This document contains final evaluation report and it is complementary to D4.1 (Requirements methodology) which defines the actual community use cases needed to capture the user requirements, D4.2 (Community Requirements Analysis Report) which covers all use cases discussed in detail, the actors and workflows for extracting the key performance indicators to evaluate the community, D4.3 (OpenMinTeD Functional Specifications), D4.4 (Community Evaluation Scenarios Definition Report) and D5.5 (Community Evaluation Methodology) which provides key performance indicators derived from the community use case evaluation scenario definition.

Topic: EINFRA-1-2014
Managing, preserving and computing with big research data
Research & Innovation action
Grant Agreement 654021
D4.5 - Community Evaluation Report

Document Description

D4.6 – Community Driven Evaluation Report

WP4 – Community Driven Requirements and Evaluation

WP participating organizations: ARC, University of Manchester, UKP-TUDA, INRA, EMBL, LIBER, OU, EPFL, BSC, USFD, GESIS, GRNET, Frontiers, UoG

Contractual Delivery Date: 31/03/2018  
Actual Delivery Date: 15/06/2018

Nature: Report  
Version: 2.1

Public Deliverable

Preparation slip

<table>
<thead>
<tr>
<th>Name</th>
<th>Organization</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>From</td>
<td>BSC</td>
<td>18/05/2018</td>
</tr>
<tr>
<td>Marta Villegas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martin Krallinger</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Montserrat Marimón</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edited by</td>
<td>BSC</td>
<td>18/05/2018</td>
</tr>
<tr>
<td>Marta Villegas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martin Krallinger</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Montserrat Marimón</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reviewed by</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Richard Eckart de Castilho</td>
<td>UKP</td>
<td>16/05/2018</td>
</tr>
<tr>
<td>Andrea Zielinski</td>
<td>GESIS</td>
<td>16/05/2018</td>
</tr>
<tr>
<td>Byron Georgantopoulos</td>
<td>GRNET</td>
<td>16/05/2018</td>
</tr>
<tr>
<td>Approved by</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Androniki Pavlidou</td>
<td>ARC</td>
<td>26/06/2018</td>
</tr>
<tr>
<td>For delivery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mike Hatzopoulos</td>
<td>ARC</td>
<td>26/06/2018</td>
</tr>
</tbody>
</table>

Document change record

<table>
<thead>
<tr>
<th>Issue</th>
<th>Item</th>
<th>Reason for Change</th>
<th>Author</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>V0.1</td>
<td>Draft version</td>
<td>Initial version</td>
<td>Marta Villegas</td>
<td>BSC</td>
</tr>
<tr>
<td>V1.0</td>
<td>First version</td>
<td>Reviewed</td>
<td>Montserrat Marimón</td>
<td>BSC</td>
</tr>
<tr>
<td>V2.0</td>
<td>Second version</td>
<td>Incorporated reviewers comments</td>
<td>Montserrat Marimón, Marta Villegas</td>
<td>BSC</td>
</tr>
<tr>
<td>V2.1</td>
<td>Final version</td>
<td>Finalised report</td>
<td>Marta Villegas</td>
<td>BSC</td>
</tr>
</tbody>
</table>
1. Table of Contents

1. INTRODUCTION ............................................................................................................................ 9

2. EVALUATION FRAMEWORK ...................................................................................................... 10

3. ANNOTATION PERFORMANCE .................................................................................................. 12

   LS-A - EXTRACT METABOLITES AND THEIR PROPERTIES AND MODES OF ACTIONS .................. 12
   LS-B - TEXT MINING FOR CURATION OF NEUROSCIENCE LITERATURE ................................ 13
   LS-C - TEXT MINING ON ARTICLES RELATED TO HEALTH STATE MODELLING IN CHRONIC LIVER DISEASES ................................................................. 13
   AS-A - TEXT MINING OVER BIBLIOGRAPHIC DATA: AGRIS & CORE .................................. 14
   AS-B - TEXT-MINING OVER RSS FEEDS: FOOD SAFETY & WATER HEALTH ..................... 15
   AS-C - MICROBIAL BIODIVERSITY ............................................................................................ 16
   AS-D - LINKING WHEAT DATA WITH LITERATURE .............................................................. 17
   AS-E - EXTRACTING GENE REGULATION NETWORKS INVOLVED IN SEED DEVELOPMENT (SeeDev) ............................................................. 17
   SS-A - FACILITATION OF COMPLEX INFORMATION LINKING AND RETRIEVAL FROM SOCIAL SCIENCES PUBLICATIONS ........................................ 18
   SC-A - RESEARCH ANALYTICS – I. FUNDING MINING SERVICES, II. DATAITE LINKING, III. DOCUMENT CLASSIFICATION ............................................ 19
   SC-B - RESEARCH PUBLICATIONS RECOMMENDATION SYSTEM AND SC-C - RESEARCH EXCELLENCE TRENDS EXPLORER .................................. 21
   SC-D - ROCK ART MINING ....................................................................................................... 21
   SC-E - TEXT MINING OF ARTICLES RELATED TO LEICA MICROSCOPES ............................ 22

4. QUALITY METRICS AND CHECKLISTS ...................................................................................... 25

5. FUNCTIONAL SPECIFICATIONS ................................................................................................ 36

   REGISTRY SERVICE .................................................................................................................. 38
   WORKFLOW SERVICE ............................................................................................................ 42

6. INTEROPERABILITY SPECIFICATIONS .................................................................................... 47

   INTEROPERABILITY GUIDELINES .......................................................................................... 47
   THE OMTD-SHARE METADATA SCHEMA .......................................................................... 50

7. THE OPEN TENDER CALLS AS AN EVALUATION OPPORTUNITY ......................................... 52

8. USE CASE FINAL REPORTS ...................................................................................................... 55

9. APPENDIX .................................................................................................................................. 58

   FUNCTIONAL REQUIREMENTS EVALUATION FORM - REGISTRY ............................................ 58
   FUNCTIONAL REQUIREMENTS EVALUATION FORM - WORKFLOW ........................................ 61
   GUIDELINES EVALUATION FORM .......................................................................................... 64
   OPEN CALL FINAL REPORT TEMPLATE .................................................................................. 66
   USE CASE EVALUATION REPORTS ....................................................................................... 68
Table of Figures

Figure 1 NLProt in the OpenEBench monitoring system .................................................................34
Figure 2 Summary of scores for functional requirements .................................................................41
Figure 3 Mandatory functional requirements (Registry) .................................................................41
Figure 4 Mandatory functional requirements + user requirements (Registry) ..................................42
Figure 5 Summary of scores for functional requirements (workflow editor) ....................................45
Figure 6 Mandatory functional requirements - workflow editor .....................................................45
Figure 7 Summary of scores for functional requirements (workflow execution) ..............................46
Figure 8 Mandatory functional requirements - workflow execution .............................................46
Figure 9 Guidelines evaluation by users .........................................................................................50

List of Tables

Table 1 Characteristics, Indicators and items (check lists) example ................................................28
Table 2 Use cases registry activities.................................................................................................37
Table 3 Registry functional requirements assessment .......................................................................40
Table 4 Summary of functional requirements scores .......................................................................41
Table 5 Functional requirements - workflow service ........................................................................44
Table 6 Summary of scores for mandatory functional requirements (workflow editor) ....................45
Table 7 Functional requirements’ scores - workflow editor ..............................................................46
Table 8 Guidelines content form .......................................................................................................48
Table 9 Guidelines assessment form ...............................................................................................49
Disclaimer

This document contains description of the OpenMinTeD project findings, work and products. Certain parts of it might be under partner Intellectual Property Right (IPR) rules so, prior to using its content please contact the consortium head for approval.

In case you believe that this document harms in any way IPR held by you as a person or as a representative of an entity, please do notify us immediately.

The authors of this document have taken any available measure in order for its content to be accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

This publication has been produced with the assistance of the European Union. The content of this publication is the sole responsibility of the OpenMinTeD consortium and can in no way be taken to reflect the views of the European Union.

The European Union is established in accordance with the Treaty on European Union (Maastricht). There are currently 28 Member States of the Union. It is based on the European Communities and the member states cooperation in the fields of Common Foreign and Security Policy and Justice and Home Affairs. The five main institutions of the European Union are the European Parliament, the Council of Ministers, the European Commission, the Court of Justice and the Court of Auditors. (http://europa.eu/)

OpenMinTeD is a project funded by the European Union (Grant Agreement No 654021).
Acronyms

API 
Application Programming Interface

AGRIS 
International System for Agricultural Science and Technology

API 
Application Programming Interface

BB 
Bacteria Biotope

BioNLP-ST 
BioNLP Shared Task

CDMI 
Component MetaData Infrastructure

ChEBI 
Chemical Entities of Biological Interest

CLARIN 
Common Language Resources and Technology Infrastructure

COCA 
Completeness, Operability, Correctness, and Appearance

EDAM 
EMBRACE Data and Methods

FAO 
Food and Agricultural Organization

FAQ 
Frequently Asked Questions

FS 
Functional Specification

GATE 
General Architecture for Text Engineering

IPR 
Intellectual Property Right

ISO/IEC 
International Organization for Standardization/International Electrotechnical Commission

KPI 
Key performance Indicator

LDC 
Linguistic Data Consortium

ML 
Machine Learning

NER 
Named Entity Recognition

NLP 
Natural Language Processing

OA 
Open Access

OMTD-SHARE 
OpenMinTeD schema

PDF 
Portable Document Format

RASFF 
Rapid Alert System for Food and Feed

RSS 
Really Simple Syndication

SeeDev 
Seed Development

SQuaRE 
Systems and software Quality Requirements and Evaluation
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TDM</td>
<td>Text and Data Mining</td>
</tr>
<tr>
<td>UIMA</td>
<td>Unstructured Information Management Architecture</td>
</tr>
<tr>
<td>WFED</td>
<td>Workflow Editing</td>
</tr>
<tr>
<td>WFEX</td>
<td>Workflow Execution</td>
</tr>
<tr>
<td>WS</td>
<td>Web service</td>
</tr>
<tr>
<td>XML</td>
<td>eXtensible Markup Language</td>
</tr>
</tbody>
</table>
Publishable Summary

This document is the sixth deliverable of Work Package 4 titled “Community Driven Evaluation Report” and provides the final evaluation report derived from the analysis of the assessment of key performance indicators previously defined by the community use cases and taking into account additionally aspects related to component interoperability as well as technical aspects underlying the use case derived text mining workflows. A central part of this evaluation task is the definition of an OpenMinTeD evaluation framework that is objective (based upon observation) and reproducible. The evaluation framework includes of a set of different components (evaluation checklists, questionnaires, form) that are attached at the end of this document in the appendix section.
1. Introduction

The OpenMinTeD project engages different thematic use case communities, namely Life Sciences, Agriculture/Biodiversity, Social Sciences and Scholarly Communication, with the aim to define real-life application scenarios which should be addressed through the OpenMinTeD infrastructure. These community use cases serve as guiding examples for the implementation and evaluation of the resulting text mining workflows. A general examination of formal aspects underlying the OpenMinTeD community use cases served as a base to define key performance indicators (KPIs) to be considered as evaluation criteria covered by the OpenMinTeD evaluation framework. KPIs refer to standard evaluation measures in case of the use case Gold Standard data and evaluation structured survey outcomes in case of the evaluation phase covering the OpenMinTeD infrastructure.

This document provides the final evaluation report derived from the analysis of the assessment of KPIs previously defined by the community use cases and taking into account additionally aspects related to component interoperability as well as technical aspects underlying the use case derived text mining workflows. A central part of this evaluation task is the definition of an OpenMinTeD Evaluation framework that is objective (based upon observation) and reproducible. The evaluation framework includes a set of different components (evaluation checklists, questionnaires, forms…) that are attached at the end of this document in the appendix section.

In order to facilitate a standardized assessment of some of the KPIs, an open community challenge shared task focusing on the technical interoperability and performance of specific components of the use case text mining workflows was carried out. The Open Call was an excellent opportunity to run the evaluation framework.

The content of this deliverable is complementary to D4.1 (Requirements methodology) which defines the actual community use cases, profiling users and building the questionnaires needed to capture the user requirements, D4.2 (Community Requirements Analysis Report) which covers all use cases discussed in detail, the actors and workflows for extracting the key performance indicators to evaluate the community, D4.3 (OpenMinTeD Functional Specifications), D4.4 (Community Evaluation Scenarios Definition Report) and D5.5 (Community Evaluation Methodology) which provides key performance indicators derived from the community use case evaluation scenario definition.
2. **Evaluation framework**

The OpenMinTeD evaluation was initially structured into two evaluation cycles (or phases) that were to be supported by the development of the OpenMinTeD evaluation framework (following a multi-step evaluation strategy). During each evaluation phase, a set of validation scenarios were defined to address the heterogeneity and particularities of (1) various community use cases on one side, and (2) end users or OpenMinTeD framework actors on the other side. For each use case, a set of evaluation criteria, derived from key indicators extracted from community use cases, were captured, analyzed and reported in the D4.5 “Community Driven Evaluation Methodology “document.

The goal of the evaluation process is to determine to which extent a given KPI is supported by the OpenMinTeD framework, and to build a use case centered evaluation infrastructure, the OpenMinTeD evaluation framework. The progress of the project and the delay in the development of the platform led us to change the methodology of the evaluation task. Components were finally deployed on the OMTD platform in later stages and, at that time, the different developer groups already started working on workflows. Consequently, the foreseen intrinsic component evaluation exercise, which was initially planned as a shared task, has been extended and (i) it included workflows (not only components), (ii) it extended evaluation to quality characteristics other than ‘accuracy’ (see below) and (iii) it focused in delivering an evaluation framework that should allow for reproducible and (hopefully) automatic evaluation tasks.

The final evaluation framework consists of a set of evaluation components addressing different aspects as listed below:

**Key Performance Indicators (KPIs):** KPIs refer to standard evaluation measures in case of the use case Gold Standard data and evaluation structured survey outcomes for evaluating the OpenMinTeD infrastructure.

- The standard evaluation measures were collected using the “Annotation evaluation performance form” (see the corresponding Appendix at the end of this document) and summarized in the section “Annotation Performance” in this document.
- The infrastructure evaluation task was performed during the phase II. In this case, we approached quality assessment focusing on three different aspects, namely: **quality metrics**, **functional specifications**, and **interoperability requirements**. For each aspect, we collected relevant input data using different methods as detailed below.

**Quality metrics:** following SO/IEC 9126-1 and ISO/IEC 25000 software evaluation standards, we defined an evaluation profile organized into different characteristics and sub-characteristics and implemented as a list of relevant indicators. This profile serves a dual purpose: to allow for objective, observable and reproducible evaluation and to provide a set of quality metrics. In the section “Quality metrics” of this document we give a detailed description.
Functional specifications: the functional specifications of the OpenMinTeD system were based on the user requirements gathered and documented on the deliverable D4.2 – Community Requirements Analysis Report. The functional specifications received corresponding priorities (“mandatory”, “important”, and “interesting”) and were grouped by service. We sent instructions and an evaluation form to the different use case partners to get information about functional specifications related to the registry and the workflow services. The form and the results are discussed in the section “functional specifications”.

Interoperability requirements: WP5 formalized the text mining infrastructure interoperability specifications and promoted them for adoption and exploitation outside the project’s boundaries. The overall objectives of the interoperability framework were to (i) identify and evaluate relevant standards, specifications, best practices for adoption in the design and implementation of the OpenMinTeD platform, (ii) identify gaps, set rules for the use of standards, specifications and best practices in APIs, data models and design patterns of the OpenMinTeD platform, (iii) produce specifications to support integration of 3rd party Text and Data Mining (TDM) artefacts with the OpenMinTeD platform and to (iv) deliver comprehensive guidelines that define and exemplify interoperability in the context of OpenMinTeD. In the section “Interoperability requirements”, we describe the evaluation task performed and the results.

Open Call (external users): The Open Call constituted an extremely useful scenario for evaluation and an excellent opportunity to get feedback and input from external users, essentially developers that were integrating their resources into the OMTD platforms following the interoperability guidelines. The final reports delivered by the participants give an overview of the main outcomes of the project, the unexpected issues, relevant experiences, etc.

In the following lines, we describe each of these evaluation tasks and related components in detail. At the end of the document, the Appendix section includes all questionnaires and forms used to collect the evaluation input data.

3. **Annotation Performance**

MS41 and MS53 documents already reported on the evaluation of the annotation performance of the community driven applications. Here we summarize and update the tasks done and give some details about the Annotation Performance form used to collect information.

We defined a questionnaire for the OMTD community use case evaluation to get information about Gold Standards produced/used and relevant prediction tasks performed. In order to get a complete picture of the different evaluation scenarios, the questionnaire collects information about different aspects including details about

- Criteria used for document selection (corpora)
- Annotation format
- Annotation editor used
- Annotation method (manual/Semi-automatic/Automatic)
- Tasks evaluated
- Evaluation scenario

The Appendix “Annotation performance evaluation form” contains the form used and the Appendix “Annotation performance evaluation responses” collects the responses we obtained. In the following subsections, we briefly describe the different evaluation scenarios reported and give some figures.

**LS-A - Extract metabolites and their properties and modes of actions**

The final ChEBI corpus annotated by the LS-A use case consists of 200 abstracts and 100 full text papers that have been annotated with named entities and relations from the biomedical domain as part of the OpenMinTeD project. This corpus facilitates the goal in OpenMinTeD of making text and data mining accessible to the users who need it most. The corpus was annotated with seven categories of entities (Metabolite, Chemical, Protein, Species, Biological Activity and Spectral Data) and four types of relations (Isolated From, Associated With, Binds With and Metabolite Of). The inter-annotator agreement (using F-score) for entities were between 0.796 and 0.892 using a strict matching protocol and between 0.875 and 0.963 using a relaxed matching protocol. For relations the inter-annotator agreement were between 0.591 and 0.693 using a strict matching protocol and between 0.744 and 0.793 using a relaxed matching protocol.

Details of the corpus are reported in the following paper:


---

2 For a complete report about use cases go to D9.4 deliverable, [http://openminted.eu/deliverables/](http://openminted.eu/deliverables/)
LS-B - Text mining for curation of neuroscience literature

Similar to LS-A, LS-B undertook a round of annotation for relevant entity types. The curation of neuroscience entities is crucial to ongoing efforts in neuroinformatics and computational neuroscience, such as those being deployed in the context of continuing large-scale brain modelling projects. However, manually sifting through thousands of articles for new information about modelled entities is a painstaking and low-reward task. LS-B engaged two computational neuroscientists to annotate a corpus of entities pertinent to neuroscience using active learning techniques to enable swift, targeted annotation and then trained machine learning models, which were tuned to recognize the entities that have been identified. We tested a traditional rule-based approach, a conditional random field and a model using deep learning named entity recognition, finding that the deep learning model was superior. Our final results show that we can detect a number of named entities of interest to the neuroscientist with a macro average precision, recall and F1 score of 0.866, 0.817 and 0.837 respectively.

This work is currently being prepared for publication in the journal Neuroinformatics. Full results and methodology will be shared upon publication.

LS-C - Text mining on Articles Related to Health State Modelling in Chronic Liver Diseases

Health State Modelling represents different health states enabling the provision of the right intervention for the right patient at the right time and dosage. Reference: https://f1000research.com/articles/6-309/v1.

Modeling of such health states should allow iterative optimization, as human data accumulate. The platform is intended to facilitate the discovery of links between opportunities related to a) the modernization of diagnosis, b) patient-centric approaches enabled by technology convergence, including digital health and connected devices, c) increasing understanding of the pathobiological, clinical and health economic aspects of disease progression stages, d) design of new interventions, including therapies as well as preventive measures, including sequential intervention approaches.

In OpenMinTED, the use-case LS-C is a tool to aid the definition of Chronic Diseases Health States. The goal is to extract Chronic Liver Diseases articles from the data repository to identify article-based health states. The first prototype focuses on Non-Alcoholic Fatty Liver Diseases and subsequent health states, such as Non-Alcoholic Steatohepatitis and End-state liver disease.

To annotate LS-C: Text Mining on Articles Related to Health State Modelling in Chronic Liver Diseases, the following semi-automatic methods were used:

The first step is to build a corpus of documents related to chronic liver diseases. This is achieved using the CORE Corpus builder tool. The output of the tool is a set of articles represented as TXT and JSON files. The text files contain the full text of the article as the JSON files contain the metadata. In our scenario, the corpus built contains 10K documents.
The second step is to annotate the documents. Three types of annotations are necessary; sentence, disease and progression keyword. This is done using the “Generic UIMA pipeline” tool. The tool is a standalone executable jar file. It takes as argument a directory containing raw text files, as well as a dictionary of items and a list of events. In our scenario, the dictionary consists of a list of liver disease built using a subset of the NCIT ontology. The keys of the dictionary are the disease including synonyms, as for the values they are the unique name of the disease. This is useful to have a single representation when different terms are matched. The events are a list of progression keywords. This list was built using n-gram over a list of sentences containing 2 diseases known to have a progression from one to the other.

Once the documents are annotated, it is time to build the aggregated report using the python scripts from the repository “Co-occurrences Graph Builder”. The first script extracts the annotations from the XMI files to JSON. This is done to decouple the generation of the report from XMI. The report could therefore be generated from different format of annotations. The JSON file contains a list of sentences, while each of them contains a list of 2 types of annotations; the items and the events.

Once the generic representation of documents is created, a second script has to be run to generate the co-occurrences. This script will aggregate the annotations of all the documents. It will loop over each sentence and build co-occurrences if the sentence contains at least 1 event and 2 different items (disease in this scenario). A sentence containing the following items, will generate the co-occurrences:

Item A [...] Item B [...] Item C [...]  
Item A -> Item B  
Item A -> Item C  
Item B -> Item C

For each co-occurrence, the reference to the sentence is added to a list. This is important to be able to validate the results by a domain expert. Once the process is completed, a table containing the number of times they appear in both directions is created. Then a threshold is applied. For example, if A -> B occurs twice more than B -> A. The current threshold is 0.7, which relates to 7 out of 10 or 7 to 3, which is a bit more than twice. The co-occurrences above the threshold are categorized as progressions. The table’s output as CSV and the user can manually edit the selected co-occurrences in case some actual progressions did not pass the threshold or in case non-progression did, which never occurred.

Finally, the last script is run. This script generates the single page application from the table of co-occurrences and the sentences stored in the pickle file.

A SME (Subject Matter Expert) from a pharmaceutical company and 3 healthcare practitioners evaluated and confirmed the Health State Model.

AS-A - Text mining over bibliographic data: AGRIS & CORE
The goal of the AS-A evaluation task was to manually check the quality of the PDF extraction module over a set of open access publications relevant to viticulture retrieved from the FAO AGRIS\(^3\) & CORE\(^4\) databases and provide insights for the calibration of the AgroVoc Extractor, and the Grape Varieties Extractor components during their development/customization. The evaluation covers the extracted terms accuracy and recall. The first phase of the evaluation provided input for the selection of the appropriate software components for the PDF extractor, and the calibration of the relevant term-based extractors.

The evaluation process used was as follows:

Four domain experts (two MSc holders and two PhD holders in Viticulture) were asked to list for a given set of 20 randomly selected Open Access papers (10 papers from AGRIS and 10 papers from CORE, when using the generic search term “vitis vinifera”, which is the scientific name of Grape Vine): (a) the recognized AgroVoc Terms, and (b) the Grape Varieties to which each paper refers.

The produced term lists were merged to create two testing term lists - one for each component (namely, the AgroVoc Extractor and Grape Varieties Extractor).

Different open-source PDF extractors were used to generate the textual information, over which the two extractors were applied. Depending on the area on which each term was located (abstract, paper body, table/figure caption), the PDF extractors were evaluated so that for each area the extractor with the best f-measure was selected.

It is worthy to mention that the best score (AgroVoc: 0.87 / Grape Varieties:0.91) was retrieved when examining the abstract section, whereas the worst score (AgroVoc: 0.62 / Grape Varieties: 0.79) was observed at table/figure captions.

The evaluation was performed with 2 groups with 15 domain experts for each group (postgraduates students from Agricultural University of Athens with Agricultural, Food and Nutrition background). The evaluation verified the conclusion of the first small-scale evaluation that the best score was retrieved when the extractors were examining the abstract section (AgroVoc: 0.82 / Grape Varieties:0.89). On the other hand, the worst score (AgroVoc: 0.60 / Grape Varieties: 0.74) was observed when the extractors were examining the captions from table and figure artifacts.

No gold standards were used as evaluation was performed over domain experts’ observations.

---

\(^3\) [http://agris.fao.org](http://agris.fao.org)

\(^4\) [https://core.ac.uk](https://core.ac.uk)
components during their development/customization. The evaluation covers the extracted terms accuracy and recall.

Similar to AS-A, the first phase of the evaluation provided input for the calibration of the relevant term-based extractors, and used the same evaluation process:

Four domain experts (two MSc holders and two PhD holders in Food and Nutrition) were asked to classify collaboratively a given set of 50 randomly selected Food Safety Alerts retrieved from RASFF’s (the European Rapid Alert System for Food and Feed) Consumer Portal\(^5\) to clusters of food recalls referring to products from the same geographical origin.

The extracted Geolocation terms for each feed were compared to the indicated (by the experts) cluster so as to compute the accuracy and recall of the extraction. The result of this evaluation guided the calibration of the term extraction modules so as to balance the trade-off between highly accurate results and inability to classify (recall in this case).

No gold standards were used as evaluation was performed over domain experts’ observations.

**AS-C - Microbial Biodiversity**

The text mining process of the Microbial Biodiversity use-case extracts information from scientific literature and database free text fields. Information extraction recognizes pieces of information of predefined specific types which are entities (i.e. terms that are of particular interest to this use-case) and relationships between these entities. We consider here three types of entities: microorganism, habitats and phenotypes; and two relationships: the “Lives_in” relation between a microorganism and its habitat(s) and the “Exhibits” relation between a microorganism and its phenotype(s). The evaluation that is reported here is the evaluation of the workflow prediction quality and usefulness.

The evaluation focuses on the quality of the prediction of these three entities and their two relationships.

The prediction of the entities and relations are evaluated by two ways:

- strict evaluation with a reference corpus. We used the framework of the Bacteria Biotope’16\(^6\) shared task of BioNLP-ST 2016 (denoted BB’16 in the following). The datasets and metrics\(^7\) are described in Deléger et al. (2016)\(^8\).

- evaluation of the potential of text-mining by comparing the automatically extracted information in a limited subject to the information of a review paper on a same subject.

---


\(^6\) [https://aclweb.org/anthology/W/W16/W16-3002.pdf](https://aclweb.org/anthology/W/W16/W16-3002.pdf)

\(^7\) [http://2016.bionlp-st.org/tasks/bb2](http://2016.bionlp-st.org/tasks/bb2)

Further details about the evaluation tasks can be found in the following article:


**AS-D - Linking Wheat Data with Literature**

The AS-D text-mining workflow automatically extracts and normalizes entities that are related to wheat genetics and phenotypes. These entities are Genes, Markers, Wheat phenotypes, Wheat-related species (taxa) and Wheat varieties.

AS-D checked the quality of the SAM blé corpus (Nédellec et al., MTSR 2014) to be used for the evaluation of the prediction of entities and their normalization.

A preliminary assessment has been carried out on the detection of two entity types on the subpart of the SAM blé corpus: genes and markers entities

- Gene detection: F1: 0.52; Recall: 0.571 ; Precision: 0.48
- Markers: F1: 0.80; Recall: 0.75; Precision 0.86
- Previous evaluation in Nédellec et al., MTSR 2014 for Phenotypes: Precision 95

Further details about the evaluation task can be found in the article:


**AS-E - Extracting gene regulation networks involved in seed development (SeeDev)**

The purpose of AS-E use-case is the extraction from scientific articles of the descriptions of genetic mechanisms involved in seed development of the model plant, Arabidopsis thaliana. This data has been integrated into the genome visualization application, FlagDB++.

The prediction of the entities and relations was evaluated using the framework of SeeDev’16 shared task of BioNLP-ST 2016 (datasets and metrics). The List of entities and relations is the one found in SeeDev.

The SeeDev corpus is used to evaluate entity and relation predictions. The entities and relations that are effectively implemented in the final application FlagDB++ are the same to a subpart of the SeeDev annotation model. We adapted the evaluation in two stages.

- We used the annotated entities of the SeeDev reference corpus the types of which are the same as the ones of the use case in order to compare the predicted entities to the reference entities. These types are Genes and Proteins.
The F1 scores for Genes and Proteins entities detection are 0.85 and 0.82 respectively. We have added to the corpus italic typography information that was not available in BioNLP-ST format of the corpus.

We evaluated the generic relationship,"Interaction” with the SeeDev corpus. As FlagDB++ contains only the Interaction relation, we had to merge the more specific relationships of the reference SeeDev corpus for alignment. We do so by aggregating five relationships of the SeeDev corpus: "Interaction, Binding, Regulation_Of_Accumulation, Regulation_Of_Expression and Regulation_Of_Molecule_Activity" that represent Interaction relations. The relevant arguments (Gene and Protein) of the relationships on both sides were normalized by the TDM workflow with the locus ID from TAIR. The resulting knowledge network is used as a reference. The performance results were

- Entity detection and categorisation and relation detection:
  - Precision: 0.23
  - Recall: 0.73
  - F-score: 0.34

The SeeDev corpus is part of the BioNLP-SP’16 shared task and can be found at: https://sites.google.com/site/bionlpst2016/tasks/seedev as well as the evaluation metrics description. [Chaix et al., 2016] reports on the detail of the knowledge model, the corpus figures and the task participant results.

Further details about the evaluation task can be found in the article:


SS-A - Facilitation of complex information linking and retrieval from social sciences publications.

The Social Science use cases cover the following tasks:

- Named Entity Recognition (NER) is the task of identifying textual mentions and classifying them into a predefined set of types. A Gold Standard Corpus for NER has been created using the WebAnno\(^9\) annotation tool. The NER resource is built out of 28 English and German articles from social sciences publications, annotated by two Social Science students. The named entity tags contain standard tags such as person and location and domain-specific tags that are particularly interesting within Social Sciences such as historical events. The tags used for annotation are PER (individuals, groups), LOC, ORG (scientific, governmental, parties, other), SUB (theoretical frames,
research methods), and OTH (Media, official texts, historical events) and described in more detail in the annotation guidelines. Both the dataset and the annotation guideline are publicly available in the project GitHub repository\textsuperscript{10}. The prediction of the entities according to the GermEval 2014 shared task\textsuperscript{11} with a small ML model (trained on 9 documents) and tested on 4 unseen documents yields the following performance: P: 0.203. R: 0.141, F-Measure: 0.167

- **Variable Detection and Linking** is the task of identifying all text spans that pertain to a set of predefined variables in scientific publications. A new Gold Standard Corpus for the task has been defined using the LDC Word Aligner. The Variable Detection and Linking resource has been created from 100 English and German social sciences publications, annotated by two Social Science students with an Inter-Annotator Agreement measured with Cohen’s Kappa of 80% on sentence level and 91.5% after a reconciliation among the annotators. The corpus and annotation procedure, annotation guidelines and baseline results are described in more detail in the references below. The annotation guidelines are publicly available in the project GitHub repository\textsuperscript{12}.

- **Keywords Assignment** is the task of extracting relevant keyword terms from a chosen vocabulary, i.e. selecting the most prominent words and phrases that appear in a document. To this end, the controlled vocabulary TheSoz, a thesaurus for the social sciences has been incorporated into the OpenMinTeD application. The predictions of the keyword assignment component have been evaluated against manually-specified keywords from TheSoz assigned by domain experts. An English dataset of 1,000 full text scientific papers in PDF, publicly shared by GESIS on the OpenMinTeD platform, is used as the gold standard for the task. Micro averaged results for all our experiments using the exact match evaluation metric and 10-fold cross-validation yield Precision: 20.67 +/- 15.38, Recall: 38.38 +/- 32.67 and F-Measure: 26.87.

Further details can be found in the following articles:


**SC-A - Research Analytics – i. Funding Mining Services, ii. DataCite Linking, iii. Document Classification**

---


\textsuperscript{11} https://sites.google.com/site/germeval2014ner/data

Three TDM research analytics applications have been integrated in OpenMinTeD:

I. The Funding Mining application mines the fulltext of research publications and extracts links to projects.

II. The DataCite linking application mines the fulltext of publications and extracts links to DataCite (https://www.datacite.org). These links are mainly citations to datasets.

III. The Document Classification application mines the fulltext of publications in order to perform content-based document classification using given taxonomies.

SC-A evaluated the funding and DataCite mining using the arXiv and Pubmed repositories (almost 2 million publications in total) and we validated manually the results. The mining results (funding, DataCite and classification) are also evaluated by community users since the algorithm runs in production in the OpenAIRE infrastructure where it text mines more than 4 million publications from various repositories.

The evaluation was done in two steps:

- The first step is offline and based on two datasets:
  - All the OA publications from arXiv (with publication dates between 2007 and 2016)
  - All the OA publications from PubMed (with publication dates between 2007 and 2016)
  - The results of the algorithm are manually validated using these datasets.

- The second step is online. The algorithm is integrated within OpenAIRE beta platform and mines more than 4 million publications that are available in OpenAIRE. These publications come from various data providers (PubMed, arXiv, JAIRO, DOAJ-Articles, CyberLeninka, LAREferencia, NARCIS, INRIA and more...). The online results of the algorithm for each funder are also manually validated before the integration with the OpenAIRE production platform. These datasets and the results of the algorithm are available through OpenAIRE API13.

The work carried on by the use cases about the generation of new gold standard corpora and annotation/evaluation tasks was reported in three recent publications (a few more are also expected):


13 http://api.openaire.eu
SC-B - Research Publications Recommendation System and SC-C - Research Excellence Trends Explorer

The use case SC-B undertook the task of producing a list of set of recommended articles for each given article in a well-defined and targeted corpus. A corpus was created with keywords regarding “Recommender systems” and filtered down by year (only year’s 2000 pubs), language (only English language was used), and publication type (only research articles was used) to produce a final small corpus of 97 articles.

The corpus was then given as input to the CORE Recommendation Web Service component and the output recommendations were stored for the later evaluation stage.

For evaluating recommendations of research articles, no ground truth (or gold standard) exists in our knowledge. Therefore, domain experts’ observations were used. The output recommendations for the small corpus was handed to 3 domain experts (Research Fellow and 2 PhD students) and asked to rate the overall quality of recommendations and indicate if there are obviously irrelevant recommendation pairs. While overall the quality of recommendation averaged very good to excellent relevance, it was reported that it lacked an element of serendipity i.e. it was unable to produce a novel possibility to the user.

The use case SC-C carried the task of enriching with analytical information a given corpus. Same as SC-B the same small corpus was created consisting of 97 articles. The corpus was then given to the citation count component (which based on Microsoft Academic Graph responds the citation count of a given article) and the outputs were stored for further evaluation.

From the given corpus, citation information (other than zero) was retrieved for 31 articles. The low recall can be justified as persistent identifiers (DOI) are missing for most articles so article identification and therefore effective querying is failing. For the retrieved dataset, a manual comparison with the number obtained from Google Scholar for the same article was carried out. It was observed that overall numbers are quite close though 2 cases of extreme difference were observed (e.g. for doi:10.1145/352871.352887 “Analysis of recommendation algorithms for e-commerce” Google scholar reports 2434 citations while the component reports only 1368).

SC-D - Rock Art Mining

This use case assists a researcher finding ancient artistic artefacts in the archaeological scientific literature. It is necessary to strictly mine scientific articles since the trust in the results is a critical factor. Artistic artefacts can be; rock paintings, petroglyphs, pictographs or engravings. The following properties must be extracted:

- Artefact type
- Site name
- Date
Dating method

Each result will be reviewed and added to a database of artefacts along with the article reference. This database is then used to consolidate the research theories.

To annotate SC-D: Rock Art Mining, the following semi-automatic methods were used:

The first step is to build a corpus of PDF documents related to the theme of research. In our use case a list of relevant PDF articles was given by the researcher. Then the “Table Extractor” tool is run over the PDF documents which extract the tables as CSV file, as well as images of the pages containing tables. The page of which the table appears on is stored in the file name. The region on which it is located on the page is stored in a JSON file.

The second step is to run the “Table Features Extractor” tool which will create a single CSV file indexing all the tables. The output file contains as columns different features indicating the number of columns, rows as well as Boolean columns indicating the presence of specific domain related words, such as paintings, petroglyphs, site and others. The tool will also classify using a simple heuristic whether an extracted table is an actual table and will only add the predicted ones to the index. This is done since the table extractor extracts a lot of tables which are not real tables, but rather paragraphs on multiple columns.

An initial prototype was created annotating the different entities using a rule-based approach. However, it was extremely hard to correlate those entities together. For example, to find the object a date was referring to, then finding the site of the object and eventually the method of the date. The accuracy of this approach was extremely low.

Therefore, another approach was suggested by the SME – Subject Matter Expert Researcher. Most of the results are stored as tables inside the PDFs. It was of great value for him to simply identify the tables and to be able to search them based on some keywords. This is the reason for the creation of those tools. It is now extremely easy to find the relevant tables and it drastically improved the speed of finding dated artefacts.

After discussions with some platform staff member, it seems that the platform is not meant to support tools doing extra activities than annotating XMI files. It however seems technically possible to integrate it. As of the time of writing this document, the tool is not wrapped in a docker image compatible with the platform, but further discussions can be carried out to maximize reusability.

SC-E - Text Mining of Articles related to Leica Microscopes

This use case demonstrates how Text and Data Mining tools and techniques can be used to count non-article based citations. The H2020 OpenUP (reference: http://openup-h2020.eu/), in its deliverable D5.5_Final-Report-on-Researcher-Impact.pdf identified the opportunity to increase the coverage of
impact metrics and to start measuring non-article based citations (i.e. data, software, cell lines, equipment, methodologies, theories).

Tools used as part of research are also a factor of their success. Identifying tools used in successful research could lead to better research. Such information is also very valuable to manufacturing companies not only to know what their most cited products are but as well when concurrent products are used instead. A better understanding of what is used could give a better insight of the demand and would therefore lead to better products.

Using Leica microscopes, this use case identifies citations of equipment in the methodology section of an article. Leica is manufacturer of optical microscope. Their products play a big role in the research community.

To annotate SC-E: Text Mining of Articles Related to Leica Microscopes, the following semi-automatic methods were used:

The first step is to create a corpus of articles containing the words; Leica and Microscope. The corpus was built through CORE API using the “CORE Corpus builder” tool. When the Leica Model annotator will be integrated on the platform, the corpus can be built directly on it. The final corpus contains 10K documents.

The next step is to annotate the documents. This is done using a UIMA pipeline integrating the Leica Model Annotator analysis engine. The annotator creates a single type of annotation highlighting the product models. Usually a very few number of annotations are generated per document.

The next step is to index all the annotations into a CSV file. This was achieved using a python script extracting the annotations from the XMI files and storing the covered text along the article id.

The final step is to generate a report. To have more accurate results, the annotations are normalized using a simple string normalization method. The report, a Jupyter notebook, contains the following insights:

1. Number of documents by number of citations
2. Most cited models
3. Correlation between top models and corpus coverage
4. Heatmap of models co-citations

The work carried on by the use cases about the generation of new gold standard corpora and annotation/evaluation tasks was reported in four recent publications (a few more are also expected):

4. **Quality metrics and checklists**

ISO/IEC 9126 is an international standard for the evaluation of software quality that distinguishes between three types of software quality: internal quality, external quality and quality in use. Internal quality is typically performed by developers and applied to the source code. External quality is measured in a simulated environment with simulated data using external metrics. Finally, quality in use measures the product when it is used in a specific environment and context.

The ISO/IEC 9126-1 model classifies internal and external quality in a structured set of characteristics and sub-characteristics as follows:

- **Functionality** - The capability of the software product to provide functions which meet stated and implied needs when the software is used under specified conditions.
  - **Suitability**: The capability of the software product to provide an appropriate set of functions for specified tasks and user objective. (Renamed as "functional appropriateness" in SQuaRE)
  - **Accuracy**: The capability of the software product to provide the right or agreed results or effects with the needed degree of precision.
  - **Interoperability**: The capability of the software product to interact with one or more specified systems. (promoted to characteristic in SQuaRE)
  - **Security**: The capability of the software product to protect information and data so that unauthorised persons or systems cannot read or modify them and authorised persons or systems are not denied access to them. (promoted to characteristic in SQuaRE)
  - **Functionality compliance**: The capability of the software product to adhere to standards, conventions or regulations in laws and similar prescriptions relating to functionality.

- **Reliability** - The capability of the software product to maintain a specified level of performance when used under specified conditions.
  - **Maturity**: The capability of the software product to avoid failure as a result of faults in the software
  - **Fault tolerance**: The capability of the software product to maintain a specified level of performance in cases of software faults or of infringement of its specified interface.
  - **Recoverability**: The capability of the software product to re-establish a specified level of performance and recover the data directly affected in the case of a failure.
  - **Reliability compliance**: The capability of the software product to adhere to standards, conventions or regulations relating to reliability.

---

14 The ISO/IEC 25000 (System and Software Quality Requirements and Evaluation) known as SQuaRE is an international standard that supersedes ISO/IEC 9126-1. The SQuaRE list of characteristics & subcharacteristics may be a bit different from the one listed here.
● **Usability** The capability of the software product to be understood, learned, used and attractive to the user, when used under specified conditions.
  ○ **Understandability**: The capability of the software product to enable the user to understand whether the software is suitable, and how it can be used for particular tasks and conditions of use.
  ○ **Learnability**: The capability of the software product to enable the user to learn its application.
  ○ **Operability**: The capability of the software product to enable the user to operate and control it
  ○ **Attractiveness**: The capability of the software product to be attractive to the user.
  ○ **Usability compliance**: The capability of the software product to adhere to standards, conventions, style guides or regulations relating to usability

● **Efficiency** - A set of attributes that bear on the relationship between the level of performance of the software and the amount of resources used, under stated conditions. The capability of the software product to provide appropriate performance, relative to the amount of resources used, under stated conditions.
  ○ **Time behaviour**: The capability of the software product to provide appropriate response and processing times and throughput rates when performing its function, under stated conditions.
  ○ **Resource utilization**: The capability of the software product to use appropriate amounts and types of resources when the software performs its function under stated conditions.
  ○ **Efficiency compliance**: The capability of the software product to adhere to standards or conventions relating to efficiency.

● **Maintainability**: The capability of the software product to be modified. Modifications may include corrections, improvements or adaptation of the software to changes in environment, and in requirements and functional specifications.
  ○ **Analyzeability**: The degree to which the software product can be diagnosed for deficiencies or causes of failures
  ○ **Changeability**: The capability of the software product to enable a specified modification to be implemented.
  ○ **Stability**: The capability of the software product to avoid unexpected effects from modifications of the software.
  ○ **Testability**: The capability of the software product to enable modified software to be validated.
  ○ **Maintainability compliance**: The capability of the software product to adhere to standards or conventions relating to maintainability.
Portability - A set of attributes that bear on the ability of software to be transferred from one environment to another.

- **Adaptability**: The capability of the software product to be adapted for different specified environments without applying actions or means other than those provided for this purpose for the software considered.
- **Installability**: The capability of the software product to be installed in a specified environment.
- **Co-existence**: The capability of the software product to co-exist with other independent software in a common environment sharing common resources.
- **Replaceability**: The capability of the software product to be used in place of another specified software product for the same purpose in the same environment.
- **Portability compliance**: The capability of the software product to adhere to standards or conventions relating to portability.

For quality in use, the model distinguishes four characteristics:

- **Effectiveness**: The capability of the software product to enable users to achieve specified goals with accuracy and completeness in a specified context of use.
- **Productivity**: The capability of the software product to enable users to expend appropriate amounts of resources in relation to the effectiveness achieved in a specified context of use.
- **Safety**: The capability of the software product to achieve acceptable levels of risk of harm to people, business, software, property or the environment in a specified context of use.
- **Satisfaction**: The capability of the software product to satisfy users in a specified context of use.

Based on this approach and with the aim of ensuring an objective (based upon observation) and reproducible evaluation task, we defined an evaluation profile used to compute and collect relevant quality characteristics. The profile is eventually implemented as a list of relevant indicators. An indicator in ISO 9126 allows measuring qualitative characteristics and it is defined as an aggregation of items. Items are questions about a product object, which are scored according to a counting rule. For example, assuming that *Documentation* is an indicator of the *Usability/Learnability* characteristics, we can easily map the interoperability specifications defined in OMTD as weighted items in a checklist as follows:
Checklists allow for objective and reproducible evaluation when dealing with qualitative characteristics. During the evaluation, the value of the items has to be determined and can be ranked. In the fictional example above, Documentation has a max score of 1 and each Item provides 0.25 points.

The evaluation profile suggested distinguishes up to 6 main (ISO based) characteristics:

1. Identification
2. Distribution
3. Functionality:
   a. Suitability
   b. Accuracy
   c. Interoperability
4. Usability, including
   a. Documentation
   b. Learnability
5. Portability, focused on Installability
The final checklist of items is collected in the table below. Items are grouped into (sub)characteristics and, for each one, we provide:

- the XPath to the corresponding element(s) in the OMTD-SHARE schema,
- a Boolean value that indicates whether the element is compulsory or not in the schema,
- the corresponding interoperability requirement and
- a tentative initial score for those items with an XPath value in the OMTD-SHARE schema. (Items aligned with as interoperability requirement are given 2 points whereas items with no mapping into any interoperability requirement are given 1 point.)

As discussed below, the fact that most of the items have a corresponding XPath information to the relevant element in the OMTD-SHARE schema will allow for a ‘quick’ and hopefully ‘automatic’ checking and evaluation task. Note, however, that some relevant characteristics/items are not included in the OMTD schema. This is especially critical for part of the Accuracy items (those referring to Gold Standards) and for input/output samples.

The Suitability items also do not have their equivalence in the metadata schema. In these cases, we assume that a user evaluation is required.

<table>
<thead>
<tr>
<th>IDENTIFICATION</th>
<th>ITEM</th>
<th>XPath in OMTD SCHEMA</th>
<th>compulsory?</th>
<th>REQ</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of component/workflow</td>
<td>/ms:componentInfo/ms:identificationInfo/ms:resourceNames</td>
<td>yes</td>
<td></td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>Main task</td>
<td>/ms:componentInfo/ms:componentType OR /ms:componentInfo/ms:applicationFunction</td>
<td>no</td>
<td>[REQ-8] Components should associate themselves with categories defined by the OpenMinTeD</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>URL to the component</td>
<td>/ms:componentInfo/ms:distributionInfos/ms:componentDistributionInfo/ms:downLoadURLs /ms:componentInfo/ms:distributionInfos/ms:componentDistributionInfo/ms:accessURL</td>
<td>no</td>
<td></td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>Description</td>
<td>/element(*,ms:componentInfoType)/ms:identificationInfo/ms:descriptions</td>
<td>yes</td>
<td></td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>Kind of component: (web app, installation, web service…)</td>
<td>/element(*,ms:componentInfoType)/ms:identificationInfo/ms:componentDistributionInfo/ms:componentDistributionInfo/ms:componentDistributionForm</td>
<td>yes</td>
<td>[REQ-8] Components should associate themselves with categories defined by the OpenMinTeD</td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>Does the component has a unique identifier?</td>
<td>/element(*,ms:componentInfoType)/ms:identificationInfo/ms:resourceIdentifiers</td>
<td>yes</td>
<td>[REQ-6] Components should have a unique identifier and a</td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>DISTRIBUTION</td>
<td>ITEM</td>
<td>XPath in OMTD SCHEMA</td>
<td>compul sory?</td>
<td>REQ</td>
<td>SCORE</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>----------------------</td>
<td>-------------</td>
<td>-----</td>
<td>-------</td>
</tr>
<tr>
<td>DISTRIBUTION</td>
<td>Does the component have licensing information?</td>
<td>/element(*,ms:componentInfoType)/ms:distributionInfos/ms:componentDistributionInfo/ms:rightsInfo/ms:licenseInfos</td>
<td>no</td>
<td>[REQ-33] Licensing information must be included in the metadata. [REQ-34] Licensing information should be expressed in a machine-readable form.</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Is the component defined as OpenAccess?</td>
<td>/element(*,ms:componentInfoType)/ms:distributionInfos/ms:componentDistributionInfo/ms:rightsInfo/ms:rightsStatement</td>
<td>yes</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does the component have clear attribution information?</td>
<td>ms:resourceCreationInfo/ms:resourceCreators/ms:resourceCreator</td>
<td>no</td>
<td>contactInfo = 2; landingPage = 1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Does the component have a contact person/point?</td>
<td>/element(*,ms:componentInfoType)/ms:contactInfo</td>
<td>no</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does the contact point include email info?</td>
<td>/element(*,ms:componentInfoType)/ms:contactInfo/ms:contactEmail</td>
<td>no</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does the component include citation instructions?</td>
<td>/element(*,ms:componentInfoType)/ms:distributionInfos/ms:componentDistributionInfo/ms:attributionTexts</td>
<td>no</td>
<td>[REQ-13] Citation information for component should be included in the metadata</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Does the component include information on funding?</td>
<td>/element(*,ms:componentInfoType)/ms:resourceCreationInfo/ms:fundingProjects</td>
<td>no</td>
<td>[REQ-47] Information on funding of resources may be included in the metadata.</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FUNCTIONALITY</th>
<th>ITEM</th>
<th>XPath in OMTD SCHEMA</th>
<th>compul sory?</th>
<th>REQ</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suitability</td>
<td>Does any mismatch exist among the component and its description (behavior, input, output, ...)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Could the component be improved in some way?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Can the component be used in the scenarios for which it was designed?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Could constraints exists for any situation (extreme or not)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>Is the component marked as 'evaluated'?</td>
<td>/element(*,ms:componentEvaluationInfoType)/ms:evaluation</td>
<td>no</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does the component includes intrinsic evaluation info?</td>
<td>/element(*,ms:componentEvaluationInfoType)/ms:evaluationTypes</td>
<td>no</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Component</td>
<td>Question</td>
<td>XPath in OMTD SCHEMA</td>
<td>Compulsory</td>
<td>REQ</td>
<td>Score</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>-----------------------</td>
<td>------------</td>
<td>-----</td>
<td>-------</td>
</tr>
<tr>
<td>Does the component includes extrinsic evaluation info?</td>
<td>/element(*,ms:component EvaluationInfoType)/ms:evaluationTypes</td>
<td>no</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does the component provides any evaluation report?</td>
<td>/element(*,ms:component EvaluationInfoType)/ms:evaluationReports</td>
<td>no</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is there any Gold Standard ready? (URL to the Gold Standard)</td>
<td>Describe the gold standard used for the evaluation of the component</td>
<td>Precision</td>
<td>Recall</td>
<td>F-score</td>
<td></td>
</tr>
<tr>
<td>Interoperability</td>
<td>Does the component uses standard input format?</td>
<td>/ms:componentInfo/ms:inputContentResourceInfo/ms:dataFormats</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does the component uses standard output format?</td>
<td>/ms:componentInfo/ms:outputContentResourceInfo/ms:dataFormats</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does the component declare the type system used in the annotation?</td>
<td>/ms:componentInfo/ms:inputResourceInfo/ms:annotationSchema</td>
<td>no</td>
<td>[REQ-9] Components should declare their annotation schema dependencies</td>
<td>2 (by rule)</td>
</tr>
<tr>
<td></td>
<td>Can the component consume input data delivered by OMTD platform in a straightforward way?</td>
<td>/element(*,ms:processingResourceInfoType)/ms:annotationSchema</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does the component declares the input's annotation schema?</td>
<td>/ms:componentInfo/ms:inputResourceInfo/ms:annotationSchema</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does the component declare language restrictions?</td>
<td>/ms:componentInfo/ms:inputResourceInfo/ms:languages</td>
<td>[REQ-43] S/W (tools, web services, workflows) must indicate whether they are language-independent or the language(s) of the resources they take as input and output</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Previous component in pipeline (if any):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Next component in pipeline (if any):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**USABILITY**

<table>
<thead>
<tr>
<th>ITEM</th>
<th>XPath in OMTD SCHEMA</th>
<th>Compulsory?</th>
<th>REQ</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Documentation</td>
<td>Does the component include a link to some documentation?</td>
<td>/ms:componentInfo/ms:resourceDocumentationInfo/ms:documentationPublications/ms:isDocumentedIn</td>
<td>no</td>
<td>[REQ-12] Components should provide documentation describing their functionality</td>
</tr>
<tr>
<td>Documentation?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>------------------</td>
<td>------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Documentation references should be versioned</td>
<td>[REQ-50] Documentation references should be versioned</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does the component includes info about the framework it belongs to?</td>
<td>/element(*,ms:componentInfoType)/ms:componentCreationInfo/ms:framework</td>
<td>yes</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Does the component declare the type of method used to perform its task?</td>
<td>/element(*,ms:componentInfoType)/ms:componentCreationInfo/ms:TDMMethod</td>
<td>no</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Learnability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does the component declares the implementation programming languages needed for allowing user contributions, or for running the tools?</td>
<td>/element(*,ms:componentInfoType)/ms:componentCreationInfo/ms:implementationLanguage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does the components have a landing page?</td>
<td>/element(*,ms:componentInfoType)/ms:contactInfo/ms:landingPage</td>
<td>yes</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Does the component includes any mailing list?</td>
<td>/ms:contactInfo/ms:mailingLists</td>
<td>no</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Does the component include training material?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does the component include information about input format?</td>
<td>/ms:componentInfo/ms:inputContentResourceInfo/ms:dataFormats</td>
<td>no</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Does the component include input samples and/or examples?</td>
<td>/ms:componentInfo/ms:inputContentResourceInfo/ms:characterEncodings</td>
<td>no</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>URL of the sample input</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does the component include information about output format?</td>
<td>/ms:componentInfo/ms:outputResourceInfo/ms:dataFormats</td>
<td>no</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Does the component include output samples and/or examples?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does the component includes info about output character encoding?</td>
<td>/ms:componentInfo/ms:outputResourceInfo/ms:characterEncodings</td>
<td>no</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>URL of the sample output</td>
<td>/element(*,ms:componentInfoType)/ms:resourceDocumentationInfo/ms:samplesLocations</td>
<td>no</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Does the component include descriptions of intended use cases.</td>
<td>/element(*,ms:componentInfoType)/ms:usageInfo</td>
<td>no</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Does the component include a link to some issue tracker system?</td>
<td>/element(*,ms:componentInfoType)/ms:resourceDocumentationInfo/ms:issueTracker</td>
<td>no</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>Does the component includes an online Help URL?</td>
<td>/ms:componentInfo/ms:resourceDocumentationInfo/ms:onlineHelpURL</td>
<td>no</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Does the component include access mode information?</td>
<td>/element(<em>,ms:componentInfoType)/ms:distributionInfos/ms:componentDistributionInfo/ms:downloadURLs /element(</em>,ms:componentInfoType)/ms:distributionInfos/ms:componentDistributionInfo/ms:accessURLs</td>
<td>no</td>
<td>[REQ-38] Access mode of resources must be included in the metadata (e.g. for s/w, whether they are downloadable or accessible as web services/workflows.</td>
<td></td>
</tr>
<tr>
<td>Are the parameters defined?</td>
<td>/ms:componentInfo/ms:inputContentResourceInfo/ms:parameterInfos</td>
<td>[REQ-21] Configuration and parameterizable options of the components should be identified and documented</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**PORTABILITY**

<table>
<thead>
<tr>
<th>ITEM</th>
<th>XPath in OMTD SCHEMA</th>
<th>compulsory?</th>
<th>REQ</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does the component declares any resource dependency?</td>
<td>/element(*,ms:componentInfoType)/ms:componentDependencies/ms:annotationResources</td>
<td>no</td>
<td>[REQ-9] Components should declare their annotation schema dependencies</td>
<td></td>
</tr>
<tr>
<td>Does the component declare its software library dependencies?</td>
<td>/element(*,ms:componentInfoType)/ms:componentDependencies/ms:softwareLibraries</td>
<td>no</td>
<td>[REQ-9] Components should declare their annotation schema dependencies</td>
<td></td>
</tr>
<tr>
<td>Does the component have version info?</td>
<td>/element(*,ms:componentInfoType)/ms:versionInfo</td>
<td>yes</td>
<td>[REQ-9] Components should declare their annotation schema dependencies</td>
<td></td>
</tr>
<tr>
<td>Is the component downloadable?</td>
<td>/element(<em>,ms:componentInfoType)/ms:distributionInfos/ms:componentDistributionInfo/ms:downloadURLs /element(</em>,ms:componentInfoType)/ms:distributionInfos/ms:componentDistributionInfo/ms:accessURLs /element(*,ms:componentInfoType)/ms:distributionInfos/ms:componentDistributionInfo/ms:componentDistributionForm</td>
<td>[REQ-28] Processing components should be downloadable (Components that are not downloadable should clearly indicate that (e.g. using a &quot;I am a service&quot; flag).</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OpenMinTeD has been collaborating with the new ELIXIR\textsuperscript{15} Benchmarking platform OpenEBench in the mapping between the different schemas. This OMTD-OpenEBench mapping will allow loading OMTD

\textsuperscript{15} \url{https://www.elixir-europe.org/}

Page 33 of 92
metadata into the technical monitoring platform. OpenEBench is a software infrastructure for Benchmarking and Technical Monitoring of bioinformatics tools, web servers and algorithms. One of the objectives of OpenEBench is to establish and refine communication protocols with communities and/or infrastructure projects. Very briefly, OpenEBench

1. imports metadata data from different registries (currently Biotools and Galaxy)
2. maps the data into its own model
3. exposes mapped data and gives quality metrics based on
   a. the metadata imported plus
   b. metadata extracted from other sources (github, pubmed, ...)

The figure below is a screenshot of OpenEBench in which we can see the analysis generated for the NLProt\(^\text{16}\) tool. Currently, the technical monitoring system measures the installability, accessibility, licensing and documentation characteristics of the tool and, additionally, provides citation references.

---

\(^{16}\) NLProt is a tool for finding protein-names in natural language-text, [https://rostlab.org/owiki/index.php/NLProt](https://rostlab.org/owiki/index.php/NLProt)
We are currently working with the OpenEBench team to automatically import metadata from OMTD registry, map the records into the OpenEBench model and compute the technical metrics. The OMTD registry API\textsuperscript{17} allows remote access to the metadata records in the registry.

\textsuperscript{17} https://services.openminted.eu/api/swagger-ui.html
5. Functional specifications

The goal of the functional specifications reported in D4.3 was to identify and define the (main) functionalities that the system must cover - that is, what the system is supposed to do at the end of the project - in order to satisfy the user’s needs. In this section, we report about the eventual evaluation task performed by the use cases. For this community driven evaluation task, we focused on the registry and workflow services described as follows:

- **Registry Service**: Validate the support to discover tools & services, scientific content, language resources within the OMTD registry service. A full documentation and registration mechanism for all types of resources available in the infrastructure, implementing widely agreed and used metadata models, covering all facets of documentation from persistent identification and versioning, technical specifications and software dependencies, rights of use, location and deployment instructions.

- **Workflow Service**: Provides users with the ability to mix and match text-mining services in workflows, i.e., by using the best of breed of text mining components for some task. Its implementation and use delves into *component level interoperability*.

The various use cases were asked to register their applications and components and:

1. Create a workflow including at least an input corpus and one of their apps/comps. They were allowed to use available corpora on the platform or upload a sample corpus by themselves.
2. Run the workflow on the testing platform.
3. Check if they were able to see the output of the workflow as their expectation.
4. In case they had any issues, they were asked to report to the technical team so that they can fix it.

Use cases reported details about their experience including: a short description of the app/components registered together with input/output details; the eventual registration status; link to the sample input used for running and testing the components (used for replication testings) and the link to the component on the testing platform. The table below collects all this information.

<table>
<thead>
<tr>
<th>Use Case</th>
<th>Component Name</th>
<th>Notes / issues</th>
<th>Component type</th>
</tr>
</thead>
<tbody>
<tr>
<td>LS-A</td>
<td>ChEBI web service</td>
<td>Successfully registered.</td>
<td>WS: <a href="http://nactem.ac.uk/api/openminted/chebi">http://nactem.ac.uk/api/openminted/chebi</a></td>
</tr>
<tr>
<td></td>
<td>Input: plain text with CAS information</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Output: text with annotated entities: Chemicals, Species, Proteins, and Metabolites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LS-B</td>
<td>Neuroscience web service</td>
<td>Successfully registered.</td>
<td>WS: <a href="http://nactem.ac.uk/api/openminted/neuroscience">http://nactem.ac.uk/api/openminted/neuroscience</a></td>
</tr>
<tr>
<td></td>
<td>Input: plain text with CAS information</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### AS-A
**Grape Varieties Extractor**
- **Input:** Plain Text Document
- **Output:** Grapevine Terms + Frequency in XML Format

Successfully registered.

DOCKER

### AS-B
**Geonames Extractor**
- **Input:** Plain Text Document
- **Output:** Geonames Terms + Frequency in XML Format

Successfully registered.

DOCKER

**Geopolitical Extractor**
- **Input:** Plain Text Document
- **Output:** Geopolitical Terms + Frequency in XML Format

### AS-C
**Habitat-Phenotype Relation Extractor for Microbes**
- **Input:** xmi
- **Output:** xmi

Component have been successfully registered

DOCKER

### AS-D
**Wheat Phenotypic Information Extractor Application**
- **Input:** Corpus (XML format or Web of Science text format)
- **Output:** Tagged text (tab-delimited format)

Component have been successfully registered

DOCKER

### AS-E
**Arabidopsis Gene Regulation Extractor**
- **Input:** xmi
- **Output:** xmi

Component have been successfully registered

DOCKER

### SS-A
**NER Component**
- **Variable Detection Component**
- **Keyword Assignment Component**

Successfully registered.

UIMA

### SC-A
**Research Analytics**
1. **Funding Mining Services**
2. **DataCite Linking**
3. **Document Classification**

Successfully registered

DOCKER

### SC-B
**Research Publications Recommendation System**
1. **CORE Research Publication recommender web service**
2. **Grobid Extractor**

Components registered successfully

DOCKER + Web Service

### SC-C
**Research Excellence Trends Explorer**
1. **Citation-count web service**
2. **ScienceParse extractor**

Components registered successfully

DOCKER + Web Service

### SC-D
**Rock Art Mining**

Components have been successfully registered

DOCKER

### SC-E
**Text Mining of Articles related to Leica Microscopes**

Components have been successfully registered

DOCKER

*Table 2 Use cases registry activities*
All use cases were able to successfully register their components and correctly provided sample corpora for testing purposes. Note, however, that AS-C and AS-E reported on some execution issues and AS-D had space memory problems.

Use cases were also required to evaluate the functionality requirements for the registry and workflow services. Their experiences were collected by means of a structured survey with four evaluation levels:

- **Full** - fully compliant
- **Part** - partially compliant. E.g., some parts of a product are compliant but not all. This is typically the case if a product is in a state of transition from a non-compliant to a compliant state.
- **No** - not compliant.
- **NA** - not applicable. This is expected to occur mainly for concrete requirements if a certain requirement is not applicable for a certain implementation, e.g. a requirement on remote API access on a tool which does not offer a remote API.

The results of this evaluation are reported in the next two sections.

**Registry service**

The following table lists the results for the structured survey we run for the evaluation of the functional requirements of the registry service. Following the naming conventions used in the D4.3 document, the table includes

- the functional specification ID prefixed with “FS” followed by an acronym of the corresponding service (e.g. “REG” for Registry) and by an incremental number,
- the description,
- the functional specification priority, that is, the level of importance of the specification, which can take the following 3 values “Mandatory”, “Important” and “Recommended” and
- the user requirements involved in each case. Note that most functional specifications cover one or more user requirements. In these cases, the user requirements are also listed and scored.
- Functional specifications have orange background in the table whereas related user requirements have white background.
- Finally note that the assessments of some use cases are grouped, this happens when the use case was developed by the same partner.

---

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FS/REG/01</td>
<td>Maintain a registry of text mining components</td>
<td>Mandatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cf.PL/REG/26</td>
<td>Users can register text mining components</td>
<td>Mandatory</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td>Cf.PL/REG/01</td>
<td>Update and manage a registry of text-mining services/components (User can update an already registered component)</td>
<td>Mandatory</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td>Partial</td>
<td>NA</td>
<td>Partial</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td>Cf.PL/REG/05</td>
<td>Users can view components characteristics</td>
<td>Mandatory</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td>Full</td>
<td>Full</td>
<td>NA</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td></td>
</tr>
<tr>
<td>FS/REG/04</td>
<td>Maintain a registry of corpora</td>
<td>Mandatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cf.PL/REG/09</td>
<td>Users can retrieve various content resources through a single query</td>
<td>Important</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td>Cf.PL/REG/09-a</td>
<td>Users can apply faceted search on textual content resources</td>
<td>Important</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td>Partial</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td></td>
</tr>
<tr>
<td>Cf.PL/REG/11</td>
<td>Identify the legal status of a given document</td>
<td>Important</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>No</td>
<td>Full</td>
<td>Partial</td>
<td>Full</td>
<td></td>
</tr>
<tr>
<td>FS/REG/05</td>
<td>Provide a user space in the registry</td>
<td>Important</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>NA</td>
<td>NA</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td>Cf.PL/REG/23</td>
<td>Export and share a search query and its results</td>
<td>Interesting</td>
<td>Partial</td>
<td>Partial</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>Full</td>
<td>Partial</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Cf.PL/REG/24</td>
<td>Update the results of a search query on the content</td>
<td>Interesting</td>
<td>Partial</td>
<td>Partial</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>Partial</td>
<td>Partial</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Cf.PL/REG/14</td>
<td>View a subset of linguistic and conceptual resources according to the users domains of interest</td>
<td>Interesting</td>
<td>Full</td>
<td>Full</td>
<td>No</td>
<td>Full</td>
<td>Full</td>
<td>NA</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td></td>
</tr>
<tr>
<td>FS/REG/07</td>
<td>An OMTD user will be able to search for text mining components, corpora, and linguistic resources.</td>
<td>Mandatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cf.PL/REG/04</td>
<td>Search the registry of text-mining services/components</td>
<td>Mandatory</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td>Cf.PL/REG/08</td>
<td>Search the registry of corpora / content datasets</td>
<td>Mandatory</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td>Cf.PL/REG/07-a</td>
<td>Browse the registry of corpora / content datasets based on domain-specific metadata</td>
<td>Interesting</td>
<td>Partial</td>
<td>Partial</td>
<td>Partial</td>
<td>Full</td>
<td>Full</td>
<td>No</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td></td>
</tr>
<tr>
<td>Cf.PL/REG/08-a</td>
<td>Faceted search on the registry of corpora / content datasets</td>
<td>Interesting</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td></td>
</tr>
<tr>
<td>Cf.PL/REG/08-b</td>
<td>Structure search results by topics</td>
<td>Interesting</td>
<td>Partial</td>
<td>Partial</td>
<td>No</td>
<td>Full</td>
<td>Full</td>
<td>NA</td>
<td>Full</td>
<td>Full</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>
### Table 3 Registry functional requirements assessment

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Level</th>
<th>Full</th>
<th>Partial</th>
<th>Full</th>
<th>Full</th>
<th>Full</th>
<th>Full</th>
<th>Full</th>
<th>Full</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cf.PL/REG/15</td>
<td>Search the registry of linguis2c resources (ontologies, etc.)</td>
<td>Mandatory</td>
<td>Full</td>
<td>NA</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td>FS/REG/08</td>
<td>An OMTD user will be able to browse the stored text mining components, corpora, and linguistic resources</td>
<td>Mandatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cf.PL/REG/03</td>
<td>Users can explore/Browse the registry of text-mining services/components</td>
<td>Mandatory</td>
<td>Full</td>
<td>Partial</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td>Cf.PL/REG/07</td>
<td>Users can explore/Browse the registry of corpora / content datasets</td>
<td>Important</td>
<td>Full</td>
<td>Partial</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
</tr>
<tr>
<td>Cf.PL/REG/07-a</td>
<td>Users can browse the registry of corpora / content datasets based on domain-specific metadata</td>
<td>Interesting</td>
<td>Partial</td>
<td>Partial</td>
<td>Partial</td>
<td>Full</td>
<td>Full</td>
<td>No</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td>Cf.PL/REG/13</td>
<td>Users can explore/Browse the registry of linguistic resources (ontologies, etc.)</td>
<td>Important</td>
<td>Full</td>
<td>Partial</td>
<td>Full</td>
<td>Full</td>
<td>No</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
</tr>
</tbody>
</table>

FS/REG/09 | Users are able to upload large files in the registry | Mandatory | NA | NA | NA | Full | Full | NA | Partial | Full | Partial |
FS/REG/10 | Users are able to download large files from the registry | Mandatory | NA | NA | NA | Full | Full | NA | Partial | Partial | Partial |
FS/REG/11 | The resources are accompanied by appropriate licenses | Mandatory | NA | NA | Full | Full | Full | Full | Full | Full | Full |

The table above includes 7 ‘mandatory’ requirements and one ‘important’ requirement. For the ‘mandatory’ ones, 4 can be further spitted into more specific user requirements. The results in these cases are quite positive as all cases get much more ‘Full’ scores than ‘Partial’ ones and very few negative scores - only 7 out of 159. Table 4 and its corresponding visualization chart in Error! Reference source not found.2 below summarize the scores for the mandatory functional requirements. As the figure shows, mandatory functional requirements get 75% of “Full” scores.
Error! Reference source not found.3 summarizes the scores for each mandatory functional requirements. Note that in this case, the user requirements involved (if any) are not explicitly listed, though their values are used to calculate the final scores. This chart gives a clear representation of the distribution of scores in each mandatory functional requirements.

Finally, Figure 44 below shows the scores for all mandatory functional requirements. Note that in this case, the chart also includes the user requirements involved, if any. The last ‘mandatory’
requirements (09 to 11) are not further split into user requirements. These have to do with upload/download capabilities and licensing.

**Mandatory functional requirements + user requirements (Registry)**

![Figure 4 Mandatory functional requirements + user requirements (Registry)](image)

Finally, the ‘Important’ functional requirement FS/REG/05 (“Provide a user space in the registry”) groups 3 user requirements. In this case, the functional requirement itself gets a ‘Full’ score whereas the user requirements involved (with priority = ‘Interesting’) get lower results.

**Workflow service**

As in the case of the registry above, the evaluation of the functional specifications relevant for the workflow service are reported in the following table. Again, for each functional requirement we give the description, the priority status and the list the related user requirements if any. Workflow service functionalities are divided into workflow editing (WFED) and workflow execution (WFEX) functionalities.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FS/WFED/01</td>
<td>An OMTD user will be able to view/inspect a workflow</td>
<td>Mandatory</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td>FS/WFED/03</td>
<td>An OMTD user will be able to configure/edit the parameters of the components of a workflow</td>
<td>Mandatory</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td>No</td>
<td>No</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td>Cf.PL/WFL/07</td>
<td>Configure and adapt pre-built workflows</td>
<td>Mandatory</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>No</td>
<td>Full</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Cf.PL/WFL/10</td>
<td>Define resources and content for a specific workflow</td>
<td>Mandatory</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cf.PL/WFL/05</td>
<td>Reproduce results of specific workflows</td>
<td>Important</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cf.PL/WFL/13</td>
<td>Input resources in the intermediate steps of a workflow</td>
<td>Important</td>
<td>Full</td>
<td>Full</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Full</td>
<td>NA</td>
<td>Partial</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>FS/WFED/04</td>
<td>Guided creation of workflows</td>
<td>Mandatory</td>
<td>Full</td>
<td>Full</td>
<td>No</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td></td>
</tr>
<tr>
<td>Cf.PL/WFL/10</td>
<td>Define resources and content for a specific workflow</td>
<td>Mandatory</td>
<td>Full</td>
<td>Full</td>
<td>No</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td></td>
</tr>
<tr>
<td>Cf.PL/WFL/16</td>
<td>Create workflows in a fully guided way</td>
<td>Important</td>
<td>Partial</td>
<td>Partial</td>
<td>No</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cf.PL/WFL/20</td>
<td>Transparency of text mining workflows</td>
<td>Important</td>
<td>Full</td>
<td>Full</td>
<td>No</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cf.PL/WFL/19</td>
<td>Use a clearly defined, standard interoperability framework for the creation of the workflows</td>
<td>Important</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td>Full</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS/WFED/05</td>
<td>An OMTD user will be able to save and share a workflow</td>
<td>Mandatory</td>
<td>Partial</td>
<td>Partial</td>
<td>No</td>
<td>Full</td>
<td>Full</td>
<td>NA</td>
<td>Full</td>
<td>No</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td>Cf.PL/WFL/11</td>
<td>Export workflows</td>
<td>Mandatory</td>
<td>Partial</td>
<td>Partial</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
<td>Full</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Cf.PL/WFL/18</td>
<td>Export workflows in standard formats</td>
<td>Important</td>
<td>Partial</td>
<td>Partial</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
<td>Full</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Requirement</td>
<td>Description</td>
<td>Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS/WFED/06</td>
<td>An OMTD user will be able to run a workflow from WFED</td>
<td>Mandatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CF.PL/WFL/01</td>
<td>Run text-mining services and workflows</td>
<td>Mandatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS/WFEX/02</td>
<td>An OMTD user will be able to download a workflow to a local machine</td>
<td>Mandatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CF.PL/WFL/15</td>
<td>Download and install pre-built text-mining workflows</td>
<td>Important</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS/WFEX/04</td>
<td>An OMTD user will be able to pause/resume/cancel the execution of a workflow</td>
<td>Important</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CF.PL/WFL/12</td>
<td>Pause workflows</td>
<td>Important</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CF.PL/WFL/13</td>
<td>Input resources in the intermediate steps of a workflow</td>
<td>Important</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS/WFEX/05</td>
<td>An OMTD user will be able to inspect the intermediate results and logs of a workflow execution</td>
<td>Important</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS/WFEX/08</td>
<td>The output of workflow execution should be managed and stored in a storage</td>
<td>Mandatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CF.PL/WFL/09</td>
<td>Export resources produced by the workflows</td>
<td>Mandatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS/WFEX/09</td>
<td>Appropriate metadata for workflow execution outputs should be generated and stored in the Registry</td>
<td>Mandatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 5 Functional requirements - workflow service**

The so called, ‘workflow edition’ functionalities are all mandatory and, though they get worse results than the registry ones, they still get high scores (61% “Full” vs. 10% “No”). Only the FS/WFED/05
functional specification gets more negative scores than positive ones. In this case, the various workflows were correctly saved and shared but the use cases reported problems when exporting them:

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Full</th>
<th>Partial</th>
<th>No</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS/WFED/01</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FS/WFED/03</td>
<td>28</td>
<td>5</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>FS/WFED/04</td>
<td>32</td>
<td>6</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>FS/WFED/05</td>
<td>7</td>
<td>3</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>FS/WFED/06</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 6 Summary of scores for mandatory functional requirements (workflow editor)

Figure 6 below, shows the distribution of scores for each functional requirement. As mentioned before, all FS get more positive values than negative ones except FS/WSED/05 (namely: “An OMTD user will be able to save and share a workflow”).

For the ‘workflow execution’ service, we only have three mandatory requirements. The scores obtained for the three mandatory requirements are summarized below. Note that whereas FS/WFEX/08 and FS/WFEX/09 get really good results, the FS/WFEX/02 specification (namely: “An OMTD user will be able to download a workflow to a local machine”) clearly fails and concentrates all negative values.
Table 7 Functional requirements’ scores - workflow editor

<table>
<thead>
<tr>
<th></th>
<th>Full</th>
<th>Partial</th>
<th>No</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS/WFEX/02</td>
<td>4</td>
<td>1</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>FS/WFEX/08</td>
<td>12</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>FS/WFEX/09</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>24</strong></td>
<td><strong>5</strong></td>
<td><strong>10</strong></td>
<td><strong>4</strong></td>
</tr>
</tbody>
</table>

Figure 7 Summary of scores for functional requirements (workflow execution)

Finally, Figure 8 below shows the distribution of scores for each mandatory functional requirement. As we can see, whereas FS/WSEX/08 and FS/WSEX/09 get rather positive scores, FS/WSEX/02 (“An OMTD user will be able to download a workflow to a local machine”) clearly fails.

Figure 8 Mandatory functional requirements - workflow execution
6. Interoperability specifications

WP5 formalized the text mining infrastructure interoperability specifications and promoted them for adoption and exploitation outside the project’s boundaries. The objectives of the interoperability framework were to (i) identify and evaluate relevant standards, specifications, best practices for adoption in the design and implementation of the OpenMinTeD platform, (ii) identify gaps, set rules for the use of standards, specifications and best practices in APIs, data models and design patterns of the OpenMinTeD platform, (iii) produce specifications to support integration of 3rd party TDM artefacts with the OpenMinTeD platform and to (iv) deliver comprehensive guidelines that define and exemplify interoperability in the context of OpenMinTeD.

Two especially relevant elements in the OMTD interoperability framework are the guidelines\(^{19}\) and the metadata schema\(^{20}\) (OMTD-SHARE schema). Guidelines address providers of the resources that OpenMinTeD targets and aim at helping them adopt interoperability specifications. The OMTD-SHARE schema aims at providing the interoperability bridge between the various resources types involved in TDM processes. In the following lines, we describe the evaluation performed concerning the guidelines and the metadata.

**Interoperability guidelines**

The interoperability guidelines\(^{21}\) (https://guidelines.openminted.eu and reported in D5.5 and D5.6) were based on the outcome of D5.1, D5.2, D5.3, and D5.4 as well as general discussions in the interoperability working groups and were extensively used, tested and evaluated by participants in the Open Call as well as partners in the consortium (use cases of WP9 and providers of apps/components).

The guidelines evaluation consists of a checklist divided into two parts: a ‘descriptive’ part and a ‘experience reporting’ part. The former collects information about the structure and content of the guidelines and the latter is designed to report users’ experience.

The descriptive checklist is listed below. The overall assessment in quite positive as the guidelines got 7 out of 9 positive scores.

---

\(^{19}\) [https://guidelines.openminted.eu/](https://guidelines.openminted.eu/)
\(^{21}\) [https://guidelines.openminted.eu/](https://guidelines.openminted.eu/) and reported in
# Availability

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is the guideline readily available?</td>
<td>yes</td>
</tr>
<tr>
<td>Document organization</td>
<td></td>
</tr>
<tr>
<td>Are the guidelines divided into chapters?</td>
<td>yes</td>
</tr>
<tr>
<td>Is there a table of content?</td>
<td>yes</td>
</tr>
<tr>
<td>Is there a term index?</td>
<td>yes</td>
</tr>
<tr>
<td>Is there an acronym glossary?</td>
<td>yes</td>
</tr>
<tr>
<td>Does the guideline include a search facility?</td>
<td>yes</td>
</tr>
<tr>
<td>Does the guideline provide a complete reference list?</td>
<td>no</td>
</tr>
<tr>
<td>Does the guideline provide a summary of its recommendations?</td>
<td>no</td>
</tr>
<tr>
<td>Dates</td>
<td></td>
</tr>
<tr>
<td>Is there a date of completion available?</td>
<td>no</td>
</tr>
<tr>
<td>Guideline developers</td>
<td></td>
</tr>
<tr>
<td>Are the developers of the guideline clearly stated?</td>
<td>yes</td>
</tr>
</tbody>
</table>

Table 8 Guidelines content form

Following the COCA quality model for user documentation\(^{22}\), we defined an evaluation survey form that focuses on four quality characteristics: Completeness, Operability, Correctness, and Appearance as defined below:

- **Completeness** is the degree to which user documentation provides all the information needed by end users to use the described software.
- **Operability** is the degree to which user documentation has attributes that make it easy to use and helpful when acquiring information that is contained in the user documentation. Operability depends on two other criteria: Completeness and Correctness.
- **Correctness** is the degree to which the descriptions provided by the user documentation are correct.
- **Appearance** is the degree to which information contained in user documentation is presented in an aesthetic way.

In order to get a complete picture, we asked volunteers with different profiles to fulfill the form. For this exercise, we distinguished between:

- **Developers**: developers that use the guidelines when integrating their comp/apps in the OMTD platform.

---

\(^{22}\) Alchimowicz, B. & Nawrocki, J.R. Software Qual J (2016) 24: 205. [https://doi.org/10.1007/s11219-014-9252-4](https://doi.org/10.1007/s11219-014-9252-4)
● **experts**: TDM experts with technical abilities that plan to integrate their comp/apps in the OMTD platform and use the guidelines to learn about the interoperability requirements.

● **non-technical users**: non-technical TDM users (and/or experts) that read the guidelines to have a general overview of the interoperability requirements.

Users were asked to evaluate different aspects using the following range of values:

Not at all (N), Weak (W); Hard to say (?); Good enough (G); Very good (VG).

<table>
<thead>
<tr>
<th>Guideline purpose and users</th>
<th>DEV</th>
<th>DEV</th>
<th>DEV</th>
<th>EXPERT</th>
<th>EXPERT</th>
<th>non TECH</th>
<th>non TECH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are the purpose and target users of the guideline stated?</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>VG</td>
<td>VG</td>
<td>VG</td>
<td>VG</td>
</tr>
<tr>
<td>The target users are clearly defined?</td>
<td>G</td>
<td>G</td>
<td>VG</td>
<td>VG</td>
<td>G</td>
<td>VG</td>
<td>G</td>
</tr>
<tr>
<td>The target situations in which the guideline is to be applied are clearly defined?</td>
<td>G</td>
<td>VG</td>
<td>VG</td>
<td>VG</td>
<td>VG</td>
<td>VG</td>
<td>G</td>
</tr>
</tbody>
</table>

**Ease of use & intelligibility**

| Is the guideline readable and easy to navigate? | G | G | VG | VG | G | G | VG |
| The guideline is organized in such a way that it is generally easy to understand, and the key recommendations are easy to identify. | VG | G | G | VG | G | VG | G |
| Users know at all times where they are in the documentation relative to the whole, and relative to where they were previously. | VG | G | VG | VG | G | VG | G |
| The recommendations are specific and unambiguous. | G | G | G | VG | VG | VG | G |
| The guideline includes examples. | N | N | G | ? | G | G | ? |

**Completeness**

| To what extent does the guideline covers all the functionality provided by the system with the needed level of detail? | G | W | W | G | ? | ? | G |
| To what extent does the guideline contain information about how to use it with effectiveness and efficiency? | ? | G | G | ? | G | ? | G |
| To what extent does the guideline provide correct descriptions with the needed degree of precision? | W | W | ? | G | G | G | G |

**Appearance**

| To what extent is the information contained in the guideline presented in an aesthetic way? | G | G | ? | G | G | G | G |

*Table 9 Guidelines assessment form*
In general, most participants gave positive scores for all aspects with 25 “Very Good”, 44 “Good” and only 2 “Not at all”. The Completeness aspect is the one with lower scores: 15 “Good” and 7 “Hard to say” and only when asking for examples, we get negative answers.

![Guidelines evaluation by users](image)

*Figure 9 Guidelines evaluation by users*

In general, there are no significative differences between users. Expert and non-tech users gave slightly higher scores than developers. This may indicate that, in general, the guidelines are useful and provide (very) good descriptions but when focusing in details they may get lower scores.

In the section “Open Calls”, we give some more details about the comments done by the Open Call participants concerning the guidelines.

**The OMTD-SHARE metadata schema**

Besides the interoperability guidelines, the OMTD-SHARE metadata schema is also a key component of the OMTD interoperability framework. The OMTD-SHARE metadata schema has been conceived and designed in order to serve as a facilitator, providing the interoperability bridge between the various resource types involved in TDM processes, and as an intermediary with the target audience, including TDM developers and end-users. It is worth mentioning that some partners have integrated OMTD-SHARE metadata into their frameworks (e.g. DKPro Core version 1.9.1) which makes such metadata available to a wider community beyond the immediate users of the OpenMinTeD platform.

As we mentioned before, we have been working together with the ELIXIR-OpenEBeanch team to guarantee interoperability between the OMTD-Share schema and the ELIXIR schema.
We are also deploying a CMDI version of the OMTD-SHARE schema following the specification given by CLARIN\textsuperscript{23}.

Note in addition that the OMTD-SHARE schema makes use of the OMTD-SHARE ontology and currently we are running different efforts to align and coordinate ontology descriptions with other EU projects in order to guarantee interoperability among them.

- **EDAM ontology**\textsuperscript{24}: EDAM is an ontology of concepts that are prevalent within bioinformatics and computational biology, including types of data and data identifiers, data formats, operations and topics. The ontology is used by \textit{bio.tools}\textsuperscript{25}, a discovery portal for bioinformatics software information developed by the ELIXIR project.

- **Foster taxonomy**\textsuperscript{26}: The FOSTER portal is an e-learning platform that brings together the best training resources addressed to those who need to know more about Open Science.

\textsuperscript{23} \url{https://www.clarin.eu/}
\textsuperscript{24} \url{http://bioportal.bioontology.org/ontologies/EDAM?p=classes}
\textsuperscript{25} \url{https://bio.tools/}
\textsuperscript{26} \url{https://www.fosteropenscience.eu/openminted/text-and-data-mining}
7. The Open Tender Calls as an evaluation opportunity

One of the objectives of OpenMinTeD is to improve the uptake of the project’s infrastructure, promoting the openness and reuse principles and the Open Calls constituted an essential part of this promoting task. Among the main aims of the tender calls were:

- Engage members of various communities to develop applications/prototypes on top of the infrastructure, and encourage them to share services, content, resources, and knowledge.
- Tender applications/prototypes should build upon and fully align with the OpenMinTeD platform (compatible with the OpenMinTeD functional and interoperability specifications).
- The calls should be used to motivate text and data mining experts to contribute missing natural language processing, document retrieval, text processing and other components to the OMTD platform.
- Encourage content providers (publishers, repositories, libraries and other holders of scholarly works) to make their content accessible through the platform to a wider research audience.
- Make the OpenMinTeD platform known outside the project consortium and improve its uptake.

The various participants in the Open Calls developed applications that covered a wide range of aspects, such as using and providing new services or providing repository or journal publishing platform plugins to increase content. A detailed description of the different participating projects is out of the scope of this document and can be found in the Deliverable D2.5. In any case, the Open Call constituted an excellent opportunity to get feedback and input from external users, essentially developers that were integrating their resources into the OMTD platforms following the interoperability guidelines.

The work plan of the Open Calls included the delivery of a technical description document and a final report. These deliverables pursue a double objective: to assess the various projects participating in the Open Call and to collect information from third party users for the evaluation of the OMTD itself. With this purpose in mind, we defined and distributed a template for the final report document. The template, included in the Appendix “Deliverable T.4 Final report”, intended to summarize in an orderly manner the main outcomes of the project, the most relevant experiences, the unexpected issues, the problems found, etc. We designed this template as a series of free text questions (and expressly avoided making a checklist) to encourage thought and freedom of expression.

The final reports delivered by the participants are included at the end of the Deliverable D2.5, in the following lines we summarize and analyze the results reported. We list and comment the positive and negative aspects referred by participants grouping them by topics and listing them in order of ‘most cited aspects’.

---

Positive aspects (assets, lessons learnt...)

1. Technical support received in the GitHub hackathon sessions:
   a. All participants highlighted and thanked the support received.
   b. Some also welcomed the idea of using GitHub as the technical support platform because of the:
      ■ ability to share experiences and code,
      ■ functionalities and easy of use and
      ■ ability to get a wider picture of what the TDM community is doing.

2. Final result:
   a. Participants highlighted the fact that the Open Call gave them the opportunity to provide a new final product which is well packaged, self-contained and well documented.
   b. The experience allowed their applications to reach a wider audience and expose their work to other communities.

3. Learning opportunity:
   a. Participants reported (and thanked) the experience they gained with:
      ■ integration, sharing and interoperability issues
      ■ packaging software (specially in reference to Docker)
      ■ knowing what others are doing
      ■ They see their experience as a proof of concept

4. OMTD platform: Participants believe that the OMTD platform
   a. Will allow non-expert users use TDM tools.
   b. Is an important step forward to interoperability between applications and resources.

Negative aspects (missing things, suggestions for improvement...):

1. Missing examples:
   In general, all participants complained about missing examples for XMI input/output files and integration templates and/or examples.

2. Need for more detailed guidelines:
   Though some participants reported that the guidelines were clear and useful for the design phase, they all complained they missed some technical details.
   They missed a step-by-step integration tutorial and a FAQ section

3. Metadata editor needs some help:
   Though some participants were happy with the XML editor they would appreciate some help and reported they did not fully understand all fields.
4. Testing their comp/apps was not easy
5. Need for more computing resources available for individual components

We consider the results above to be very positive. As can be seen above, positive aspects are many and varied whereas negative aspects have to do with lack of examples and documentation problems. It is worth noting that these aspects have been addressed so that during the Open Calls improvements have been added to the documentation as demanded by participants. As a consequence of the interaction with the Open Call participants, the eventual guidelines were improved and should cover a wide range of issues.
8. **Use case final reports**

The final report template used to assess the selected project in the open call was used also to get relevant information from use cases. As mentioned before, the template is a short document intended to summarize in an orderly manner the main outcomes of the project, the most relevant experiences, the unexpected issues and factors that affected the project results, the problems found, etc.

The appendix “Use case evaluation reports” collects all reports sent by the use cases, in the following lines we summarize and list the most cited and highlighted aspects as we did in the previous section:

**Positive aspects (assets, lessons learnt...)**

1. **Availability of resources:**
   - The platform provides easy access to a good amount of important linguistic tools.
   - The creation of a platform where data sources and “algorithms” live in the same place.
   - The platform provides easy access to a huge amount of data (via protocols to access content).
   - This “cloud” nature greatly simplifies the sharing of either computational components or linguistic resources.

2. **Community:**
   - The platform provided the opportunity to work with a wide community of TDM and domain experts.
   - Collaboration between OMTD partners and participants to the tender calls.
   - Cooperation between developers and users.

3. **Technical improvements acquired:**
   - Packaging and distribution: use cases end with well-packaged components/apps.
   - Testing: use cases end with tested and more robust components/apps.
   - The project motivated several improvements and updates.

4. **Promotion and reusability of resources:**
   - The registry makes comp/apps more discoverable and promotes reusability.
   - Content and service providers gain visibility.
   - The platform allows and promotes sharing resources and tools.

5. **Interoperability:**
   - The platform is seen as a step towards interoperability.
   - Use cases gain awareness about interoperability (including UIMA type system).
6. OMTD ontology

Negative aspects (missing things, suggestions for improvement...):

Contrary to the ‘positive’ aspects listed above, the ‘negative’ aspects reported by the use cases were more heterogeneous and less repeated. Consequently, to avoid the notion of ranking, we use bullets instead of a numbered list.

- Technical issues at various levels when trying to deploy the components to the OMTD platform. In general, the technical problems referred to were due to the complexity of the tasks, the need for better examples and to unexpected events. Note also that most of the technical issues reported are specific to each use case.

- The delay in the platform and in the testing phase.

- The lack of a complete tool for manipulating and extracting information from PDF documents (PDF files pose many problems when identifying spans).

- Format conversions: the efforts needed to translate inputs/outputs from/into XMI. Note, however, that type sharing and XMI are valued positively.

- Need for some more examples: some more practical examples would greatly benefit the preparation of the components, thus resulting in accelerating the integration process.

- Licensing: intellectual property rights make the collaborative curation of corpora annotated by the OMTD platform very restricted.

- Metadata requirements for the registry are considered too complex. These are already difficult for technical minded people, so they will be difficult for non-technical users.

- Some actions take too much time: developers suggest to speed up the corpus creation time and the time it takes to run workflows

Despite the technical issues and the delay in the development of the platform (as well as some other minor issues reported), again the use cases gave a positive assessment and they highlighted and agreed in a large number of positive aspects. Not only they reported more positive aspects than negative ones, but also they showed more agreement in their positive assessments. Note that most of the technical problems reported are specific to each use case. This heterogeneity reflects the difficulty of the task: the integration of different uses cases in the same platform involving different requirements, different resources and tools, different input and output formats, different protocols, etc.
The open call participants and use cases agree in positively assessing the availability of resources, the experience and technical improvements acquired, the gain in visibility etc. Note, however, that whereas the open call participants highlighted the support received, the use cases valued the ‘community’ aspect of the project. This is because the open call was developed during a very short period of time and technical support was crucial.
9. Appendix

The appendix section comprises sample structured questionnaires from the OpenMinTeD evaluation framework.

Functional requirements evaluation form - Registry

<table>
<thead>
<tr>
<th>FUNCT REQ ID</th>
<th>DESCRIPTION</th>
<th>STATUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS/REG/01</td>
<td>Maintain a registry of text mining components</td>
<td>Mandatory</td>
</tr>
<tr>
<td>Cf.PL/REG/26</td>
<td>Users can register text mining components</td>
<td>Mandatory</td>
</tr>
<tr>
<td></td>
<td>Update and manage a registry of text-mining services/components (User can update an already registered component)</td>
<td>Mandatory</td>
</tr>
<tr>
<td>Cf.PL/REG/01</td>
<td>Users can view components characteristics</td>
<td>Mandatory</td>
</tr>
<tr>
<td></td>
<td>Maintain a registry of corpora</td>
<td>Mandatory</td>
</tr>
<tr>
<td>Cf.PL/REG/09</td>
<td>Users can retrieve various content resources through a single query</td>
<td>Important</td>
</tr>
<tr>
<td>Cf.PL/REG/09-a</td>
<td>Users can apply faceted search on textual content resources</td>
<td>Important</td>
</tr>
<tr>
<td>Cf.PL/REG/11</td>
<td>Identify the legal status of a given document</td>
<td>Important</td>
</tr>
<tr>
<td>FS/REG/05</td>
<td>Provide a user space in the registry</td>
<td>Important</td>
</tr>
<tr>
<td>Cf.PL/REG/23</td>
<td>Export and share a search query and its results</td>
<td>Interesting</td>
</tr>
</tbody>
</table>
### D4.5 - Community Evaluation Report

| CF/PL/REG/24 | Update the results of a search query on the content | Interesting |
| CF/PL/REG/14 | View a subset of linguistic and conceptual resources according to the users domains of interest | Interesting |

| FS/REG/07 | An OMTD user will be able to search for text mining components, corpora, and linguistic resources. | Mandatory |
| CF/PL/REG/04 | Search the registry of text-mining services/components | Mandatory |
| CF/PL/REG/08 | Search the registry of corpora / content datasets | Mandatory |
| CF/PL/REG/07-a | Browse the registry of corpora / content datasets based on domain-specific metadata | Interesting |
| CF/PL/REG/08-a | Faceted search on the registry of corpora / content datasets | Interesting |
| CF/PL/REG/08-b | Structure search results by topics | Interesting |
| CF/PL/REG/15 | Search the registry of linguistic resources (ontologies, etc.) | Mandatory |

| FS/REG/08 | An OMTD user will be able to browse the stored text mining components, corpora, and linguistic resources | Mandatory |
| CF/PL/REG/03 | Users can explore/Browse the registry of text-mining services/components | Mandatory |
| CF/PL/REG/07 | Users can explore/Browse the registry of corpora / content datasets | Important |
| CF/PL/REG/07-a | Users can browse the registry of corpora / content datasets based on domain-specific metadata | Interesting |
| CF/PL/REG/13 | Users can explore/Browse the registry of linguistic resources (ontologies, etc.) | Important |
| FS/REG/09 | Users are able to upload large files in the registry | Mandatory |
| FS/REG/10 | Users are able to download large files from the registry | Mandatory |
| FS/REG/11 | The resources are accompanied by appropriate licenses | Mandatory |
### Functional requirements evaluation form - Workflow

<table>
<thead>
<tr>
<th>FUNCT REQ ID</th>
<th>DESCRIPTION</th>
<th>STATUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS/WFED/01</td>
<td>An OMTD user will be able to view/inspect a workflow</td>
<td>Mandatory</td>
</tr>
<tr>
<td>FS/WFED/03</td>
<td>An OMTD user will be able to configure/edit the parameters of the components</td>
<td>Mandatory</td>
</tr>
<tr>
<td></td>
<td>of a workflow</td>
<td></td>
</tr>
<tr>
<td>CF.PL/WFL/07</td>
<td>Configure and adapt pre-built workflows</td>
<td>Mandatory</td>
</tr>
<tr>
<td>CF.PL/WFL/10</td>
<td>Define resources and content for a specific workflow</td>
<td>Mandatory</td>
</tr>
<tr>
<td>CF.PL/WFL/05</td>
<td>Reproduce results of specific workflows</td>
<td>Important</td>
</tr>
<tr>
<td>CF.PL/WFL/13</td>
<td>Input resources in the intermediate steps of a workflow</td>
<td>Important</td>
</tr>
<tr>
<td>FS/WFED/04</td>
<td>Guided creation of workflows</td>
<td>Mandatory</td>
</tr>
<tr>
<td>CF.PL/WFL/10</td>
<td>Define resources and content for a specific workflow</td>
<td>Mandatory</td>
</tr>
<tr>
<td>CF.PL/WFL/16</td>
<td>Create workflows in a fully guided way</td>
<td>Important</td>
</tr>
<tr>
<td>CF.PL/WFL/20</td>
<td>Transparency of text mining workflows</td>
<td>Important</td>
</tr>
<tr>
<td></td>
<td>Use a clearly defined, standard interoperability framework for the creation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>of the workflows</td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
<td>Requirement</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>FS/WFED/05</td>
<td>An OMTD user will be able to save and share a workflow</td>
<td>Mandatory</td>
</tr>
<tr>
<td>Cf.PL/WFL/11</td>
<td>Export workflows</td>
<td>Mandatory</td>
</tr>
<tr>
<td>Cf.PL/WFL/18</td>
<td>Export workflows in standard formats</td>
<td>Important</td>
</tr>
<tr>
<td>FS/WFED/06</td>
<td>An OMTD user will be able to run a workflow from WFED</td>
<td>Mandatory</td>
</tr>
<tr>
<td>Cf.PL/WFL/01</td>
<td>Run text-mining services and workflows</td>
<td>Mandatory</td>
</tr>
<tr>
<td>FS/WFEX/02</td>
<td>An OMTD user will be able to download a workflow to a local machine</td>
<td>Mandatory</td>
</tr>
<tr>
<td>Cf.PL/WFL/15</td>
<td>Download and install pre-built text-mining workflows</td>
<td>Important</td>
</tr>
<tr>
<td>FS/WFEX/04</td>
<td>An OMTD user will be able to pause/resume/cancel the execution of a workflow</td>
<td>Important</td>
</tr>
<tr>
<td>Cf.PL/WFL/12</td>
<td>Pause workflows</td>
<td>Important</td>
</tr>
<tr>
<td>Cf.PL/WFL/13</td>
<td>Input resources in the intermediate steps of a workflow</td>
<td>Important</td>
</tr>
<tr>
<td>FS/WFEX/05</td>
<td>An OMTD user will be able to inspect the intermediate results and logs of a workflow execution</td>
<td>Important</td>
</tr>
<tr>
<td>FS/WFEX/08</td>
<td>The output of workflow execution should be managed and stored in a storage</td>
<td>Mandatory</td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
<td>Status</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Cf.PL/WFL/09</td>
<td>Export resources produced by the workflows</td>
<td>Mandatory</td>
</tr>
<tr>
<td>FS/WFEX/09</td>
<td>Appropriate metadata for workflow execution outputs should be generated and stored in the Registry</td>
<td>Mandatory</td>
</tr>
</tbody>
</table>
Guidelines evaluation form

[yes/no answers]

Availability
- Is the guideline readily available?

Document organization
- Are the guidelines divided into chapters?
- Is there a table of content?
- Is there a term index?
- Is there an acronym glossary?
- Does the guideline include search option?
- Does the guideline provide a complete reference list?
- Does the guideline provide a summary of its recommendations?

Dates
- Is there a date of completion available?
- Does the guideline provide an anticipated review date?
- Does the guideline provide dates for when literature was included?

Guideline developers
- Are the developers of the guideline clearly stated?

[Not at all (N), Weak (W); Hard to say (?) Good enough (G); Very good (VG)]

Guideline purpose and users
- Are the purpose and target users of the guideline stated?
- The target users are clearly defined?
- The target situations in which the guideline is to be applied, are clearly defined?

Ease of use & intelligibility
- Is the guideline readable and easy to navigate?
- The guideline is organized in such a way that it is generally easy to understand, and the key recommendations are easy to identify.
- Users know at all times where they are in the documentation relative to the whole, and relative to where they were previously.
- The recommendations are specific and unambiguous.
- The guideline includes examples.
D4.5 - Community Evaluation Report

- To what extent does the guideline cover all the functionality provided by the system with the needed level of detail?
- To what extent does the guideline provide information which is helpful in deciding whether the system is appropriate for the needs of prospective users?
- To what extent does the guideline contain information about how to use it with effectiveness and efficiency?
- To what extent is the guideline easy to use and helpful when operating the system documented by it?
- To what extent does the guideline provide correct descriptions with the needed degree of precision?
- To what extent is the information contained in the guideline presented in an aesthetic way?
Open Call Final report template

As software provider, you shared your resource(s) as (choose):
  ○ UIMA/GATE components uploaded via maven central repository
  ○ Docker images
  ○ Web services

Briefly, explain/justify your choice:
...

List the highlights and key success factors of the project:
...

List and identify anything you found useful:
...

List and describe any unexpected events
(unexpected events that occurred during the project and the impact that those events may have had on the project and the action(s) taken to address them if any. Common problems are normally due to (please identify)
  ● the complexity of the task,
  ● lack of documentation,
  ● lack of suitable supporting tools and/or examples,
  ● over-ambitious project aims,
  ● time schedule,
  ● other ...

List any issue you had regarding (interoperability) requirements related to
(please use none if you had no relevant issues):
  ● Licensing.
  ● Metadata description.
  ● Registry.
  ● Input / output restrictions & formats.
  ● Preparing & Packaging (for Maven, Docker and web services).
  ● Other
  ● None

The amount of work to adapt and integrate your tools into OMTD was
  ● as expected
  ● much bigger
  ● Other ....
Do you think the effort done is justified? (yes, no, to what extend)  
...

List and summarize any lessons learned from this project.  
....

Regarding the use case you presented, list the aspects that OMTD can (or could) benefit the use case.  
....

List what you would have done differently and/or your recommendations for improvement:  
.....
Use case evaluation reports

Use Case evaluation report: AS-A Text mining over bibliographic data: AGRIS & CORE

List the highlights and key success factors of the project:

The main highlight of the AS-A use case is the extraction of structured information from unstructured bibliographic resources for the domain of Viticulture. The key success factors are:

- The use of rich & established sources of bibliographic resources, like AGRIS (http://agris.fao.org) and CORE (https://core.ac.uk). AGRIS provides about 8 million bibliographic records from all over the world, and CORE provides access to about 37 million open access research publications. Both sources provide the means to access their content programmatically.
- The use of established and well documented linguistic resources like Agrovoc (http://aims.fao.org/standards/agrovoc/concept-scheme) as well as the creation of linguistic resource, when needed, that are based on formal sources like the Grape Varieties Ontology. In this case, we have used the complete list of grape varieties as published by the International Organisation of Vine and Wine (http://www.oiv.int/public/medias/2273/oiv-liste-publication-2013-complete.pdf).
- The use of mature text manipulation and search tools like Lucene (https://lucene.apache.org/core/) and PDF Miner.

List and identify anything you found useful:

The maturity of important linguistic resources that are directly connected to the use case, as well as the existence of rich sources that are easily accessible. Also, the final evaluation provided satisfactory results, which showed that the components could really benefit the community of the specific use case namely the viticulture research community.

More specifically, the evaluation was performed by 2 groups with 15 domain experts for each group (postgraduate students from Agricultural University of Athens with Agricultural, Food and Nutrition background). As in the first evaluation the domain, experts were asked to list (a) the recognized AgroVoc Terms, and (b) the Grape Varieties of a larger set of testing input. The size of the testing input was 50 papers from each source and the evaluation performed on the workflow that has been deployed on OMTD platform. The produced term lists were merged to create two testing term lists - one for each component (namely, the AgroVoc Extractor and Grape Varieties Extractor).

Different open-source PDF extractors were used to generate the textual information, over which the two extractors were applied. Depending on the area on which each term was located (abstract, paper body, table/figure caption), the PDF extractors were evaluated so that for each area the extractor with the best f-measure was selected.

It is worthy to mention that the best score (AgroVoc: 0.82 / Grape Varieties:0.89) was retrieved when examining the abstract section, whereas the worst score (AgroVoc: 0.60 / Grape Varieties: 0.74) was observed at table/figure captions.

List and describe any unexpected events
The lack of a complete tool for manipulating and extracting information from PDF documents. To overcome this obstacle, we encapsulated several different tools and combined their best features in order extract as much as information possible from the PDF documents.

List any issue you had regarding (interoperability) requirements related to
(please use none if you had no relevant issues):

● The only issue related to interoperability, was to adjust the components to be compatible with the input requirements of the project. Originally the components used in the context of the use case AS-A, required as input a document that contains plain text. To fulfill the requirements of the project the components were modified to handle also XMI documents (https://www.omg.org/spec/XMI/About-XMI/).

The amount of work to adapt and integrate your tools into OMTD was

● As expected

Do you think the effort done is justified?

● yes

Regarding the use case you presented, list the aspects that OMTD can (or could) benefit the use case.

OMTD brought two main benefits to the AS-A use case. The first one is that, through OMTD a researcher or a domain expert can not only use complex computational components in workflows important for their work but also those workflows can be shared to the corresponding community, thus increasing the reusability of the created workflow.

The second one is the “cloud” nature of OMTD. This “cloud” nature greatly simplifies the sharing of either computational components or linguistic resources. The fact that OMTD is compatible with the Docker Hub and supports Web Services greatly simplifies the sharing and use of computational components and linguistic resources. Moreover, the simple yet thorough user interface makes available almost instantly the desired component or linguistic resource for use by the community.

List what you would have done differently and/or your recommendations for improvement:

Although the process of integrating a component to the OMTD platform is quite straightforward, the lack of a tangible example slowed the whole process. Some more practical examples would greatly benefit the preparation of the components, thus resulting in accelerating the integration process.
Use Case evaluation report: AS-B Text-mining over RSS Feeds: Food Safety & Water Health

List the highlights and key success factors of the project:

The main highlight of the AS-B use case is the extraction of geolocation information from unstructured bibliographic resources targeting the Food Safety and Water Health Community. The key success factors are:

- The use of established and well documented linguistic resources like Geonames (http://api.geonames.org/) and the FAO Geopolitical Ontology (http://www.fao.org/countryprofiles/geoinfo/en/)
- The use of mature text manipulation and search tools like Lucene (https://lucene.apache.org/core/) and PDF Miner.

List and identify anything you found useful:

The maturity of important linguistic resources that are directly connected to the use case as well as the existence of rich sources that are easily accessible. Also, the final evaluation provided satisfactory results that were used to guide the calibration of the term extraction modules to balance the trade-off between highly accurate results and inability to classify (recall in this case).

More specifically, an evaluation was performed by 2 groups with 15 domain experts for each group (postgraduate students from Agricultural University of Athens with Agricultural, Food and Nutrition background). They asked to classify collaboratively a given set of 200 randomly selected Food Safety Alerts retrieved from RASFF’s (the European Rapid Alert System for Food and Feed) Consumer Portal (https://webgate.ec.europa.eu/rasff-window/consumers/) to clusters of food recalls referring to products from the same geographical origin.

The extracted Geolocation terms for each feed were compared to the indicated (by the experts) cluster to compute the accuracy and recall of the extraction (Geonames: 0.80 / FAO Geopolitica:0.85).

List and describe any unexpected events

The lack of a complete tool for manipulating and extracting information from PDF documents. To overcome this obstacle, we encapsulated several different tools and combined their best features in order to extract as much as information possible from the PDF documents.

List any issue you had regarding (interoperability) requirements related to

(please use none if you had no relevant issues):

- The only issue related to interoperability, was to adjust the components to be compatible with the input requirements of the project. Originally the components used in the context of the use case AS-B, required as input RSS Feeds documents. To fulfill the requirements of the project the components were modified to handle also XMI documents (https://www.omg.org/spec/XMI/About-XMI/).
The amount of work to adapt and integrate your tools into OMTD was
  - As expected

Do you think the effort done is justified?
  - yes

Regarding the use case you presented, list the aspects that OMTD can (or could) benefit the use case.
OMTD brought two main benefits to the AS-B use case. The first one is that, through OMTD a researcher or a domain expert can not only use complex computational components in workflows important for their work but also those workflows can be shared to the corresponding community, thus increasing the reusability of the created workflow.

The second one is the “cloud” nature of OMTD. This “cloud” nature greatly simplifies the sharing of either computational components or linguistic resources. The fact that OMTD is compatible with the Docker Hub and supports Web Services greatly simplifies the sharing and use of computational components and linguistic resources. Moreover, the simple yet thorough user interface makes available almost instantly the desired component or linguistic resource for use by the community.

List what you would have done differently and/or your recommendations for improvement:
Although the process of integrating a component to the OMTD platform is quite straightforward, the lack of a tangible example slowed the whole process. Some more practical examples would greatly benefit the preparation of the components, thus resulting in accelerating the integration process.
Use cases evaluation reports: AS-C Microbial Biodiversity, AS-D Linking Wheat Data With Literature and AS-E Extracting gene regulation networks involved in seed development (SeeDev)

List the highlights and key success factors of the project:
- OMTD-SHARE and the ontology
- The guidelines
- The registry of resources
- Proposed interoperability based on the UIMA type systems
- Collaboration between OMTD partners and participants to the tender calls

List and identify anything you found useful:
- OMTD-SHARE and the ontology
- The registry and the guidelines
- The UIMA type systems
- The resources in the registry
- OpenMinTeD AERO Protocol

List and describe any unexpected events:
(unexpected events that occurred during the project and the impact that those events may have had on the project and the action(s) taken to address them if any. Common problems are normally due to (please identify)
- Lack of navigation buttons, user profile edition and transparent logs for users on the registry interface
- The usage of Galaxy framework was over-ambitious: The Galaxy API seemed less easy than what was expected. Others interesting functionalities of Galaxy (main interface, toolshed, conda packages, etc.) are not used.
- There is misunderstanding of some instructions of the guidelines (i.e. instructions to define and deploy docker-based components)
- Lack of information from the infrastructure stack to better prevent issues related to the size of the docker images, RAM memory
- Lack of annotation editors

List any issue you had regarding (interoperability) requirements related to
(please use none if you had no relevant issues):
- Despite several efforts, the creation of workflows is not enough straightforward from the point of view of users: lack of help about the components from the Galaxy Workflow Editor. In addition, the metadata of the components/workflows are not full visible on the workflow service.
● The bugs that occur when using the registry are not explicitly and well presented to the end-users
● Creating metadata for the resources by hand is time consuming and error prone

The amount of work to adapt and integrate your tools into OMTD was
● as expected
● much bigger
● Other....

Do you think the effort done is justified?
● yes,
● no,
● to what extend?

List and summarize any lessons learned from this project.
● Get quickly a first simplified version that works and try to enhance after
● Better schedule the priority tasks
● Better manage the bugs (submission, notification, reparation)

Regarding the use case you presented, list the aspects that OMTD can (or could) benefit the use case.
● Share domain resources and specialized TDM tools
● Facilitate the reuse of different resources
● Invite communities to OpenMinTeD
● Reuse of components between different applications, e.g. molecule entities detection developed in the ChEBI use-case could be useful for the Microbiotope use-case.

List what you would have done differently and/or your recommendations for improvement:
● Usage of the Galaxy framework
● Enhance the user spaces (e.g., integrate groups, permission levels)
● Add annotation editors (e.g., based on OpenMinTeD AERO Protocol)
● Enhance the interoperability between the components (e.g., syntactic and semantic)
● Better automate and document the painful functionalities
● Better anticipate the issues (e.g., related to memory and slowness of the platform)
Use case evaluation report: LS-A Extract metabolites and their properties and modes of actions

List the highlights and key success factors of the project:

- A new corpus to support text mining for the curation of metabolites in the ChEBI database was created. The corpus specifically contains entities of Metabolite, Chemical, Protein, Species, Biological Activity, and Spectral Data. It was also marked up with four relations between these entities:
  - Isolated_From (Metabolites, Species),
  - Associated_With (Chemical/Metabolite, Biological Activity/Spectral Data),
  - Binds_With (Chemical/Metabolite, Protein),
  - Metabolite_Of (Metabolite, Chemical).

The resulting corpus consists of 200 PubMed abstracts and 100 PMC full papers.

- An application that can automatically identify Metabolite, Chemical, Protein, Species, Biological Activity, and Spectral Data entities in text has been published as a standalone web service as well as a component via the OpenMinTeD platform. Output of the application can be used to enhance the quality of the curated database entries and the productivity of curators.


List and identify anything you found useful:

- We have had opportunities to work closely with the community to make TDM applications practical and useful.
- Registry makes our web service more discoverable. Promotes reuse, etc.
- Potential to use other services for chemistry, should they be made available on the platform.

List and describe any unexpected events

(unexpected events that occurred during the project and the impact that those events may have had on the project and the action(s) taken to address them if any. Common problems are normally due to (please identify)

- the complexity of the task,
- lack of documentation,
- lack of suitable supporting tools and/or examples,
- over-ambitious project aims,
- time schedule,
- other …
We expected that the platform would be ready much earlier in the project. Due to the lack of an available platform, we halted progress on our workflow. This meant that we completed less of the workflow than we originally anticipated (e.g., the relation extraction).

List any issue you had regarding (interoperability) requirements related to (please use none if you had no relevant issues):

- Licensing.
- Metadata description.
- Registry.
- Input / output restrictions & formats: The output file is in a format that is difficult for a non-specialist to process, making visualization of the output awkward.
- Preparing & Packaging (for Maven, Docker and web services).
- Other:
  - It took a long time to make the wrapped web service run properly on the platform.
  - Uploading of corpus documents for mining and their subsequent processing on the platform is currently rather slow.
- None

The amount of work to adapt and integrate your tools into OMTD was

- as expected
- much bigger
- Other ....

Do you think the effort done is justified?

- yes,
- no,
- to what extend?

List and summarize any lessons learned from this project.

- None

Regarding the use case you presented, list the aspects that OMTD can (or could) benefit the use case.

- While targeted at database curators, the scenario will eventually benefit the life-science community at large, as it will lead to a much more dynamic and systematic accumulation of data in widely used chemistry and metabolism databases around the globe. The TDM solution could be used as a stand-alone workflow for individual researchers to extract the same or similar types of data from a smaller, specialized corpus.

List what you would have done differently and/or your recommendations for improvement.
- Speed up corpus creation time
- Speed up time it takes to run workflows (or give user an estimated completion time)
- Reduce metadata requirements for the registry - these are difficult for technical minded people, so will be difficult for non-technical users. Possible these could be pre-filled with sensible values.
- Provide export functions of the output
Use case evaluation report: LS-B Text mining for curation of NeuroScience Literature

List the highlights and key success factors of the project:

- A corpus of sentences coming from scientific papers has been manually annotated by EPFL with nine types of entities: Neurons, Brain Regions, Scientific Values, Scientific Units, Ionic Currents, Ionic Channels, Ionic Conductances, Synapse, and Model Organisms. Another corpus has been manually annotated by EPFL with five types of modelling parameters: Neuron Density, Maximal Ionic Conductance, Resting Membrane Potential, Volume of Brain Region, and Ohmic Input Resistance.
- UNIMAN has published an application that can automatically identify the above-mentioned entities in text as a standalone web service as well as a component via the OpenMinTeD platform. Output of the application can be used to enhance the productivity of NeuroCurator.
- UNIMAN has also developed a workflow to normalize four categories of entities (Neuron, Brain Region, Ionic Current, and Model Organisms) to the NIFSTD ontology. The workflow will be registered into the OpenMinTeD platform soon.
- EPFL has had some preliminary results on integrating the text mining output into NeuroCurator.

List and identify anything you found useful:

- We had opportunities to work closely with the community to make TDM applications practical and useful.
- The registry makes our web service more discoverable, promotes reuse, etc.
- Potential to use other services for neuroscience, should they be made available on the platform.

List and describe any unexpected events

(unexpected events that occurred during the project and the impact that those events may have had on the project and the action(s) taken to address them if any. Common problems are normally due to (please identify)

- the complexity of the task,
- lack of documentation,
- lack of suitable supporting tools and/or examples,
- over-ambitious project aims,
- time schedule,
- other …

We did not anticipate the complexity in implementing an interoperable platform. The release of the platform was significantly delayed, leading to difficulties in reaching the key objectives identified for this use case. Especially, the first output to be considered to be integrated into NeuroCurator was available

28 A desktop application (GUI) to perform systematic and collaborative curation of neuroscientific literature, created at EPFL in the context of the Blue Brain Project. [https://github.com/BlueBrain/neurocurator](https://github.com/BlueBrain/neurocurator)
only at the beginning of April. Before that time, the data model, the file format, and the annotations themselves were uncertain.

List any issue you had regarding (interoperability) requirements related to

(please use none if you had no relevant issues):

- Licensing:
  - Intellectual property rights make the collaborative curation of corpora annotated by the OpenMinTeD platform very restricted.
- Metadata description.
- Registry.
- Input / output restrictions & formats.
  - Scientific articles as PDF files have several challenges. The extracted text (1) is not always in the correct reading order, (2) doesn’t separate the main text from the figure texts or the table contents, and (3) is different from one software library to another. These challenges lead to several consequences. The most critical for this use case is the difficulty to directly highlight the annotations made by the OpenMinTeD platform on the corresponding PDFs. EPFL needs this information in their literature curation workflow based on the desktop application NeuroCurator.
  - The OpenMinTeD component recognizing the entities and NeuroCurator work on extracted texts which are different for the same PDF. This is because two different software libraries are used. The OpenMinTeD component extracting the text doesn’t provide in the annotated corpora the position on the page of the words part of an annotation. NeuroCurator uses a software library corresponding to its ecosystem (Qt) which gives this information.
  - The annotations performed by OpenMinTeD have small spans (typical output of NERs), which makes their location even more difficult because it leads to ambiguity. One identified solution to solve this issue is to add context around the annotation (i.e., text before and after). This solution is challenged by the fact that the surrounding text might be different between extracted texts.
- Preparing & Packaging (for Maven, Docker and web services).
- Other:
  - It took a long time to make the wrapped web service run properly on the platform.
  - Uploading of corpus documents for mining and their subsequent processing on the platform is currently rather slow.
- None

The amount of work to adapt and integrate your tools into OMTD was

- as expected
- much bigger
- Other …. 

Do you think the effort done is justified?

- yes,
- no,
List and summarize any lessons learned from this project.
- None

Regarding the use case you presented, list the aspects that OMTD can (or could) benefit the use case.
- Blue Brain Project researchers and neuroscientists can use the TDM solution to speed up the manual literature curation required for their research.

List what you would have done differently and/or your recommendations for improvement:
- Speed up corpus creation time
- Speed up time it takes to run workflows (or give user an estimated completion time)
- Reduce metadata requirements for the registry. These are already difficult for technical minded people, so they will be difficult for non-technical users. These could possibly be pre-filled with sensible values.
- Provide export functions of the output
Use Case evaluation report: SS-A: Facilitation of complex information linking and retrieval from social sciences publications

List the highlights and key success factors of the project:

Built a NER model for the social sciences domain on data that was annotated in the project. Packaged and deployed the model to a Maven repository such that it can be used on the OpenMinTeD platform. Enhanced DKPro Core components to be able to make use of this model on the OpenMinTeD platform. Updated and released several DKPro projects which are dependencies of the SSH components. Implemented and evaluated Variable Mention Detection and Linking component. Implemented and evaluated Keyword Assignment component (based on Maui).

List and identify anything you found useful:

The project motivated several improvements and updates in various DKPro projects, some of which had been dormant for a while. The deployment of the use case components contributed to the testing of the OpenMinTeD platform and uncovered various issues which could be fixed.

List and describe any unexpected events

(unexpected events that occurred during the project and the impact that those events may have had on the project and the action(s) taken to address them if any. Common problems are normally due to (please identify)

- Dependencies on other projects that required updates
- Technical problems at various levels trying to deploy the components to the OpenMinTeD platform
- Project participants changed during the course of the project causing some loss of time and knowledge

List any issue you had regarding (interoperability) requirements related to

(please use none if you had no relevant issues):

- Many attempts at packaging, deploying, and running where necessary due to bugs, unconsidered cases or edge cases in the specifications. However, all could be resolved.

The amount of work to adapt and integrate your tools into OMTD was

- More than expected

Do you think the effort done is justified?
Yes. It causes improvements to the components that were integrated, to dependencies of the components that were integrated, and to the platform itself.

List and summarize any lessons learned from this project.

- This project shows once more that it is necessary of developers to put themselves into the roles of users and test the system as early as possible.

Regarding the use case you presented, list the aspects that OMTD can (or could) benefit the use case.

- OpenMinTeD permits easy access to and use of the deployed components.

List what you would have done differently and/or your recommendations for improvement:

- Start testing earlier.
Use Case evaluation reports: SC-B Research Publication Recommendation System and SC-C Research Excellent Trends Explorer

List the highlights and key success factors of the project:

A key success of the project is the creation of a platform that both data sources and “algorithms” live in the same place. This enables to avoid loading/unloading overhead, collection struggles and interoperability headaches (how to make components work together)

List and identify anything you found useful:

Dr. Inventor Text mining toolkit (one of the external Use cases) is something that intrigued our team and we are looking ways of how to incorporate it to our pipelines. The fact that they will be ready components of it in OMTD platform gives us a huge head start and the possibility to use it from OMTD directly.

List and describe any unexpected events

Our task mainly composed of “translating” our existing services and packaging them in OpenMinTed schema. This was not particularly straightforward as we had to handle strenuous format (CAS) but eventually we found open solutions to resolve these issues.

Documentation of schema, web services schema, docker format is complete, perhaps what was missing was an example tutorial demonstrating the whole process how to package components (wrapping up, execute, test, validate outputs, ..)

List any issue you had regarding (interoperability) requirements related to

- Licensing: we used default open licenses (Apache 2.0)
- Metadata description: registration of components via the xml editor, no validation errors
- Registry: none
- Input / output restrictions & formats: none (after a few iterations of getting it right)
- Preparing & Packaging (for Maven, Docker and web services) none (few iterations needed but that was expected)

The amount of work to adapt and integrate your tools into OMTD was

- as expected: In total two weeks was required (scattered though in a period of three months)

Do you think the effort done is justified?

- yes,

List and summarize any lessons learned from this project.
Interoperability is much more important than people think! Making your components work with others adds a huge value to them.

**Regarding the use case you presented, list the aspects that OMTD can (or could) benefit the use case.**

Our use case components provide external metadata for publications - that cannot be inferred from the publication itself - by invoking external web services and utilizing external datasets. This is essential for applications that require analytical tasks (e.g. evaluating impact of publication via citation count).

**List what you would have done differently and/or your recommendations for improvement:**

As stated above an end-to-end example of how to package a component in either web service or docker image would tremendously benefit external developers to quickly add their components in OMTD registry.
Use Cases evaluation reports: LS-C: Text Mining on Articles Related to Health State Modelling in Chronic Liver Diseases, SC-E: Text Mining of Articles Related to Leica Microscopes Leica Microscope and SC-D: Rock Art Mining.

List the highlights and key success factors of the project:

- Frontiers developed three Scholarly Communication Uses Cases to highlight how articles can be used for Text and Data Mining. Rock Art, uses TDM techniques to find the word ‘Rock Art’ or its synonyms on articles, confirm data methodologies and the position (i.e. latitude & longitude) where the Rock Art was found. Leica Microscope uses TDM techniques to read through the Methodology section of articles to find the Leica Microscope used in the research. HSM uses TDM techniques to build a health state progression on chronic liver diseases.
- The first key success factors in building the use case is the involvement of SME - Subject Matter Experts in the domain. For Rock Art, the SME is the researcher writing a paper on rock art. For Leica, the SME is the marketing manager from Leica. For HSM, the SME is the author of the paper on Health State Modelling and Precision Medicine.
- The second key success factor is access to NLP expert resources. This is critical to guide the work and find the appropriate tools and resources.
- The third key success factor is availability of resources to define the requirements, design, build and system test the components, as well as user test the components with the SMEs.

List and identify anything you found useful:

- Building the use case was a good exercise to understand Text and Data Mining, how publishers enable text and data mining by making sure the articles are discoverable, and deposited in repositories and how publishers can continue to support text and data mining by complying to meta-data standards.
- The use case highlighted the urgency of having access to all the articles. Today there are approximately 50M articles published and every year an estimated 2.5M new articles are published. There are two issues. First, access to the articles are critical. Approximately 95% of articles are behind the paywall. Second, a researcher once quoted the 90-90-90 rule. 90% of researchers time is spent acquiring the article, 90% of the time the article is closed or hidden behind the paywall, and another 90 TB is required to store all of Open Access literature, and that is just the Open Access part

List and describe any unexpected events:

- For the Rock Art Use Case, the complexity of the requirement and task made the use case difficult. It was overambitious to think that we could replace the whole analysis of Rock Art related documents. We had to iterate several times to find a solution that is useful to the researcher. The scoped-down work is to extract meaningful tables and text for the researcher to perform analysis.
- For the Leica Microscope Use Case, the first challenge was agreeing with the SME the appropriate use case to develop. Frontiers wanted to show how TDM could enable citation of non-article based citations (i.e. use of laboratory equipment), but the SME wanted to extract email addresses of authors based on a domain and field for marketing purposes. While Frontiers proved that TDM resources can fulfill this requirement, policies on data privacy inhibited Frontiers from redistributing the results to Leica. In the end, Frontiers decided to build the use case on non-article based citation.
For the HSM Use Case, there are no unexpected events. Frontiers successfully developed a progression model on Chronic Liver Disease with the SME.

List any issue you had regarding (interoperability) requirements related to (please use none if you had no relevant issues):
- On interoperability, the difficulty we encountered is the ability to download the OMTD results to local resources due to licensing restrictions, so we can integrate them with other components. To expound, to be useful OMTD steps can be defined as 1. Find the relevant articles, 2. Annotate the relevant articles, 3. Analyze and consolidate the relevant articles, 4. Summarize and prepare reports, 5. Display results on a user interface. Today, OMTD is only able to do #1 Find the relevant articles and #2 Annotate the relevant articles. Our use case can work with #1 and #2 provided we are able to download the results and upload it into our own resources, so we can execute #3 to #5. Due to licensing limitations, we are unable execute the download, hence our use case delivers very restricted functionality.

The amount of work to adapt and integrate your tools into OMTD was
- Much bigger - it was difficult to build OMTD components when the platform is not yet available. Once the platform became available, it was equally difficult to make the components work on Maven and Docker images.

Do you think the effort done is justified?
- Yes, I agree that the effort done is justified. Building a platform of services is difficult. It is ideal that the platform is ready before components are built on top of the services, but this did not happen.

List and summarize any lessons learned from this project.
- Understand the functionalities that will be delivered. Try to work within that scope.

Regarding the use case you presented, list the aspects that OMTD can (or could) benefit the use case.
- Rock Art is reusable and useful to archaeologists looking for keywords, dating methodologies and latitude/longitude in articles.
- Leica Microscope is useful to scientometricians who want to build metrics for non-article based citations.
- Health State Modelling on Chronic Liver Disease is useful for researchers who want to build models on progressions of chronic diseases.

List what you would have done differently and/or your recommendations for improvement:
- Start on the Scholarly Communication Use-cases much earlier in the project, define the requirement and validate what the OMTD platform can and cannot do.
Use Case: SC-A Research analytics Document Classification

List the highlights and key success factors of the project:

- This is one of the three TDM research analytics use cases that we integrated in OMTD. The Classification application mines the fulltext of publications in order to perform content-based classification using given taxonomies.
- A key success factor has been that this TDM application was initially developed as part of OpenAIRE’s Inference (by mining) workflow to enrich the OpenAIRE information space, and is already actively used as a standalone workflow in several E.U. funding & research evaluation tenders providing useful insight and timely intelligence of systematic research in E.U.
- Several vocabularies for various well-known taxonomies are used as knowledge base, namely: arXiv, MeSH (Medical Subject Headings), ACM (Association for Computing Machinery), and DDC (Dewey Decimal Classification/System).
- If additional taxonomies are required in the future, this will require training a new classifier for the particular vocabulary and retraining the pre-classifier used to select which taxonomies apply to a particular document.

List and identify anything you found useful:

- It is important to provide users with good classification results. Since this TDM module had been developed and tuned over a couple of years as part of the OpenAIRE inference workflow, it has been refined to produce reliable classifications.

List and describe any unexpected events:

(unexpected events that occurred during the project and the impact that those events may have had on the project and the action(s) taken to address them if any. Common problems are normally due to (please identify)

- None

List any issue you had regarding (interoperability) requirements related to

(please use *none* if you had no relevant issues):

- None. The application does not use any external web services.

The amount of work to adapt and integrate your tools into OMTD was

- Much bigger - there was considerable more effort involved that expected in creating and registering the component, combining it into a workflow and integrating the final application.

Do you think the effort done is justified?

- Yes - it is worthwhile having our content-based document classification service running on a platform that many users can reuse for their own needs.

List and summarize any lessons learned from this project.
Regarding the use case you presented, list the aspects that OMTD can (or could) benefit the use case.

- Content-based classification provides information that is useful to researchers and research communities, scholarly communication organisations, research officers, content providers, publishers, repository managers, etc.

List what you would have done differently and/or your recommendations for improvement:

- In the OMTD platform our component needs to be called, started and executed for each PDF of the input corpus, one at a time, instead of batch processing them. Also, the results returned are one JSON file per document, instead of one JSON file with all the results. This makes it inefficient and very time consuming.
Use Case: SC-A Research analytics DataCite linking

List the highlights and key success factors of the project:
- This is one of the three TDM research analytics use cases that we integrated in OMTD. The DataCite linking application mines the fulltext of publications and extracts links to DataCite (https://www.datacite.org). These links are mainly citations to datasets.
- A key success factor has been that this TDM application was initially developed as part of OpenAIRE’s Inference (by mining) workflow to enrich the OpenAIRE information space, and is already actively used as a standalone workflow in several E.U. funding & research evaluation tenders providing useful insight and timely intelligence of systematic research in E.U.
- A key success factor for the this TDM service to continue to be useful is that its database needs to be updated when a new DataCite dump is available.

List and identify anything you found useful:
- It is important to provide users with reliable results. Thankfully, due to the fact that this TDM module had been developed and tuned over many years as part of the OpenAIRE inference workflow, it has been refined to produce links that are 97-98% accurate.

List and describe any unexpected events:
(unexpected events that occurred during the project and the impact that those events may have had on the project and the action(s) taken to address them if any. Common problems are normally due to (please identify)
- None

List any issue you had regarding (interoperability) requirements related to
(please use none if you had no relevant issues):
- None. The application does not use any external web services.

The amount of work to adapt and integrate your tools into OMTD was
- Much bigger - there was considerable more effort involved that expected in creating and registering the component, combining it into a workflow and integrating the final application.

Do you think the effort done is justified?
- Yes - it is worthwhile having our DataCite linking service running on a platform that many users can reuse for their own needs.

List and summarize any lessons learned from this project.

Regarding the use case you presented, list the aspects that OMTD can (or could) benefit the use case.
- DataCite linking provides information that could be useful to researchers and research communities, scholarly communication organisations, research officers, content providers, publishers, repository managers, etc.
List what you would have done differently and/or your recommendations for improvement:

- In the OMTD platform our component needs to be called, started and executed for each PDF of the input corpus, one at a time, instead of batch processing them. Also, the results returned are one JSON file per document, instead of one JSON file with all the results. This makes it inefficient and very time consuming.
Use Case: SC-A Research analytics Funding Mining Services

List the highlights and key success factors of the project:

● This is one of the three TDM research analytics use cases that we integrated in OMTD. The Funding Mining application mines the fulltext of research publications and extracts links to projects.

● In order for the Funding Mining application to work, it requires updated project lists from supported funders (i.e. list of project numbers/identifiers/grants, project acronyms if available, etc.) We currently support the projects from the following funders: EC (FP7/H2020), NSF (National Science Foundation, USA), NIH (National Institute of Health, USA), Wellcome Trust, FCT (Fundação para a Ciência e a Tecnologia, Portugal), ARC (Australian Research Council), NHMRC (National Health and Medical Research Council, Australia), CSF/HRZZ (Hrvatska Zavlada Za Znanost, Croatia), MSES-MZOS (Ministarstvo Znanost, Obrazovanja i športa, Croatia), SFI (Science foundation Ireland), NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek, Netherlands), but new funders are added regularly.

● A key success factor has been that this TDM application was initially developed as part of OpenAIRE’s Inference (by mining) workflow to enrich the OpenAIRE information space, and is already actively used as a standalone workflow in several E.U. funding & research evaluation tenders providing useful insight and timely intelligence of systematic research in E.U.

● A key success factor for the this TDM service to continue to be useful is that project lists need periodic updating. The frequency of the updates depends on the frequency of calls of each funder; usually 2-3 times per year. This update requires some manual work to download and insert the new projects (or the new funders) in the existing SQLite DB.

● Finally, the more funders we include, the more useful the application becomes to users. However, the application needs updating when a new funder is included. This happens because we tune the algorithm to mine the new funder’s projects with higher accuracy. This is a manual process which happens when there are new funders. Which new funders are included is currently been driven by the requirements of the OpenAIRE project.

List and identify anything you found useful:

● It is important to provide users with reliable results. Thankfully, due to the fact that this TDM module had been developed and tuned over many years as part of the OpenAIRE inference workflow, it has been refined to produce link that are more than 99% accurate.

List and describe any unexpected events:
(unexpected events that occurred during the project and the impact that those events may have had on the project and the action(s) taken to address them if any. Common problems are normally due to (please identify)

● None

List any issue you had regarding (interoperability) requirements related to
(please use *none* if you had no relevant issues):

- None. The application does not use any external web services.

**The amount of work to adapt and integrate your tools into OMTD was**

- Much bigger - there was considerable more effort involved that expected in creating and registering the component, combining it into a workflow and integrating the final application.

**Do you think the effort done is justified?**

- Yes - it is worthwhile having our project funding mining service running on a platform that many users can reuse for their own needs.

**List and summarize any lessons learned from this project.**

Regarding the use case you presented, list the aspects that OMTD can (or could) benefit the use case.

- Funding mining provides useful information especially for funders and policy makers in deducing research outputs of funded projects. In combination with other information, funding mining can be used for different scientometric analysis.

**List what you would have done differently and/or your recommendations for improvement:**

- In the OMTD platform our component needs to be called, started and executed for each PDF of the input corpus, one at a time, instead of batch processing them. Also, the results returned are one JSON file per document, instead of one JSON file with all the results. This makes it inefficient and very time consuming. As an example, running it on a sample 2 PDF corpus requires almost 3 mins running on the OMTD platform, whereas on its own it requires just 5-6ms processing time.

*Annotation performance form*
OMTD evaluation phase II: Gold Standards

Questionnaire for the OMTD community use case evaluation phase 1 & 2: to get information about Gold Standards produced/used. Only one member per use case should complete this questionnaire.

Important: You should fill in this form by 30th November 2017.

*Required

1. Provide your full name: *

2. Provide your institution name: *

3. Provide your contact e-mail: *

4. Provide your community use case identifier: *

   Mark only one oval.

   - AS-A
   - AS-B
   - AS-C
   - AS-D
   - AS-E
   - LS-A
   - LS-B
   - SS-A
   - SC-A
   - Other:

5. If "other", specify it:

6. Select the types of components or tasks that are relevant to your evaluation scenario? *

   Tick all that apply.

   - Named entity recognition
   - Automatic term recognition
   - Term/entity grounding to ontologies/thesauri
   - Event extraction
   - Other:

7. Briefly describe the evaluation scenario *

   ________________________________________________________________
   ________________________________________________________________
   ________________________________________________________________
   ________________________________________________________________

Gold Standard (identification information)

If you use more than one Gold Standard, please fill different forms
8. Name: *

9. Citation: *

10. URL:

11. Identifier:

12. Language: *

13. Short description of the GS: *

14. Status
   Mark only one oval.
   - Released
   - Unreleased

15. License:

16. Author(s):

17. Maintainer(s):

Gold Standard (creation info)

18. Corpus creation/source: *
   Tick all that apply.
   - Manually created by the the use case team
   - Automatically created
   - Provided by an "evaluation challenge/task"
   - Other:

19. Document types: *
   Tick all that apply.
   - Titles and abstracts
   - Full text
   - Journals
   - Other:
20. URL of the "evaluation challenge task" (if any)

21. Name of the annotation tool used (Brat,...):

22. Document selection criteria (Random, journal-based, keyword/text search based, ...)

23. Annotation tool URL:

24. Is there a training set?
   *Mark only one oval.*
   - Yes
   - No

25. Is there a test set?
   *Mark only one oval.*
   - Yes
   - No

26. Annotation guidelines URL:

27. Was the corpus consistency measured? *
   *Mark only one oval.*
   - Yes
   - No

28. Inter-annotator agreement:

---

**Gold Standard (corpus info)**

29. Annotation format (Json, BioC,...): *

30. How many entity annotation types?

31. How many entity annotations?

32. How many annotation relation types?
33. How many annotation relations?

34. List of entity name(s) (and link to ontology / vocabulary):

35. List of relation name(s) (and link to ontology / vocabulary):

36. Mime Type:

37. Size (bytes):

38. Size (characters):

39. Size (tokens):

40. Size (words):

41. Size (sentences):

42. Size (documents):

**Prediction data**

43. Input document format (HTML, PDF, XML, ..): *

44. Prediction output format *

45. Prediction data url (in case you ran an evaluation test)
46. URL to a collection input dataset (to be used by an evaluation task)

**Prediction results**
In case you ran an evaluation task, please provide the following information

47. **Precision & score:**

48. **Recall & score:**

49. **Accuracy & score:**

50. **F-Measure & score:**

51. **Other metrics & scores:**

**Participation in a shared task**
If you participated in a shared task, please indicate:

52. **Name of shared task:**

53. **URL:**

54. **Comments**