Such information is extracted using the Knowledge Extractor GeToVA component and then represented internally in GeToVA as RDF using the Ontology Manager component. In this way we can generate structured, semantic representations of the companies profiles which will be used by COMPlus to enrich the knowledge base. In addition the information is transformed in other formats using the Converter component. The information in RDF is than imported in the COMPlus ontological based where it can be queried and reasoned upon for the COMPlus improved network transparency case.
In total we have processed a total of 76 LED companies profiles and generated virtual assets from them using GeToVA functionality. An example of COMPlus LED company information generated by GeToVA is provided below.

```
_:g2173020640 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> 
<http://fitman.sti2.at/company/Company> .
_:g2173020640 <http://fitman.sti2.at/company/name> "spectral" .
_:g2173020640 <http://fitman.sti2.at/company/hasLocality> "Freiburg" .
_:g2173020640 <http://fitman.sti2.at/company/country> "Germany" .
_:g2173020640 <http://fitman.sti2.at/company/postalCode> "79111" .
_:g2173020640 <http://fitman.sti2.at/company/hasWebsite> 
"http://www.spectral-online.de" .
_:g2173020640 <http://fitman.sti2.at/company/hasMail> "info@spectral-online.de" .
_:g2173020640 <http://fitman.sti2.at/company/locatedInRegion> "Germany" .
_:g2173020640 <http://fitman.sti2.at/company/hasStreetAddress> 
"Bötzinger Straße 31" .
_:g2173020640 <http://fitman.sti2.at/company/produces> "Arbeits- und 
Leseleuchten" .
_:g2173020640 <http://fitman.sti2.at/company/produces> "LED-Strahler" .
_:g2173020640 <http://fitman.sti2.at/company/produces> "Lichtböden / -
decken" .
```

### 2.3 Lessons learned

When using GeToVA in the trials we faced several challenges. On one hand the diversity of information sources (unstructured, semi-structured and structured) had required adaptation of the GeToVA Knowledge Extraction component to be able to handle heterogeneous input. Luckily GATE, the tool we based our extraction component on provides some flexibility and we could use it straightforward in most case. For some however the additional code we develop for extraction had to be customized to cover all cases.

Using Text-based clustering to discover relations between company data has also proved sometimes a challenge. The lag of discovering any semantic relations between assets misses out some opportunities. Text-based Clustering is not leveraging the advantages the semantic environment has in which the assets are handled. In future work it would be nice to incorporate semantic clustering. Also for clustering large datasets are usually required.
3 Advanced Management of Virtualized Assets (MoVA)

3.1 Short overview of MoVA SE

MoVA has been deployed and used in the context TANET FITMAN Virtual Factory Trial. The architecture has already been described in previous reports. Thanks to the MoVA flexibility it could be applied without any problem.

3.2 Experiments and Results

3.2.1 Data Modelling

In the context of the TANET trial MoVA is used to identify new clusters (groupings of SMEs) responding to a new tender opportunity. The search needs to assure that the cluster fits in terms of competence and tender requirements. Therefore the following structure was modelled in MoVA.

![Figure 5 MoVA Startscreen and model](image-url)
3.2.2 Cluster modelling

Each asset allows for a certain capability, for example the ability to manufacture Exhausts Systems. This is defined by each supplier; as they know what they are able to do. The facilitator will estimate the capabilities of an asset based on significant factors such as Capability, Cost, Quality etc. This is important to include the tacit knowledge of the facilitator who knows best how the assets actually perform. Using natural language such as High, Medium, Low allows the facilitator to make easy judgements.

![Figure 6 MoVA Asset](image)

The domain entities provide the description of a tangible or intangible asset in terms of the capability. A domain entity may be a composite of other entities which are the requirement for the real life creation of parts or products. Example: an “Exhaust System” can either be produced as a whole by a supplier, or a network of companies with complementary capabilities can do it. Therefore it is necessary to divide the “Exhaust System” domain entity in sub-domains, e.g. Bracket, Pipe, Catalytic Converter and to add activities such as “Welding” and Time/Project Management. The composed entity has a 0-100% rating on the significance of each child entity, this is a simplified representation to allow non-technical staff to make “gut decisions” while building the entity tree. Normalisation (Sum = 100%) not needed. The following figure shows the decomposition of the domain entity “Exhaust System”.
3.2.3 Importing

Using the MoVA Backend and Plugin API the import was implemented. The Import Routines are implemented using the MoVA Plugin API and MoVA Backend API.

The plugin are integrated in the repository by creating a folder for the plugin in the repositories plugin directory. In the file register_application_components.php the following code activates the import code.

```php
$plugInPath = '../plugins/tanet/';

{ # Additional Functions
   $subPlugInPath = $plugInPath.'code/';
   { #
      $r->register_JavaScriptFile($subPlugInPath.'importSuppliers.js');
      $r->register_JavaScriptFile($subPlugInPath.'importTenders.js');
   }
}
```

The special import code for the two import functions is integrated in two JavaScript Files which defines the menu structure for calling the import which is running on the server. Each of the both import function has its own file with their own server side code.

```javascript
function importSuppliers() {
    [...]
}
```
function importTenders() {
    [...]
}

For importing the suppliers the JSON content from GeToVa URL http://fitman.sti2.at/tanet.json is read and parsed into an array. For each new content element there is created a new supplier in MoVA. This can be done very easy by using the MoVA Backend API. The values of the supplier can also be set using the MoVA API.

```
#Create new object
$O_Suppliers = $OT_Suppliers->addObject();

$O_Suppliers->setAttributeValue_noCheck( $OA_UUID_Suppliers_Supplier_name, "value_text", $companyName, true );

	if (property_exists( $importedAttributes, "c:hasDescription" ) ) {
        $O_Suppliers->setAttributeValue_noCheck( $OA_UUID_Suppliers_Description, "value_text", trim($importedAttributes->{"c:hasDescription"}), true );
    }

$addressParts = array();
	if (property_exists( $importedAttributes, "c:hasLegalAddress" ) ) {
        #ignore hasHQAddress as the data is the same as hasLegalAddress, but with less information
        $O_Suppliers->setAttributeValue_noCheck( $OA_UUID_Suppliers_Address, "value_text", trim($importedAttributes->{"c:hasLegalAddress"}), true );
        $addressParts[] = $importedAttributes->{"c:hasLegalAddress"};
    }

	if (property_exists( $importedAttributes, "c:country" ) ) {
        $O_Suppliers->setAttributeValue_noCheck( $OA_UUID_Suppliers_Country, "value_text", trim($importedAttributes->{"c:country"}), true );
    }
    else if (property_exists( $importedAttributes, "c:locatedInRegion" ) ) {
        $O_Suppliers->setAttributeValue_noCheck( $OA_UUID_Suppliers_Country, "value_text", trim($importedAttributes->{"c:locatedInRegion"}), true );
    }

	if (property_exists( $importedAttributes, "c:hasMail" ) ) {
        $O_Suppliers->setAttributeValue_noCheck( $OA_UUID_Suppliers_Email, "value_text", trim($importedAttributes->{"c:hasMail"}), true );
    }
    if (property_exists( $importedAttributes, "c:hasWebsite" ) ) {
        $O_Suppliers->setAttributeValue_noCheck( $OA_UUID_Suppliers_Website, "value_text", trim($importedAttributes->{"c:hasWebsite"}), true );
    }
```

### 3.2.4 Cluster Search

Using the MoVA Backend API and Plugin API the cluster search was implemented. The Cluster Search is started navigating to Requirements. Several Requirements are already imported.

The article which should be produced is set under Domain Keyword.
Next, the facilitator can specify the criteria regarding quality, capacity and cost when setting up a cluster search. Again, this brings the tacit knowledge and the experience of the facilitator into the game. He knows the tender opportunities and their focus: is high quality the most critical challenge or low price or a quick delivery time?

The cluster search has 3 levels, level 0, 1 and 2. In level 0 it searches for assets directly linked to the linked Domain Keywords. This means that a supplier can produce the whole piece the tender is asking for, no networking would be required. The result list is sorted by the best fitting asset. Internally the cluster search is working with the algorithm:

\[
O_{RequirementsAssetsBenchmark} = \text{requirements['O_RequirementCapacity']} \times O_{AssetCapacity} + \text{requirements['O_RequirementCost']} \times O_{AssetCost} + \text{requirements['O_RequirementQuality']} \times O_{AssetQuality};
\]

In level 1 and level 2 search the search algorithm first descends into the subdomains of the related domain and then searches for related assets. This means that the system is divided into subdomains and that the search looks for clusters that jointly cover all required subdomains.
The benchmark is calculated and then sorted by benchmark beginning with the best fitting result.

![Figure 10 MoVA: Result of the cluster search](image)

### 3.3 Lessons learned

Before you can start implementing the import routine you’ve to get familiar with the MoVA Plugin and Backend API. If you’ve done this orientation you can implement the import very easy.

The data provided by GeToVa for the suppliers is very good structured. So it was very easy to specify the import. Also the data for the Assets and Domain Entities seems to be good. The import could be done very easy. The GeToVa fields are mapped to the MoVA structure. By inspecting the assets and domain entities there are was found a lot of unstructured data, which GeToVa itself couldn’t filtered out, because the source data itself contains this wrong unstructured data. Therefore the imported data has to be inspected and validated by hand.

As only a few requirements contain a link to a Domain Entity a cluster search can only started for this small amount of requirements. The cluster search is working good and results suitable data. As the Domain Entity model is still not fully built, the cluster search is only working for level 0 and 1 at the moment.

The next steps are to clean up the Domain Entities from the wrong data and to build up the full domain hierarchy. After that the cluster search can be evaluated again to test the results. The MoVA SE proves to be very useful for the TANET trial as it combines hard facts (domain keywords, sub-domains) with the human way of working (searching for either high quality or quick delivery). MoVA is able to handle such fuzzy values and to integrate them into the technically exact model.
MoVA requires apparently a little experience in modelling complex information systems. The GUI fully supports the system and shows step by step what is happening. Hence, it supports the user in gaining this experience.
4 Conclusion

In this deliverable we have reported on the first round of experiments, results and lessons learned from using the two SEs developed as part T14.1 namely:

- **advanced Management of Virtualized Assets (MoVA)** aiming to support Virtual Factories (VF) in intuitively generating, composing, and transforming virtual representations of in-/tangible assets (VAaaS) within Manufacturing Ecosystems. Design and implementation of an intuitive user-centric graphical interface for dynamic discovery and flexible composition of Virtualized in-/tangible Assets (as a Service) targeting at team building applications as well as advances in production networks;

- **Generation and Transformation of Virtualized Assets (GeToVA)** aiming to support Virtual Factories (VF) in semi-automatic generation and clustering of Virtualized intangible Assets (VAaaS) from real-world semi-structured enterprise and network resources. GeToVa enables as well multi-format ontology transformation between various representations of Virtualized in-/tangible Assets.

The Specific Enablers have been incorporated into the FITMAN Architecture and have been used in the FITMAN pilots, especially in the Virtual Factories (WP6) that created the specification for the open calls to which these SEs fulfil.

Based on the lessons learned we are continuing to fine-tune the two SEs, to further experiment and evaluate them in the TANET and COMPlus trials.