EOPEN

opEn interOperable Platform for unified access and analysis of Earth observatioN data

H2020-776019

D7.1

Pilots implementation and 1st prototype evaluation report

Dissemination level:	Public
Contractual date of delivery:	Month 20, 30/06/2019
Actual date of delivery:	Month 21, 05/07/2019
Workpackage:	WP7 Evaluation and pilot implementation
Task:	T7.1 Pilot implementation and evaluation
Type:	Report
Approval Status:	Final Draft
Version:	7
Number of pages:	78
Filename:	D7.1- Pilots implementation and 1st prototype

	evaluation report _2017-11-01_v0 7_subm.docx

Abstract

This deliverable reports on the pilots' implementation and the 1st platform prototype evaluation through the Pilot Use Cases (PUCs). This deliverable is under WP7 and Task 7.1 Pilot implementation and evaluations and it is the first of two iterations (D7.1 and D7.2). In this first iteration we establish the evaluation methodology and report on the current status of implementation and the preliminary evaluation results. The documents starts with setting the framework of evaluation that defines i) what me measure, ii) who evaluates and iii) which is the implementation-evaluation plan. Then the scientific and technologies activities (Key Results- KRs) of the project are individually reported with respect to their established Key Performance Indicators (KPIs). Subsequently the PUCs report on the current status of implementation and set the timeline for completion. The PUCs are then positioned within the EOPEN framework, elaborating on how its generic modules are utilized in dedicated workflows to address tangible user needs. A set of potential test scenarios of evaluation are suggested for the pilot phase of the PUCs.

co-funded by the European Union

History

Version	Date	Reason	Revised by	Approved By
1	27/06/2019	Initial Draft		
2	01/06/2019	Further Contributions		
3	02/06/2019	Further Contributions		
4	03/06/2019	Further Contributions		
5	04/06/2019	Further Contributions	Ioannis Papoutsis	Vasileios Sitokonstantinou
6	04/06/2019	Review	Gabriella Scarpino, Dennis Hoppe, Silvia Scifoni (Serco)	Vasileios Sitokonstantinou
7	05/06/2019	Final Draft	Vasileios Sitokonstantinou	Ioannis Papoutsis

Author list

Organization	Name	Contact Information
NOA	Vasileios Sitokonstantinou	vsito@noa.gr
NOA	Ioannis Papoutsis	ipapoutsis@noa.gr
CERTH	Stelios Andreadis	andreadisst@iti.gr
AAWA	Francesco Zaffanella	francesco.zaffanella@distrettoalpiorientali.it
AAWA	Francesca Lombardo	francesca.lombardo@distrettoalpiorientali.it
FMI	Petteri Karsisto	petteri.karsisto@fmi.fi
SPACEAPPS	Bernard Valentin	bernard.valentin@spaceapplications.com
HLRS	Li Zhong	li.zhong@hlrs.de

Executive Summary

This deliverable, D7.1 *'Pilots implementation and 1st prototype evaluation report'*, is the first of two parts for the Task 7.1: '*Pilot implementation and evaluation'*. This document reports on the thus far implementation of EOPEN pilots and evaluates the first platform prototype both based on the individual Key Results (KRs) and from PUC perspective.

The document initially reports on the different scientific and technological KRs of the platform. Each KR subchapter includes the scope of the activity, the scientific framework and implementation, the Key Performance Indicators (KPIs) and finally a discussion subsection positioning the KR within the context of the PUCs. Then the three different PUCs are individually reported in Chapter 5, including a brief recap of the use case story (D2.1), the timeline of implementation, the monitoring of implementation against the user requirements, the implementation specifics – i.e. the developed products and services and the exploitation of EOPEN KRs within the PUC implementation workflow. Finally, potential test scenarios are suggested for the pilot phase of the PUCs.

The purpose of the EOPEN Pilot Use Cases (PUCs) is to demonstrate the potential of the 1st prototype of the EOPEN platform, through case specific workflows, utilizing its components (scientific and technological) to address tangible user needs. The evaluation of the 1st prototype of the EOPEN platform is performed in the context of the 3 PUCs. The evaluation methodology is designed based on four evaluation axes and that is the alignment of implementations with i) the call requirements, ii) the scientific and iii) technological requirements of the project, as specified in the GA, and finally with iv) the collected user requirements, as described in deliverable *D2.2: User Requirements*.

Abbreviations and Acronyms

AOI	Area Of Interest
API	Application Programming Interface
ARI	Adjusted Rand Index
COM	Current Operating Model
DCNN	Deep Convolutional Neural Network
DEM	Digital Elevation Model
DSM	Digital Surface Model
DSM	Digital Surface Model
GA	Grant Agreement
HPC	High Performance Computing
HPDA	High Performance Data Analytics
IoU	Intersection over Union
KPI	Key Performance Indicator
LDA	Latent Dirichlet Allocation
MNDWI	Modified Normalized Difference Water Index
NER	Name Entity Recognition
NMI	Normalized Mutual Information
OA	Overall Accuracy
OWL-DL	Ontology Web Language Description Language
PA	Producer's Accuracy
PUC	Pilot Use Case
RDF	Resource Description Framework
RF	Random Forests
RNN	Recurrent Neural Networks
SAR	Synthetic Aperture Radar
SWIR	Short-Wave InfraRed
TOM	Target Operating Model
UA	User's Accuracy
WFS	Web Feature Service
WMS	Web Map Service
WMTS	Web Map Tile Service

Table of Contents

1 INTRODUCTION

The document is structured as follows. In the second chapter a high level description of the evaluation methodology is provided, setting the questions to be answered in the coming chapters, i.e. of *what we measure, who measures, how we measure.* According to the GA, EOPEN brings in certain scientific and technological activities, established as Key Results (KRs) of the project. The third and fourth chapters elaborate on all scientific and technological activities with respect to their measurable performance indicators but also their exploitation within the context of the PUCs. The Key Performance Indicator (KPI) tables, in these sections, will be fully completed in D7.2. In the fifth chapter, we introduce each PUC individually, stating the respective use case stories and the problem to be tackled; and reporting on the implementation status. Some preliminary evaluation results are also provided, along with a list of potential test scenarios that will be fully reported in D7.2. In chapter six we introduce the four different evaluator categories and their role in the overall process; while in chapter seven we describe how the evaluation will be completed in the two deliverables D7.1 and D7.2. Finally, in chapter nine we describe how in the second iteration of the deliverable (D7.2) we will extract and then report lessons learnt, best practices and recommendations for future users of the platform. In this section we will also provide the high level qualitative and quantitative evaluation of the EOPEN platform against the requirements of the call, as seen in the evaluation framework described in Chapter 2.

2 EVALUATION FRAMEWORK

The overall evaluation framework for the 1st and 2nd prototype of the platform and the PUCs is elaborated below. Part of this holistic framework is addressed within this document; while the overall evaluation process will be concluded with D7.2 in M28. In order to perform the evaluation of the EOPEN platform we identify the measures, the tools and the people, as described below.

1. **Evaluation measures –** Quantitative and qualitative measures based on the described KPIs in the GA and the collected user requirements. Below are described the four main axes of evaluation.
   a. **High level Call requirements** –Once the PUCs have been completed we will evaluate the EOPEN platform at higher level, identifying its key competences – showcased through the PUCs - and how these explicitly address the call requirements, such as the ones indicated below.
      i. Enable value adding services on generic data and information storage and processing facilities which can allow public and commercial users effective production environment to interact with and serve their user base without deploying their own storage and processing facilities.
      ii. Make access to the Copernicus data and information easy and user friendly through scalable dissemination and exploitation software based on international standards.
      iii. Foster the establishment of interoperable access facilities to all EU Member States.
      iv. Provide user community tools including best-practices.
      v. Optimize the use of Copernicus data by non-traditional user communities to meet societal challenges.

   b. **Scientific requirements** – Showcase how the following scientific activities of the EOPEN platform were exploited for the purposes of the PUCs. Deliverables 7.1 and 7.2 report on the individual implementations of scientific KRs but also establish and monitor pertinent performance indicators.
      i. Knowledge discovery and content extraction
         1. Change detection in EO data
         2. Concept and Event Detection in Social Media
      ii. Multimodal clustering and retrieval
         1. Multimodal fusion from diverse data sources for information retrieval
         2. Clustering of EO and non EO data
      iii. Community detection and tracking
         1. Animation of the network of user communities
      iv. Knowledge representation and ontology construction for decision making
         1. Ontology construction and reasoning support
         2. Linking of open EO data

c. **Technological requirements** – Showcase how the following technological activities of the EOPEN platform were exploited for the purposes of the PUCs. Deliverables 7.1 and 7.2 report on the individual implementations of technological KRs but also establish and monitor pertinent performance indicators.
   - EO and non-EO data acquisition
     - EO data acquisition
     - Social media crawling
     - Meteo data acquisition
   - Standardized interfaces
   - High Performance Computing
   - Fast and Secure Data Exchange
   - EOPEN platform
   - DIAS - EOPEN would represent a typical Third Party User interested to exploit the DIAS services to design and operate a Front End EOPEN service.

d. **PUC Requirements**
   - **PUC Implementation and Evaluation**
     - Objectives – problem statement and suggested EOPEN solution
     - Users involved – main PUC stakeholders and their role in implementation and/or evaluation
     - Timeline of implementation – record achieved and monitor forthcoming milestones
     - Potential test scenarios – establish testing and evaluation scenarios to take place during the pilot phase of PUCs
     - Feedback mechanism – introduce the methods of receiving feedback from the EOPEN platform users and the pertinent scheduled events (i.e. workshops, info days etc.)

2. Evaluators

a. *Self-assessment* – This refers to the performance indicators, qualitative and quantitative, set and monitored by the developers of the product/module/service in question.

b. *Cross-partner assessment (peer)* – Partners of the consortium that use but have not developed the Key Result in question evaluate their experience. This can be done either by re-evaluating the self-assessment KPIs, if applicable, or setting other relevant, more user oriented ones.

c. *EOPEN user ecosystem* – The end users partners of the consortium (AAWA) and the PUC engaged stakeholders will evaluate the EOPEN platform through their feedback in the form of questionnaires, Q&A sessions, deliverables (in the case of AAWA) etc. that will be based on platform demonstrations and/or hands-on exercises.

d. 3rd party users
   - External users – This refers to non PUC specific users of the platform. Their feedback will be extracted in the dedicated dissemination events of the project, such as the South Korea ACRS workshop and ‘Infoday’ of PUC1

- Other projects of the H2020-EO-2017 call

  A hackathon has been organized (7-8/11/2019) among five consortia to test the individual platforms (EOPEN, BETTER, CANDELA, openEO, Perceptive Sentinel).

3. Implementation - Evaluation plan

- Two iterations
  - Iteration 1 – This first deliverable (D7.1) of Task 7.1 focuses more on the self-assessment of the individual scientific and technological KRs of the platform and establishes the overall evaluation methodology (M20)
  - Iteration 2 – D7.2: The second deliverable (D7.2) of Task 7.1 will more user-oriented, strengthening the evaluation from the PUC perspective as well, while fulfilling the overall evaluation framework (M28) established in this document.

3 SCIENTIFIC ACTIVITIES

Key Result preliminary evaluation and establishment of KPI[1]s to monitor

3.1 Change detection in EO data (KR01)

3.1.1 Purpose of Scientific Activity

The Remote Sensing (RS) based Change Detection module's aim is to identify differences in the state of land features for the detection of flooded areas (PUC1), rice fields (PUC2) and passable roads (PUC3).

The first module (PUC1) of this activity maps the extent of flood events based on satellite images either of Optical or Radar origin (mainly Sentinel-1 and Sentinel-2). A baseline approach has been developed, directly estimating the water-bodies of an input satellite image[2]. The second module (PUC3) takes as input a Sentinel-2 image inferring whether a road is not blocked by a major flood or snow, thus is passable. The third module (PUC2) refers to the land cover map update (rice map), relevant to the PUC 2 implementation. In this module past rice maps, which are freely available, are updated to reflect the truth for the year of inspection, through a change detection workflow using Sentinel-1 data. It essentially refers to an outlier detection method to prepare a refined training dataset for the subsequent machine learning based rice classification (see PUC 2 Implementation section).

The first version of the flood detection (PUC1) module has been successfully deployed as a process on the EOPEN platform and is currently running, generating water-masks of the Vicenza area, that are shown on the website using a Web Map Server. The current version of the road passability (PUC3) module runs on CERTH's side, while land cover map update (PUC2) module runs on NOA's side. Some preliminary information on the Change Detection has been included in D4.1, while more details will follow in D4.4.

Currently, for the evaluation of the flood monitoring module, 10m pixel resolution images of the Copernicus Sentinel Program were used. In the future, adaptations will be performed to support Very High Resolution Images. For the evaluation of the road passability module, the images provided by MediaEval 2018 Satellite Task "Emergency Response for Flooding Events" were used.

3.1.2 Description of the scientific framework

In the first module (PUC1), in order to estimate the water body areas, two different approaches were developed. The first implementation utilizes Sentinel-2 data and the Modified Normalized Difference Water Index (MNDWI), based on the Green and SWIR bands. The second implementation uses Sentinel-1 data. In this case, pre-processing via SNAP is required. Finally, a GeoTiff raster with the water-bodies mask is generated. Future work includes the improvement of model's performance, especially for Sentinel-1 data by

[1] The KPI tables in Section 2 include the field 'relevance to user requirements'. The user requirement IDs included can be found in the ANNEX of D2.2., which is a living a document and will be soon resubmitted.

[2] For more information see D4.1

using Digital Elevation Model (DEM) files and filtering the falsely detected water areas due to high slope.

In the implementation of the second module (PUC3), pre-trained VGG-19 DCNN was used to further train the prediction model. New coming images are fed to the model to infer weather they are passable or not. Both modules are driven by a process (Umbrella API[3]) that searches for Sentinel-1 and Sentinel-2 products of the previous day for the desired Areas of Interest (AoI) and is executed on a daily basis.

In the implementation of the third module (PUC2), the VH backscatter (Sentinel-1) difference between the year of inspection and the year of the land cover map release is used. It is then combined with the water mask produced by the first module to remove obvious outliers, namely parcels of the rice map that are not cultivating rice in the year of inspection. For more information on the implementation of all three modules can be found in D4.1.

3.1.3 Results and metrics

An evaluation of the baseline approach has been made. Main metrics for the accuracy are the Intersection over Union and the F1-Score which combines precision and recall, whereas future evaluation of the change detection module will include the Kappa Coefficient (Table 1). Since, this is the baseline approach the time performance is not yet measured.

Three Lakes of Italy were monitored (Garda, Maggiore, Trasimeno). Due to lack of existing ground-truth in the examined PUCs, a first step towards the evaluation was to generate ground-truth annotations through manual annotation. The shapefiles of the permanent Lakes and Rivers of Italy were used as a base for the annotation. Then recent high resolution images of Open Street Maps, Google Maps and Bing Maps were used to correct the initial shapefiles. Evaluation results for the flood assessment can be found in Table 2.

For the evaluation of the road passability module (PUC3), multiple pre-trained models were tested, with VGG-19 demonstrating the best overall performance (Table 3), which is 76.6% for the validation set and can be considered a satisfactory result.

Table 1: Flood-detection related metrics

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Precision	%	US_V_02 US_V_02b	The model avoids misclassifying water areas as non-water, rice fields as non-rice
Recall	%		The model detects all water areas and rice paddies respectively

[3] T3.1 EO data acquisition service – see D3.1 for more information

F-score	%		The model performs well for imbalanced classes (water and non-water, rice and non rice)
Accuracy	%		
IoU	%		The detected water areas matches the actual water areas The detected passable /non-passable tiles match the actual ones
K	%		kappa index-of-agreement
Time performance	millisecond (ms)		

Table 2: Performance evaluation of two methodologies – Sentinel-1 based and Sentinel-2 based - over Italian Lakes

Lake	Sentinel-1 F1-Score	Sentinel-1 IoU	Sentinel-2 F1-Score	Sentinel-2 IoU	Date
Garda	91.9	85.0	98.2	96.4	30/03/2019
Maggiore	90.3	82.3	97.1	94.3	23/03/2019
Trasimeno	62.3	45.3	97.5	95.2	21/03/2019

Table 3: Two of the best Neural Networks achieved validation set accuracy

DCNN	Learning Rate	Optimizer	Batch Size	Development Set Accuracy	Validation Set Accuracy
VGG-19	0,01	Adam	128	0.8610	**0.7667**
ResNeet-50	0,001	Adam	256	0.8820	0.7323

The evaluation of the third module (land cover map update – PUC2) can be twofold:

1) Evaluate using photointerpretation to determine if identified outliers do not indeed represent land covered by rice. Photointerpretation is performed against a pseudocolor RGB composite VH, VV, VH/VV backscatter products (Sentinel-1) and the true colour composite of 10th September 2018 (Sentinel-2) when differences between rice and non-rice fields are visually maximum.

2) Evaluate if indeed the updated training set offers better accuracy results for the machine-learning based rice paddy classification.

3.1.4 Discussion

This module monitors the flood levels of the selected locations exploiting the available Copernicus products. In the future, monitored areas will be fetched via search criteria, providing the end users the ability to monitor new locations. This module also specifically addresses the needs of PUC 2 via preparing refined training datasets for rice classification.

3.2 Event detection in social media (KR02)

3.2.1 Purpose of Scientific Activity

The Event Detection module aims to identify possible events based on social media or weather data. The current implementation focuses on the crawled tweets, derived from the social media monitoring task, and uses the fluctuation of the number of relevant tweets per day to discover events. At this point a respective notification is produced, but a future step is to provide more insights on each event, such as the location of the happening and some representative keywords. The first version of the module has been successfully deployed as a process on the EOPEN platform and is currently running, generating JSON files with notifications for any detected event. Some initial information about Event Detection has been included in D4.1, while more details will follow in D4.4.

3.2.2 Description of the scientific framework

In order to determine whether an event occurs in a given date, we calculate the *z-score* of that date's number of relevant tweets. The z-score indicates how many standard deviations this day is from the sample's mean, assuming a Gaussian distribution, where sample is the previous thirty days (previous month). If a date's z-score is above a constant threshold, then we can assume there is an event on that day. Bibliography suggests a threshold of 2.5-3.5 and after internal experiments we have settled to a threshold of 3[4].

3.2.3 Results and metrics

At the time of writing the evaluation of the Event Detection module has not been carried out yet. Due to lack of ground-truth and domain-specific datasets in the examined PUCs, a first step towards the evaluation is to generate ground-truth annotations through manual annotation. Afterwards, the following key performance indicators will be calculated:

Table 4: Performance indicators and scientific metrics for event detection in social media

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Precision	%	US_V_02a4	
Recall	%	US_V_02a5	
F-score	%	US_V_08a	
Time performance	millisecond (ms)	US_F_02d1	

[4] https://en.wikipedia.org/wiki/68%E2%80%9395%E2%80%9399.7_rule

3.2.4 Discussion

This module discovers possible events exploiting the social media data that has been crawled by the corresponding task. Each dataset is use case-specific by defining different search criteria during the crawling procedure, so events are detected per each use case separately and independently. In detail, events about floods refer to PUC1; events about snow coverage refer to PUC3 and events about food security to PUC2. Every end user that is interested to a certain use case will be only informed about events relevant to that use case. Future event detection on weather data will focus mainly on flood and snow prediction (PUC1, PUC3).

3.3 Concept extraction in social media (KR03)

3.3.1 Purpose of Scientific Activity

The Concept Detection module aims to extract high-level content (i.e. concepts) from textual and visual low-level information in order to be able to retrieve relevant content and to mark multimodal content as relevant or not to a target event (i.e. flood, heatwave). The current implementation focuses on extracting concepts from the images included in the crawled tweets, derived from the social media monitoring task. At a later stage, textual information will be also used for extracting concepts out of Twitter texts. This first version of the module runs on CERTH's side, but all the information is stored in a database that the platform has access to. Some initial information about Concept Detection has been included in D4.1, while more details will follow in D4.4.

3.3.2 Description of the scientific framework

In order to detect the concepts in a non-EO image, we trained a 22-layer GoogleNet network on 5055 ImageNet concepts, which means that the initial classification layer of the trained network which is a fully connected layer had dimension equal to 5055. In the sequel, in order to obtain as output the probabilities of 345 SIN TRECVID concepts, we fine-tuned the network which involves replacing the classification layer, also fully connected, with one that has dimensionality equal to 345. Finally, we added an extra fully connected layer before the classification layer, as it seems to boost its performance.

3.3.3 Results and metrics

The proposed concept detection methodology has been evaluated in TRECVID datasets but not inside the EOPEN framework, which is due to the fact that it requires annotation of significant large number of images and concepts. Therefore, given that the manual annotation of an EOPEN image collection is labour-intensive, the evaluation will be limited on existing annotated datasets such as TRECVID.

Table 5: Performance indicators and scientific metrics for concept extraction in social media

Indicator/Scientific metric	Value	Relevance to User Requirements
Mean Extended Inferred Average	30.63%	US_V_02a1

Precision		US_V_02a2 US_V_02a5 US_F_01b1 US_F_01b3

Other metrics that will be considered in next versions will be the Mean Average Precision, Accuracy and F-score.

3.3.4 Discussion

This module assigns high level concepts to images which are collected by the social media crawler (KR12), and it allows filtering of tweets according to specific concepts (i.e. flood, snow) by using the tweet images. Given that the content of social media images differs significantly per PUC, this module allows the retrieval of visually similar images, and thus images of a specific PUC.

3.4 Similarity fusion (KR04)

3.4.1 Purpose of Scientific Activity

The Similarity Fusion module aims to retrieve similar content found inside the EOPEN collection. Given that the module considers different modalities for retrieving similar EO and non-EO content, there is the need to fuse these multimodal objects in a scalable way, taking into account memory and computational complexity, in order to retrieve similar content in response to a query. The current implementation is a single modality retrieval module for non-EO data which are described by several modalities (i.e. visual, textual, spatiotemporal, and concepts). At a later stage, the fusion of EO data and non-EO data as well the transfer of the service to the EOPEN platform will be realized. This first version of the module runs on CERTH's side. Some initial information about Similarity Fusion has been included in D4.1, while more details will follow in D4.3.

3.4.2 Description of the scientific framework

The single modalities retrieval modules that have been developed are the following:

Visual concepts-based retrieval: This module is based on the visual concepts and it allows ranking the images by considering the probabilities of the top 10 concepts per image. Fast indexing and retrieval is achieved by inserting in MongoDB these top 10 concepts.

Visual features-based retrieval: Regarding the visual features, they are DCNN-based descriptors and they are the output of the last pooling layer of the fine-tuned GoogleNet architecture described in the visual concepts extraction module. For the Nearest Neighbour search, we create an Asymmetric Distance Computation index and then, we compute the K-Nearest Neighbours from the query image. In order to accelerate the querying and retrieval process, a web service is implemented that loads the index on the RAM memory,and queries the index upon user demand.

Text-based retrieval: Regarding the textual representation of tweets, we used the word2vec representation and the Euclidean distance for retrieving similar tweets to the query tweet.

3.4.3 Results and metrics

At the time of writing the evaluation of the Similarity Fusion module in the EOPEN dataset has not been carried out yet. However, the method proposed has been evaluated in other external datasets such as WIKI11 (Tsikrika, et al. 2010), the IAPR-TC12 (Grubinger, et al. 2006), the MediaEval dataset from the diversity task of 2015 (Ionescu, et al. 2014) and the TRECVID 2011 and 2012 (Over, et al. 2011) datasets. In order to evaluate it in EOPEN, ground-truth and domain-specific datasets in the examined PUCs are required, and thus a first step towards the evaluation is to generate ground-truth annotations through manual annotation. Afterwards, the following key performance indicators will be calculated:

Table 6: Performance indicators and scientific metrics for similarity fusion

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Precision	%	US_V_02a2 US_V_02b3	
Recall	%		
F-score	%		
Processing time	ms		
Memory complexity	GB		

3.4.4 Discussion

This module allows the user to retrieve similar EO and non-EO content given a query tweet or EO product. Given that the tweet content, which involves both the text and image included (if it exists) and the EO products are rather different per use case, the similar content that will be retrieved will differ per use case.

3.5 Clustering of EO data (KR05)

3.5.1 Purpose of Scientific Activity

The purpose of this module is to cluster EO data including satellite images and social media images into groups, which can be used by end users to find the interested patterns. Currently the module has been implemented and tested on the images coming from social media. And the next step is to do fine tuning of this model and apply it on the satellite images. Also the service will be integrated into the EOPEN platform and will be available to end users.

3.5.2 Description of the scientific framework

The module for EO data clustering adopts the technology of deep neural network, specifically autoencoder to do the dimension deduction, so that the key features are extracted. And then the output of autoencoder can be fed to a clustering algorithm, like K-means, based on which the images are clustered into different groups. When trying to visualize the final clusters of the images, t-SNE is employed and the result of image clusters could be shown in both 2D and 3D spaces.

3.5.3 Results and metrics

The proposed EO image clustering methodology has been evaluated internally at HLRS, however, the evaluation only focused on the performance on test data, and the module has not been assessed yet inside the EOPEN platform.

Table 7: Performance indicators and scientific metrics for Clustering of EO data

Indicator/Scientific metric	Value	Value Achieved	Relevance to User Requirements	Comments
Improve Normalized Mutual Information (NMI) by at least 5%	%	N/A	US_V_02a3 US_F_01b US_F_01c US_F_04	
sImprove Adjusted Rand Index (ARI) by at least 5%	%	N/A		
Reduce Processing Time by at least 15%	%	N/A		
Achieve a Likert score of 4 and above in a user-based evaluation	Likert scale (1-5)	N/A		

3.5.4 Discussion

When calling the service for image clustering, the use case has to be defined as input. Then, the module uses the images and feeds to the model that well trained and returns the cluster of this image. End users have the ability to call the service for the use case they are interested in and find interesting patterns of EO data including satellite images and social media images.

3.6 Text clustering of non-EO data (KR06)

3.6.1 Purpose of Scientific Activity

The purpose of this module is to cluster textual information from crawled tweets into groups, which can be perceived as topics that are trending on social media. Currently the module has been implemented, deployed as a service on the EOPEN platform and is available to be called at any time. The service receives as an input the use case of interest (e.g., Italian tweets about flood events) and responds with the clusters detected in the 300 most recent relevant tweets, so as to capture the trending topics. Each topic has the tweets it comprises and some keywords that reflect the most common terms mentioned in these tweets. A first description of the Text Clustering module can be found in D4.1; a more detailed description will be documented in D4.2.

3.6.2 Description of the scientific framework

The module for clustering textual streams of data uses a hybrid density-based framework, which combines the DBSCAN-Martingale process and the well-established Latent Dirichlet Allocation (LDA). After extracting the word *n*-grams of the recent tweets, the robust to noise DBSCAN-Martingale progressively estimates the number of clusters in the given dataset and LDA assigns the tweets to topics.

3.6.3 Results and metrics

Although the proposed text clustering methodology has been evaluated before, the module has not been assessed yet inside the EOPEN framework. On one hand, the evaluation will concern automatic calculations (e.g., NMI, ARI) and, on the other hand, it will involve people, in order to demonstrate that the developed clustering techniques are beneficial to end users.

Table 8 : Performance indicators and scientific metrics for text clustering of non-EO data

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Normalized Mutual Information (NMI)	%	US_V_02a3 US_V_02a5	
Adjusted Rand Index (ARI)	%		
Processing Time	millisecond (ms)		
User-based evaluation	Likert scale (1-5)		

3.6.4 Discussion

When calling the service for text clustering, the use case has to be defined as input. Then, the module uses tweets that have been crawled specifically for this use case and returns the detected topics. End users have the ability to call the service for the use case they are interested in and get informed about the relevant trending topics on social media, e.g. current floods and snow incidents (PUC1, PUC3) or news about food security (PUC2).

3.7 Community detection (KR07)

3.7.1 Purpose of Scientific Activity

The objective of this module is to detect end-user communities through their relationship. The first version of the module focuses on social media accounts and their "following" interaction, while it can be easily adapted to focus on users of the EOPEN platform and the similarity of their behaviour. The module has been deployed on the EOPEN platform as a service and is currently online. The inputs to the service are the use case of interest and a date range. The response includes the pairs of connected users, the detected communities, and a list of the most influential users. The Community Detection module has been described briefly in D4.1, but further information will follow in D4.4.

3.7.2 Description of the scientific framework

The methodology of this module serves two functionalities: the detection of key-communities and the identification of key-players. After mapping users and their relationships to nodes and edges, the Louvain algorithm is adopted to discover communities (Blondel et al., 2008). On the detected community of nodes, the degree measure is computed per node, defined as the number of its neighbours. The nodes with the largest degree are the key-players.

3.7.3 Results and metrics

The Community Detection method will be evaluated in the forthcoming months. In case of available ground truth, i.e. a network where each node's community is known a priori, we are able to compute metrics such as NMI and ARI. Otherwise, we can calculate the modularity of networks, which measures the strength of division of a network into communities. In both cases, the module will also be evaluated by end users in terms of usefulness.

Table 9 : Performance indicators and scientific metrics for community detection

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Modularity	[-1,1]	US_V_02a US_F_01b	
Normalized Mutual Information (NMI)	%		
Adjusted Rand Index (ARI)	%		
Processing Time	millisecond (ms)		
User-based evaluation	Likert scale (1-5)		

3.7.4 Discussion

The implemented service searches for formed communities inside a network of user-to-user interactions, where users are social media accounts whose posts have been collected for EOPEN. Since the collections of crawled tweets are use case-based, the results of the service refer to specific use cases.

3.8 Ontology construction and reasoning support framework (KR08)

3.8.1 Purpose of Scientific Activity

The developed framework provides the appropriate semantic knowledge structures and vocabularies to represent metadata coming from EOPEN modules and other services, like the localisation module, and build intelligent reasoning mechanisms and decision making solutions to extract implicit conceptualisations and interpretations. In the current version of the framework the metadata that are supported come from the localisation and the topic extraction modules. These metadata are transformed from JSON into RDF format and saved in the Knowledge Base. This version of the framework, as well as the Knowledge Base, has been successfully deployed in the EOPEN platform as dockerised services. More information about the module can be found in D5.1.

3.8.2 Description of the scientific framework

The Ontology construction and reasoning support framework is responsible for:

- generating the appropriate data mapping
- populating the Knowledge Base with the information
- enriching the results using the reasoning module
- retrieving the results using the semantic queries

Furthermore, the localisation module is responsible for:

- retrieving location and organisation entities that appear in social media posts

3.8.3 Results and metrics

Since the semantic framework has not been finalized, some metrics have not yet been utilised. The completeness and conciseness, and the consistency are taken into account at the time that we develop the ontology. The consistency, the pitfalls and the completeness and conciseness are part of the measurement of the quality of the ontology, while the accuracy and precision are associated with the reasoning framework. The response time of the framework will also be taken into account (Table 10). Lastly, to measure the performance of the English/Finnish location recognition metadata models and the respective named entity recognition (NER) configurations, values for the precision, recall and F1-score measures were computed (Table 11,Table 12).

Table 10: Ontology-related metrics

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Completeness and Conciseness	Number	PUC1_GA2, PUC1_GA3, PUC1_GA7, PUC2_GB4, PUC3_GC2, PUC3_GC5, PUC3_GC7, PUC3_GC12	The ontology follows the competency questions
Consistency	Number		The ontology is consistent
Pitfalls	String (Critical, important, minor)		
Accuracy	Percentage		
Precision	Percentage		
Response time	Number of seconds		1.01100 seconds approximately

The ontology-related metrics results are satisfactory. More specifically, the ontology is consistent and follows the competency questions that mainly came from the PUC2 users. The response time of the data saving service in the Knowledge Base is extremely fast (1.01100 seconds approximately).

Table 11: Performance evaluation of the proposed NER system (EN) vs. the baseline system and a state-of-the-art approach

System (CoNLL2003 dataset)	Precision	Recall	F1-score
Our system	90.95	90.94	90.97
Next target value			92.00
Best-scoring shared task system: Florian et al., 2003	88.99	88.54	88.76
Baevski, A. et al. 2019	(not reported)	(not reported)	93.5

As is evident from Table 11, English language results are already satisfactory for the purposes of PUC 1. Although our scores are lower than the current state-of-the-art (Baevski, A. et al. 2019), they are not far off, while the model still outperforms the baseline method (Florian et al., 2003). Results are expected to increase by 1-2% when we further tweak current model parameters and test more modern models.

Table 12: Performance evaluation of the proposed NER system (IT) vs. the baseline system and a state-of-the-art approach

System (EVALITA2009 dataset)	Precision	Recall	F1-score
Our system	75.49	75.60	75.37
Next target value			80.00
Best-scoring shared task system: FBK_ZanoliPianta	84.07	80.02	82.00
Nguyen and Moschitti, 2012	85.99	82.73	84.33

The Italian dataset results are for the purposes of PUC 1. The model is still being worked on and fine-tuning is under process. As can be seen in Table 12 the model's experimental character is apparent since it is outperformed by the baseline method. Accuracy is expected to increase considerably when we update the dataset, the annotation format (to BILOU) and decide on final parameters.

3.8.4 Discussion

This module saves metadata from topic extraction in the Knowledge Base. The metadata are converted in semantic format and connected in terms of a semantic graph. Also, semantically similar terms are found in order to detect the similarity among different topics, which applies to all PUCs. For the localisation module in particular, we look for Italian location entities in tweets for PUC1, and for Finnish ones for PUC3, in order to ensure that the effort done is directly reflected to the users' particular needs.

3.9 Linking Open EO Data (KR09)

3.9.1 Purpose of Scientific Activity

The module is developed as a process that may be integrated in processing workflows to transform the metadata of Earth Observation products (such as Sentinel scenes) from their native format to Linked Data. The process behaviour is controlled using sets of mapping rules. In the context of EOPEN, mapping rules will be developed allowing to convert XML-encoded Sentinel metadata as well as the JSON records obtained from the Umbrella Sentinel Hub (KR10). Outputs are RDF graphs intended to be stored in the Knowledge Base (graph database, via KR08). The target Linked Data structure is compliant with the GC Discussion Paper OGC 16-074 *EO Metadata Discovery using Linked Data*, published in March 2016.

3.9.2 Description of technological framework

The module makes use of the FOSS tool GeoTriples to perform the conversions. The module is stateless (no need for a database or session management, for example) as it receives the data to convert through input parameters and provides back the generated RDF graphs through output parameters. The actual storage of the generated graphs in the KB must be performed downstream in the workflow.

When deployed, the process exposes an OGC WPS 1.0.0 compatible interface and allows users to issue execute requests to it. When integrated in a workflow, all the interactions with the WPS interface are managed by the platform. The Use Case developer only needs to graphically integrate the process in workflows to use it.

3.9.3 Functionalities and performance

The module receives as inputs the EO product metadata (by value or by reference) and a reference to the mapping rules to apply. It returns back the generated RDF graph (by value) or the reference of a file that contains that graph.

Table 13: Performance indicators and scientific metrics for Linking Open EO data

Indicator/Scientific metric	Value	Value achieved	Relevance to User Requirements	Comments
Conversion speed	Number		US_SK_03 US_SK_05 US_SK_06a	The number should be averaged.
Number of EO products (metadata) successfully converted	Percentage of test conversions			Test data to be prepared.
EO products metadata successfully stored in the KB	Percentage of test storage			Test storage scripts to be prepared.
EO products metadata searchable in the KB	Percentage of test searches			Test queries to be prepared.
Successful cross-concept queries	Percentage of test queries			Test queries to be prepared. Example of cross-concepts query involve EO products and tweets.

3.9.4 How does the technology enable the PUC implementation

Encoding the EO products metadata using the same specifications as the tweets and the events makes it possible to store, manage and retrieve linked entities in a seamless manner, using either stSPARQL or GeoSPARQL (provided by KR08). For example, it becomes possible to search in a single query for all the messages tweeted and all the satellite imagery acquired in the area and the time frame of a given event.

It also becomes possible to use an RDF browser to navigate in the tweets, events, EO products, and any other concepts that are encoded using the same classes.

4 TECHNOLOGICAL ACTIVITIES

Key Result preliminary evaluation and establishment of KPIs[5] to monitor

4.1 EO data acquisition (KR10)

4.1.1 Purpose of the technological activity

The developed Umbrella Sentinel Hub, under the EO data acquisition Key Result, provides a transparent and seamless access to Sentinel data to the user, enabling the Earth Observation (EO) based implementations of the PUCs. This EOPEN platform service offers a single point of access to multiple Sentinel hubs, allowing access to all Sentinel missions, with no geographic restrictions, high performance, and minimal delays.

4.1.2 Description of technological framework

The application developed connects to the APIs of multiple Sentinel hubs, searches and stores new metadata and chooses the most appropriate source from which a requested product will be downloaded from. The Umbrella application database frequently collects metadata of newly ingested Sentinel products, keeps up to date by deleting no longer available products and scores the different connected hubs according to their download speed.

4.1.3 Functionalities and performance

The API application creates a REST API via Django views module and allows users to make GET requests to it. In this request, the users are able to define their parameters (mission, product, model, date etc.) based on their needs. The Umbrella hub gives back a result set that contains metadata from the most efficient source to download from.

Table 14: Performance indicators and scientific metrics for EO data acquisition

Indicator/Scientific metric	Value	Value achieved	Relevance to User Requirements	Comments
Number of connected hubs	Number	4	US_V_01c US_V_05a3 US_SK_03 US_SK_06c US_SK_07a US_SK_12a US_F_01a2	By the end of the project more than 10 hubs will be connected including ONDA DIAS
Mean download speed compared to Open Access Hub	Percentage of increase			Ensured timeliness for time sensitive applications
Mean latency compared to Open	Percentage of decrease			Ensured timeliness for time sensitive applications

[5] The KPI tables in Section 3 include the field 'relevance to user requirements'. The user requirement IDs included can be found in the ANNEX of D2.2., which is a living a document and will be soon resubmitted.

Access Hub			US_F_03 US_F_04a US_F_04a3	
Access to all Sentinel missions	Boolean	Yes		Access to S-1, S-2, S-3 and S-5p allowing the generation of diverse datasets from soil moisture to bathymetry
Search data for a user defined area of interest	Boolean	Yes		
Search data on a specific time range	Boolean	Yes		
Number of broken links	Number			

4.1.4 How does the technology enable the PUC implementation

This Umbrella Sentinel API functions as the basic building block in all EO based PUC workflows, enabling EO data acquisition in a comprehensive manner. On top of this API, customized PUC-specific processors are built. For the case of PUC2 where time-series of Sentinel imagery is required to capture the phenological stages of the crops, the *time-series* processor has been implemented by the PUC developers. This is another general process utilizing the Umbrella application to create Sentinel-1 and Sentinel-2 time-series, which can be parameterized by the user (start month, end month, max number of images per month etc).

4.2 Social media crawlers (KR12)

4.2.1 Purpose of the technological activity

The Social Media Crawling module is responsible to gather information from Twitter that serves as a non-EO data source for EOPEN. Based on suggestions by end users, proper search filters are used to crawl valuable social media posts. In the current version, after being collected, we detect whether tweets are fake, we try to detect their location if mentioned, and we extract visual concepts from their images. In the future versions, we will check for nudity in images, while textual and visual information of tweets will be used in order to automatically estimate whether they are relevant or not to the examined use cases. At this point the crawling procedure and the complementary components only run on CERTH's side, but all the information is stored in a database that the platform has access to. An introduction to this module has been included in D3.1, but future deliverable D3.3 will be dedicated on it.

4.2.2 Description of technological framework

In order to gain access to the global stream of Twitter data, we use Twitter API, a free option to stream real-time tweets. The retrieval options can be keywords that must be included in the posts, open user accounts and locations in form of bounding boxes. Using these search criteria, the API constantly retrieves new tweets in a JSON format and stores them in a MongoDB database. Any information derived from the analysis of the crawled tweets (e.g., extracted locations and concepts, relevancy estimation, nudity detection, etc.) is appended to the JSON before storing.

4.2.3 Functionalities and performance

The evaluation of this module will focus on the visual and textual classification techniques regarding the relevance of collected tweets. The developed models will be language- and use case-specific, so they will be evaluated separately. The metrics to be calculated can be seen in the following table. It should be noted here that ground truth is strongly required and it can be obtained by human annotation, which is actually an already ongoing task.

Table 15: Performance indicators and scientific metrics for social media crawlers

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Precision	%	US_V_02a1 US_V_02a2 US_F_01b	
Recall	%		
F-score	%		
Accuracy	%		
Response time	millisecond (ms)		

4.2.4 How does the technology enable the PUC implementation

As mentioned above, the social media crawling procedure uses predefined search criteria, which are tailored to the needs of each PUC. For example, for PUC 1 we search for Italian keywords regarding flooding events inside the bounding box of north-east Italy. For PUC 2 we follow accounts relevant to food security and search for Korean words that refer to agriculture. Finally, for PUC 3 we search for Finnish terms about snow coverage. As a result, every collection of tweets is directly connected a use case.

4.3 Meteo data acquisition (KR13)

4.3.1 Purpose of the technological activity

The developed Weather Data Wrapper Module is a collection of functionalities designed for extracting, transforming, and storing meteorological and climatological data in the EOPEN platform.

4.3.2 Description of technological framework

The module provides a collection of functions and methods, written in Python. EOPEN developers can use the module to automate data ingestion to the EOPEN platform. The data is stored into a common database, from which the user can obtain the requested data via suitable database queries. In addition to PUCs, the module can also serve other EOPEN modules, such as Event Detection module.

4.3.3 Functionalities and performance

The module connects to multiple open APIs which provide meteorological and climatological data, extracts the data from the services, and transforms the data into JSON data format. During the project, the development is focused on the needs of the use cases, however, the module is designed to be reusable for the future needs of EOPEN users.

Table 16: Performance indicators and scientific metrics for meteo data acquisition

Indicator / Scientific metric	Value	Value achieved	Relevance to User Requirements	Comments
Number of connected services providing meteo data	Number	2	US_V_08 US_SK_02 US_SK_03 US_SK_05 US_F_01 US_F_07 US_F_08 US_F_09 US_F_10 US_F_11	We foresee connecting to up to 5 different services during the project.
Module is extendable to include additional APIs or services	Boolean	No(*)		Extendability is not directly related to the needs of the use cases. However, the module must allow for inclusion of new services as required by future EOPEN users. (*)Final design is not complete

4.3.4 How does the technology enable the PUC implementation

The module connects to open APIs, which provide the meteorological and climatological data as requested by the Use Cases. In order to limit the requests made to the services, this module is designed to be a lower-level functionality. Namely, the data ingestion and transformation are developed as separate processes from the Use Cases; the Use Case developers should query the EOPEN database in their processors. For the Use Cases, the module will be developed to connect to FMI Open Data service (PUC1, PUC3), Korean Meteorological Administration Open API (PUC2), Copernicus services (PUC1, PUC3), and Sodankylä National Satellite Data Centre datasets (PUC3).

4.4 Standardized Interfaces (KR14)

4.4.1 Purpose of the technological activity

The use of standard interfaces, including formats and protocols, is the key to bring interoperability among independently designed and developed systems and environments. To be fully interoperable, a system must adhere to standard interfaces on all its ends, including back-ends interfaces used to communicate with remote services and data providers, and its front-ends allowing client users and applications to interact with it in a seamless manner.

Standardized interfaces thus allow systems to not operate in an isolated manner, to benefit from data and services offered by other compatible systems, and on its turn to offer its data and services to client applications.

4.4.2 Description of technological framework

The EOPEN platform is an environment that allows integrating, configuring and executing any kind of services. An example of custom service is the harvesting and processing (geo-location, concept extraction, community detection) of tweets. Other services automatically harvest meteorological products encoded in netCDF, convert them to JSON, and store the results in MongoDB. The platform also integrates processes that allow publishing products (such as GeoTIFF files) as GIS layers through GeoServer. The adoption of standards must thus not only be considered as the platform level but also at the processes and applications level.

For the user point of view, it is important to be able to discover and access the available data, as well as discover and user the available services in a standard manner.

4.4.3 Functionalities and performance

Table 17: Performance indicators and scientific metrics for Standardized Interfaces

Indicator/Scientific metric	Value	Value achieved[6]	Relevance to User Requirements	Comments
Ability for a user to authenticate using a standard protocol	Boolean	Dev. Platform: True Dashboard: False	US_SK_14	Relevant protocols: OAuth2, OpenID Connect, CAS
Ability to discover	Boolean	False	US_F_05a	Relevant interfaces:

[6] Most indicators are "Mostly True" or "Mostly False" at this stage. Indicating "True" here does not mean we do not need to still improve the solution and indicating “False” does not mean it will not be developed later. It merely records the current status.

the available collections and products through a standard interface				OGC CSW, OpenSearch, OpenAPI, SPARQL
Ability to access the product data files through standard interfaces	Boolean	True	US_F_05a	Relevant interfaces: http, ftp
Ability to access GIS layers through standard interfaces	Boolean	True	US_SK_06b	Relevant interfaces: OGC WMS/WMTS, WMS-T, WFS, WCS
Ability to remotely execute an application through a standard interface	Boolean	False	US_F_05b	Relevant interface: OGC WPS

4.4.4 How does the technology enable the PUC implementation

PUC stakeholders already use tools to perform their tasks. EOPEN bring enhancements in providing new products and services. The use of standard formats and protocols allows PUC users to interconnect the tools they are already using with services run in the EOPEN platform and they will be able to fetch the data available in EOPEN to use them in their tools. Concrete examples include the possibility to visualise GIS layers (OGC WMS, WFS) served by EOPEN in any standard-compliant GIS client such as QGIS and ArcGIS. Opening water body masks encoded as raster data (GeoTIFF) or vector data (Shapefile, GeoJSON) is also possible in user-chosen software. Moreover, these standards allow visualising and fusing data served by EOPEN with user owned data and any other data served by third-parties.

4.5 High Performance Computing (KR15)

4.5.1 Purpose of the technological activity

The goal of this module is to integrate HPC and HPDA infrastructures into the EOPEN platform for deployment of embarrassingly-parallel tasks and data analytics applications. Specifically, Cloudify orchestrator will be implemented to enable access to both High Performance Computing (HPC) and High Performance Data Analytics (HPDA) infrastructures in a seamless and user-friendly manner.

4.5.2 Description of technological framework

To enable fully utilization of the computing resource of HPC and HPDA through Cloudify orchestrator, further Cloudify plugin needs to be developed. The Cloudify plugin is composed of two parts: Cloudify HPC plugin and Cloudify HPDA plugin.

Cloudfiy HPC plugin, which is designed to organize the interaction between portal and HPC systems, is consisted of three main modules: workflows, tasks, cli_client and workload_manager. It was already developed in the former GoeGSS project and has been implemented in EOPEN. A sample BASEMENT simulation for hydro- and morphodynamic modelling has successfully run on the HPC at HLRS.

Cloudfiy HPC plugin will be implemented as an extension of the Cloudfiy HPC plugin, which could support Mesos as the resource manager. Thus, it required significant changes in submodules workload_manager and cli_client. These changes will be made in the next steps.

To effectively process and analyze Big Data on the HPDA system, a stack of software frameworks is established. Apache Hadoop is enabled for processing Big Data at large-scale. Apache Spark, which allows in-memory processing and a large set of libraries for Big Data processing, including common clustering and classification tasks is also enabled. Cray Graph Engine is supported to provide the ability to perform graphic analysis. Besides, Tensorflow is supported to enable deep learning algorithms performed on the image analysis. Further information is delivered in D6.1.

4.5.3 **Functionalities and performance**

The performance will be measured via a monitoring framework such as ATOM, which is going to be deployed, and that performance measurements will be provided in deliverable D6.3.

Table 18: Performance indicators and scientific metrics for concept extraction in social media

Indicator/Scientific metric	Value	Value Achieved	Relevance to User Requirements	Comments
Access to HPC	Boolean	True	US_V_01 US_V_02 US_V_08d US_F_08c1 US_F_10a1 US_F_10b1 US_F_11a1	Access to HPC enabled via ssh and Cloudify
Access to HPDA	Boolean	False		Access via ssh is enabled and HPDA plugin is in progress
At least three application kernels parallelized	Number	2		Applications like community detection and text analysis are implemented in parallel. Clustering of EO images to be done.
Improved application performance by at least 15%	Percentage	N/A		Prerequisite for benchmarking is the deployment of ATOM which is to be done.
At least 5 workflows	Number	3		Components from

ported to HPC/HPDA				WP4 and WP7 executed on HPC/HPDA
Workload manager as a component	Boolean	True		Employed Cloudify as orchestrator for HPC/HPDA; Integrated with EOPEN platform
Adopted state of the art data analytics frameworks	Boolean	True		Apache Spark, Cray Graph Engine and TensorFlow, etc. are available for processing.

4.5.4 How do we measure

For monitoring HPC jobs at HLRS, CrayPAT a preinstalled tool is adopted. With CrayPAT, the performance data generated during the execution of the program on Hazel-Hen can be captured and analysed. The information collected and analysis produced by using CrayPAT can help to identify the speed of the program as well as the method for accelerating it. The information which are used for the analysis includes: call graph profile, communication statistics, time-line view for Communication and IO, activity view, pair-wise communication statistics and text reports. Based on the analysis, the user can easily monitor the load imbalance, excessive communication, network contention, excessive serialization and I/O Problems. For more information about CrayPAT, please refer to: https://kb.hlrs.de/platforms/index.php/CRAY_XC40_Tools#Perftools_:_Performance_Analysis_Tool_Kit

Besides the preinstalled CrayPAT for HPC task monitoring, ATOM will be implemented in next to monitoring the HPDA framework at HLRS. The objective of the ATOM Monitoring Framework is to enable the application optimization based on the understanding of both software non-functional properties and hardware quality attributes with regards to performance and energy consumption.

The monitoring workflow is divided into two parts as shown in next Figure. The devices keep being monitored by the **MF-Client** (Infrastructure-level), and the instrumented applications keep monitored using the **MF Library** (Application-level). The metrics collected are sent to the **MF-Server**, which tasks are to store the metrics and to provide an interface for query and analyse them.

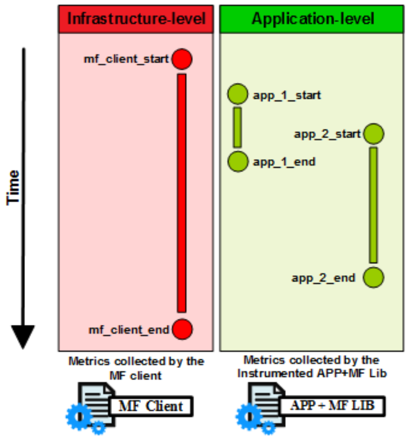


Figure 1 Infrastructure-level and Application-level monitoring.

The metrics that can be acquired cover a wide range of functions that target different aspects of the hardware, such as memory, IO, processor, GPU, and network utilization as well as energy consumption. Next figure shows also a classification of metrics:

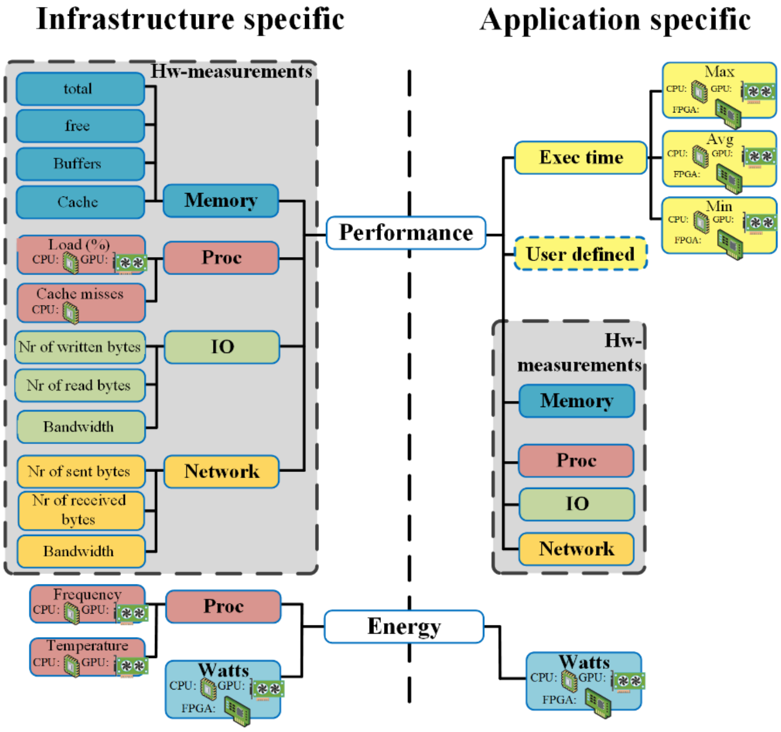


Figure 1: PHANTOM Monitoring Client metrics' taxonomy

For more detailed information, please refer to: https://github.com/PHANTOM-Platform.

4.5.5 How does the technology enable the PUC implementation

For PUC1, it is required to provide faster and more effective emergency responses to extreme weather by increasing the speed of risk analysis and by improving in-field coordination of emergency response. This requirement can only be satisfied by exploiting large-scale distributed computing power provided by HPC and HPDA. EOPEN builds on state-of-the-art methods in HPC and HPDA, and as the basic building block in all EO based PUC workflows, the extreme performance of the computing systems will be exploited by the EOPEN analytics platform. The High Performance Computing and High Performance Data Analytics (HPDA) can provide two different strands for PUC: batch job processing for static workloads as well as stream processing for live data analytics. One great example is PUC 2 that deals with long time-series of Sentinel imagery that extends over multiple years. Such applications, at the national scale, constitute big data problems and can be solely tackled by high performance computing systems.

4.6 Fast and Secure Data Exchange (KR16)

4.6.1 Purpose of the module

The purpose of this module is to enable a fast and secure data exchange so that the full bandwidth of the underneath high-speed data connection could be utilized. Moreover, a secure transfer and storage will be ensured through the maintenance of standards in encryption. Besides, data locality is of utmost importance in order to achieve maximum performance, thus input data are required to be transferred in a secure and fast manner to these systems before processing can start.

4.6.2 Description of technological framework

A secure and fast data transfer from and to HLRS infrastructure is planned to be realized via Rucio. Rucio is a scientific data management framework that is developed by CERN with the intention to transfer very large data. EOPEN is following here best practices for authentication and authorization, and the actual transfer will benefit from parallel data streams such as provided by the GridFTP protocol to ensure a fast and secure data transfer between sites.

Besides, as the images for the rice classification Sentinel time-series is refreshed at quite low frequency, therefore as an addition to transferring data via Rucio, the Umbrella API which was implemented by NOA will provide the metadata to HLRS and then the products will be downloaded with a simple script directly at HLRS. Thus the workload for transferring data from NOA to HLRS is eliminated.

4.6.3 Functionalities and performance

Table 19: Performance indicators and scientific metrics for fast and secure data exchange

Indicator/Scientific metric	Value	Relevance to User Requirements	Comments
Service Coverage	%	US_V_02a US_V_02b	Number of sites which are offering data services (goal is to have a 100% coverage)

Service Availability	%	US_V_02c US_V_02d1 US_V_05a US_V_05b US_V_09a US_V_11 US_SK_07	Uptime is taken into consideration for monitoring the service availability. If the service is required to be available all the time, then the goal is to have 100%.

4.6.4 How does the technology enable the PUC implementation

Secure and fast data exchange is a fundamental requirement for all PUCs. Data transfer with Rucio by GridFTP could provide authentication and encryption to file transfers, with user-specified levels of confidentiality and data integrity. Also for cross-server transfers it could enable all PUCs achieve a high data rate on high bandwidth connections when dealing with the data. Also, the download script which can fetch the metadata and download links provided by the umbrella API will be implemented at HLRS locally. Therefore, the workload of data transfer would be eliminated, and the efficiency of processing is improved.

4.7 The EOPEN Platform (KR17)

4.7.1 Purpose of the technological activity

The EOPEN Platform provides core components that allow orchestrating the execution of processes organised in worflows. The platform includes built-in processes that allow performing generic tasks such as sending email notifications and registering a product in a catalogue. Common libraries and toolboxes are also available in the core platform that allow developing custom algorithms and import them as processes.

The EOPEN Platform is accompanied by a generic Web-based user interface allowing to interact with the platform, request for workflow executions, and for discovering and visualising the available data.

The EOPEN Platform is thus meant to be used by developers for implementing new modules and applications at setup-time and used as a back-end for collecting, processing and serving data toend-users during normal operations.

4.7.2 Description of technological framework

The EOPEN Platform is based on the ASB Framework initially developed for ESA. In particular, the ASB core components included in the EOPEN Platform are used by the module and Use Case developers to import their algorithms (as processes and services), create, test and execute processing workflows, and publish these as applications (processors). In the course of the EOPEN project, a number of modules (some are implementing the KRs) must be integrated allowing the preparation of the Pilot Use Cases.

4.7.3 Functionalities and performance

The core functionalities of the EOPEN Platform reside in the ability to import algorithms, configure workflows and execute processors. These features are available through a Web interface. The performance may be related to the amount of modules integrated and available for implementing each of the Use Cases.

Table 20: Performance indicators and scientific metrics for the EOPEN platform

Indicator/Scientific metric	Value	Value achieved	Relevance to User Requirements	Comments
Amount of integrated services	Number	9*		
Amount of operational processes	Number	11*		
Amount of operational processors	Number	>5		
Amount of scheduled processors	Number	6		
Amount of available database engines	Number	3	US_F_01	PostgreSQL (RDBMS), MongoDB (Document), GraphDB (Triplestore)
Amount of product types managed in the platform	Number	>7	US_F_01	Tweets, Sentinel-1/2/3 images, water bodies, SMOS Freeze-Thaw, GlobSnow SWE L3A, ...
Amount of processes executable in the HPC	Number	0	US_F_05b	Proofs-of-concept have been implemented but no process currently integrated in EOPEN workflows.
Amount of DIAS platforms available for deploying EOPEN processes	Number	(soon: 2)	US_F_01d	The mid-term objective is to support them all.
Amount of product types accessible in the Dashboard	Number		US_SK_14	
Amount of product types accessible through APIs	Number		US_SK_06c	This includes custom and standard APIs, including database API, OGC Web Services, etc.

* As documented in D2.3 Application Developer User Guide v1.0.

4.7.4 How does the technology enable the PUC implementation

The EOPEN Platform, with its core components and its integrated modules allows PUC-specific processing chains to be implemented and executed (immediately or differed) and the collected and generated data to be displayed to the end-users. In this context, the usefulness of the Platform not only resides in the built-in features but also on the various modules that will be integrated (some being directly related to the KRs).

5 PUC IMPLEMENTATION AND PRELIMINARY EVALUATION REPORTING

5.1 Joint Decision & Information Governance Architecture Framework

5.1.1 Overview

The Joint Decision & Information Governance Architecture Framework (JDIG) is an architecture framework designed for EOPEN to support service improvement in the PUC's. Based upon TOGAF 9.2, it builds Current (COM) & Target (TOM) Operating Decision-making models, to identify how services are currently implemented in the PUC, and how they could be implemented using EOPEN; through a combination of People, Process, Information, Governance & Technology components.

5.1.2 Baseline Evaluation

As part of the JDIG each PUC underwent a baseline Holistic Capability Assessment to identify the maturity of its capabilities to deliver their desired services, as seen in Figure 2, Figure 3 and Figure 4 below. Rationale behind the evaluation and the process to define these are described in detail in *D2.3: Current Operating Decision-Making Models.* The colour coding is as follows:

- Green – Optimal Functionality
- Amber – Satisfactory but can be improved
- Red – Improvement needed
- Black – Not applicable/ Not enough information provided

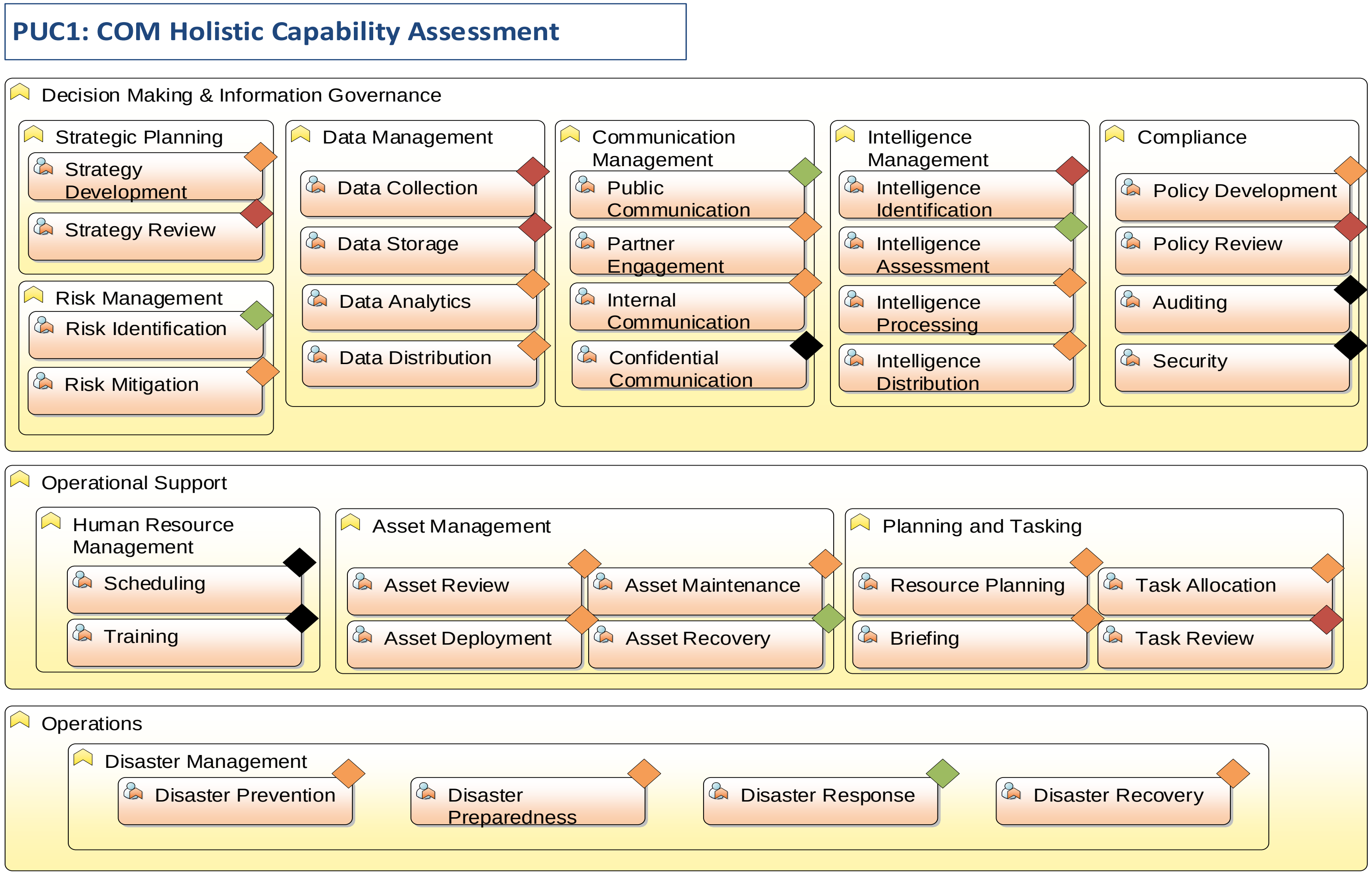


Figure 2: PUC1 - COM Holistic Capability Assessment

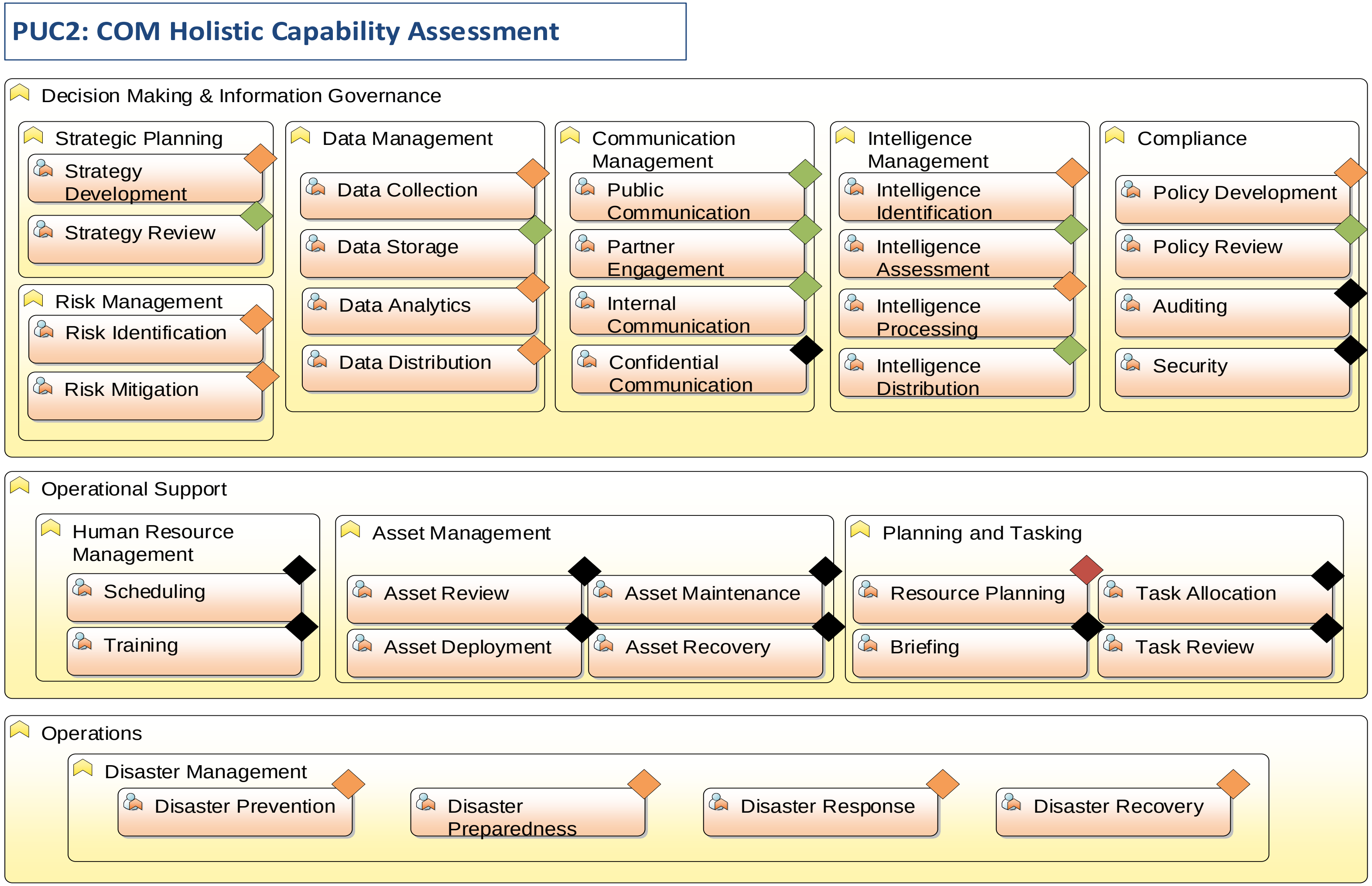


Figure 3: PUC2 - COM Holistic Capability Assessment

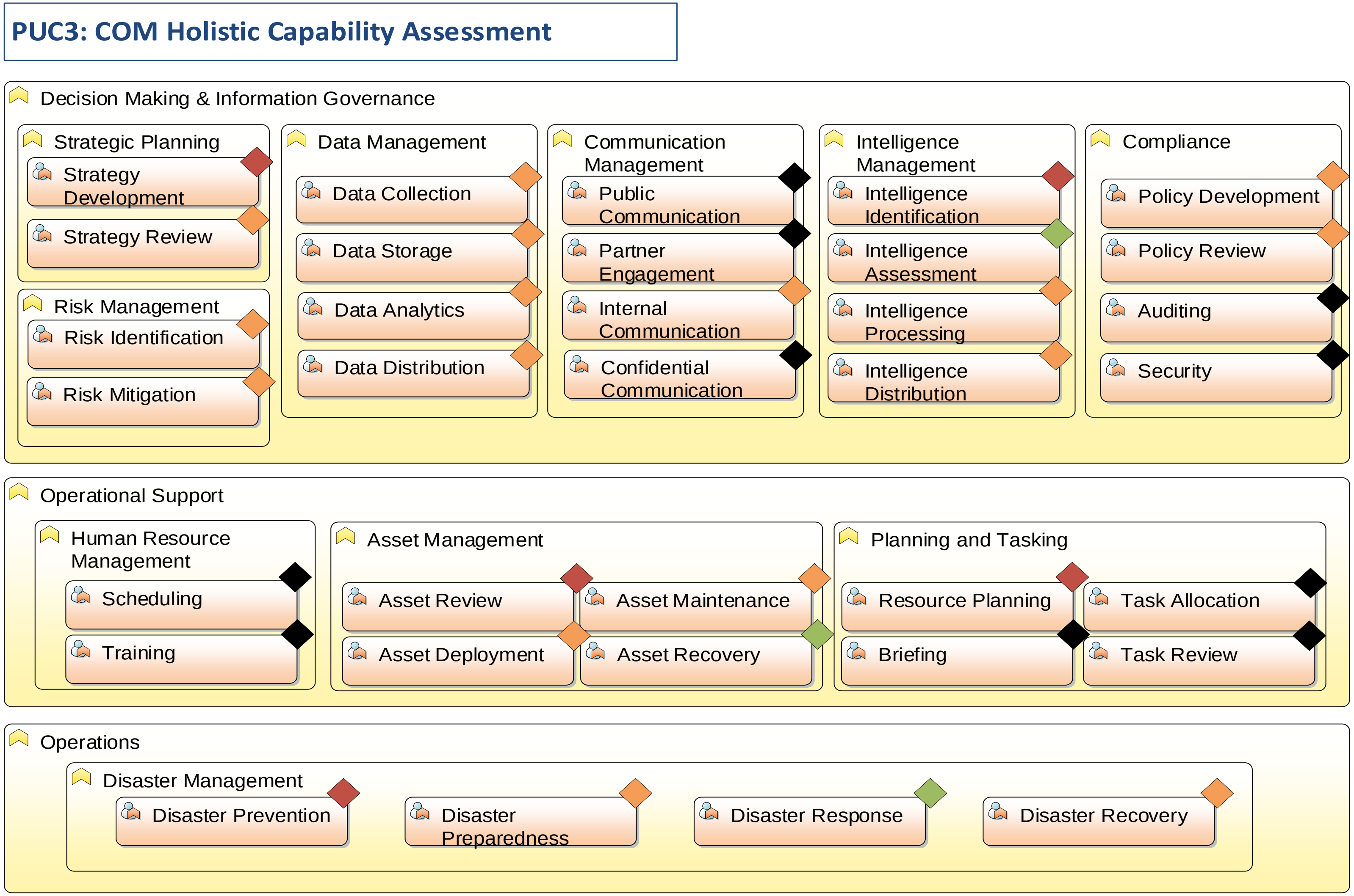


Figure 4: PUC3 - COM Holistic Capability Assessment

5.1.3 Future Evaluation

As part of the JDIG TOM process, PUC's will undergo a final Holistic Capability Assessment, to verify improvements provided by the EOPEN platform have enhanced their capabilities. This capability assessment will be based upon the outputs of the individual PUC evaluations described in the following chapters.

5.1.4 Desired Improvements

EOPEN aims to provide direct and indirect improvements to the below capabilities. Direct meaning EOPEN will provide this capability itself; Indirect meaning EOPEN does not provide this capability, but it is improved as a result of a direct improvement.

EOPEN aims to provide direct improvements to the following capabilities across all PUC's:

- Data Management (Data Collection, Data Storage, Data Analytics & Data Distribution)
- Risk Management (Risk Identification)
- Communication Management (Public Communication, Partner Engagement, Internal Communication)
- Intelligence Management (Intelligence Identification, Intelligence Assessment, Intelligence Processing, Intelligence Distribution)
- Planning & Tasking (Task Allocation, Task Review)

EOPEN aims to have indirect improvements to the following capabilities across all PUC's:

- Strategic Planning (Strategy Development, Strategy Review)
- Risk Management (Risk Mitigation)
- Compliance (Policy Development, Policy Review)
- Asset Management (Asset Review, Asset Maintenance, Asset Deployment, Asset Recovery)
- Planning & Tasking (Resource Planning, Briefing)
- Disaster Management (Disaster Prevention, Disaster Preparedness, Disaster Response, Disaster Recovery)

5.2 PUC 1 – Flood risk assessment and prevention

5.2.1 Use Case story

The test scenario of PUC1 was built based on the flood event of 2010 in Vicenza; the storyboard provided in Deliverable 2.1 underlines the most important figures involved in the emergency management and their role in the "Command Chain". EOPEN provides a unique platform for all actors involved in the emergency management; the use of satellite images and ground truth information in a scalable environment is the added value that the PUC1 wants to test.

During emergency all forces work together with a common target that is the emergency support and safety for citizens; the main problem is the fragmentation of information due to the number of actors involved. During emergency it is very difficult to share information in traditional ways (papers, emails, fax and so on) and it requires a lot of time, so the decision maker can't have a precise and timely view of the emergency. Another issue is that at the end of emergency the lack of a unique database results in losing acquired information from the event.

The lack of powerful computer infrastructure is mainly a limit for the use of satellite images that require Earth Observation knowledge and an appropriate ICT infrastructure. What works is mostly the command chain that is known by all, during emergency every actor works separately with their team, information and decisions are shared at the manager/director level that are physically inside each organization (from COC to DICOMAC).

Table 21: PUC 1 main stakeholders

Stakeholder	Description
ARPAV	This administration is the environmental agency of Veneto Region; inside this office there is also the Copernicus contact person for Veneto Region
Regione Veneto Difesa Suolo	This Administration is the office of Veneto Region Administration responsible of land use, water, environmental planning, and waste management. This office is also responsible for cartography.
Regione Veneto Protezione Civile	This Administration is the office of Veneto Region Administration responsible of Civil Protection Volunteer (at regional scale, so about formation, guidelines etc.), for Bulletins (Alert bulletins)
Genio Civile di Vicenza	This Administration is the operative office of Veneto Region Administration in water management (river maintenance, river project, dikes etc.)
Corpo Nazionale dei Vigili del Fuoco di Vicenza	The firefighters of Vicenza.
Provincia di Vicenza protezione Civile	This is the provincial office of Civil Protection, similar to Veneto Region Civil Protection office but at province level.
Comune di Vicenza	Municipality of Vicenza.
Consorzio APV	This Administration is a Land reclamation authority responsible for

	''Alta Pianura Veneta'' basin.
Consorzio Brenta	This Administration is a Land reclamation authority responsible for ''Brenta'' basin.
AAWA (Autorità di Bacino dei fiumi Isonzo Livenza Piave Brenta-Bacchiglione)	This Administration is the higher water authority in Veneto, Trentino-Alto-Adige and Friuli region; it is also responsible for some international basins like Timavo Basin (Slovenia). It is an office directly dependent from the Italian Environmental Ministry.

At the moment the most important stakeholder for PUC 1 is AAWA that is also part of the EOPEN Consortium; AAWA is following step by step the development of EOPEN platform. Other important stakeholders will be involved in EOPEN development during the "Infoday" in September 2019 organized by AAWA to present the first prototype of the platform. During this event we will provide to stakeholders with an overview of the platform linked explicitly with the user requirements acquired in April 2018. At the end of the "Infoday" we will provide a second iteration of questionnaires to get a feedback from participants. Questionnaires were provided to the stakeholders that participated in the Vicenza Meeting (April 2018, User Requirement meeting, listed in the table above) and to other stakeholders (general public). The feedback from the first category of stakeholders (the most important) will be the backbone of the evaluation; other stakeholders could also provide ideas and feedback that could be an added value to the EOPEN platform.

5.2.2 Implementation

Flood detection product

The flood detection product utilized in PUC1 is part of the change detection module and more information can be found in D4.1. Additionally the relevant performance indicators for the flood detection product are mentioned under the Change Detection subsection (3.1), earlier in this document. The product will also integrate during the validation phase the tweet analysis products developed by CERTH. Tweets will be used as proxies of ground truth information to ensure that flooded areas are correctly identified. Tweets detected as relevant by CERTH algorithm are geolocalised and a map in the interface will show different layers (tweets, flood maps etc.) to give to decision makers all information needed in a comprehensive manner.

Hydraulic model (prototype)

The hydraulic model prototype will be addressed by the end of July (M21); and it will be integrated on the platform. Users will load model parameters (geometry, boundary conditions etc.) and run the model inside the platform. All results and that is depth and velocity maps will be stored and visualized on the platform. The model will be updated during the integration phase (Hydrologic-Hydraulic, M25) with the hydrological module to produce an Early Warning system that can work as predictor for flooded areas, so as to request in advance pertinent satellite products.

Questionnaires

A second iteration of questionnaires will be produced after the "Infoday" to evaluate the status of the platform and to acquire a feedback from stakeholders respect to the original requests.

Flood detection product (final)

This the final product for flood detection that follows the test and the development of the baseline product. This product will be able to detect flooded areas from optical and SAR imagery but will also include advanced filtering and noise cancelling post-processing techniques. The algorithm will also account for tweets and the results of hydrological-hydraulic model that works as a predictor.

5.2.3 Timeline for PUC 1 implementation

Milestone	Completion Month /Expected completion month	Status (done, in progress, to be done)	Comments
EOPEN User Requirement Meeting	M6,	Done	Meeting with PUC 1 stakeholders in Vicenza about EOPEN User Requirement
Questionnaires	M8	Done	
Requirements gathering	M8	Done	Both user and data requirements are defined. These may be refined during the project based on the feedback of the end users.
Development of baseline flood detection method	M16	Done	Developed by CERTH
Integration of baseline flood detection module with the platform	M19	Done	Developed by CERTH

Hydraulic model (1st prototype)	M21	In progress	
Questionnaires	M24	To be done	2nd iteration during infoday in September
Flood detection product (baseline)	M23	In progress	Collaboration with CERTH
Hydrologic-Hydraulic model	M26	In progress	
Development of refined/sophisticated flood detection method	M26	To be done	
System workflow integration module	M28	To be done	The tool to query AAWA products in EOPEN platform
Integration of refined/sophisticated flood detection module with the platform	M28	To be done	
Flood detection product (final)	M28	To be done	Collaboration with CERTH

5.2.4 Monitoring of PUC 1 implementation against the User Requirements

Use case G-id	Users of the EOPEN platform:	Status
PUC1_GA1	Must be provided with capabilities for data dissemination and integration of EO data, weather information and relevant social media text and images.	Development (1)
PUC1_GA2	Must be allowed geo-localisation of social data and real time control.	Done (*)
PUC1_GA3	Should be provided with an intuitive online platform with the possibility to visualise EO data (e.g. from webcam or mobile phone) and possibility to send text message (e.g. SMS or tweet) and analyse its semantics for meaningful	Development (2)

	automatic decision-making.	
PUC1_GA4	Must be enabled to merge different administrative database and formats in a unique platform with all data shared.	Done (*)
PUC1_GA5	Should be enabled to implement EWS and add maps in the platform.	Development (3)
PUC1_GA6	Should be provided with an intuitive and robust interface.	Development
PUC1_GA7	Must be provided with an interactive archive of each event; all data from social network communities and from satellites should be stored in a specific database to provide a history of each event.	Not dev.
PUC1_DA1	DEM/DSM 1m (e.g. Airbus Pleiades) from Copernicus	Delivered (1)
PUC1_DA2	Snow maps with a resolution < 20m	Delivered (2)
PUC1_DA3	Soil moisture maps with resolution < 10m	Delivered (3)
PUC1_DA4	Flood maps	Development (4)
PUC1_DA5	Damage maps	Development (5)
PUC1_DA6	Water presence maps	Development (6)
PUC1_DA7	Bathymetry of coast, lakes, rivers	not prov.
PUC1_DA8	Orthophoto with resolution of 50 cm (e.g. WorldView4)	Delivered (7)
PUC1_DA9	Vegetation presence	Delivered (8)
PUC1_DA10	Land cover	Delivered (9)
PUC1_DA11	LAI and other vegetation indexes	Delivered (8)
PUC1_DA12	Other maps (thermal or multispectral data ready to be processed) with high resolution	Not. prov

PUC1_DA13	Weather forecast	Delivered (10)

User requirements (GA_1 to GA_7)

1- This capability will be better developed in the second prototype, at now there are different interface for the visualization of maps and layers and for the processing
2- This capability will be better developed in the second prototype, at now there are different interface for the visualization of maps and layers and for the processing, at now we can see tweets georeferenced
3- The hydrological-hydraulic model is under development, we will provide a first version before the Infoday limited to Vicenza City; the hydraulic solver is going to be updated with the use of CPU's and GPU's The hydraulic model is an important instrument to predict flood; now it is running on a traditional Workstation with high computation cost and time; the new version of the solver, the 3.0, can work using CPU and GPU with a reduction of the computational time as show in the tables below for the Circular Dam Breach benchmark.

* Done with the baseline product, which will be updated during the project. All work is performed by CERTH with the help of AAWA as a co-designer.

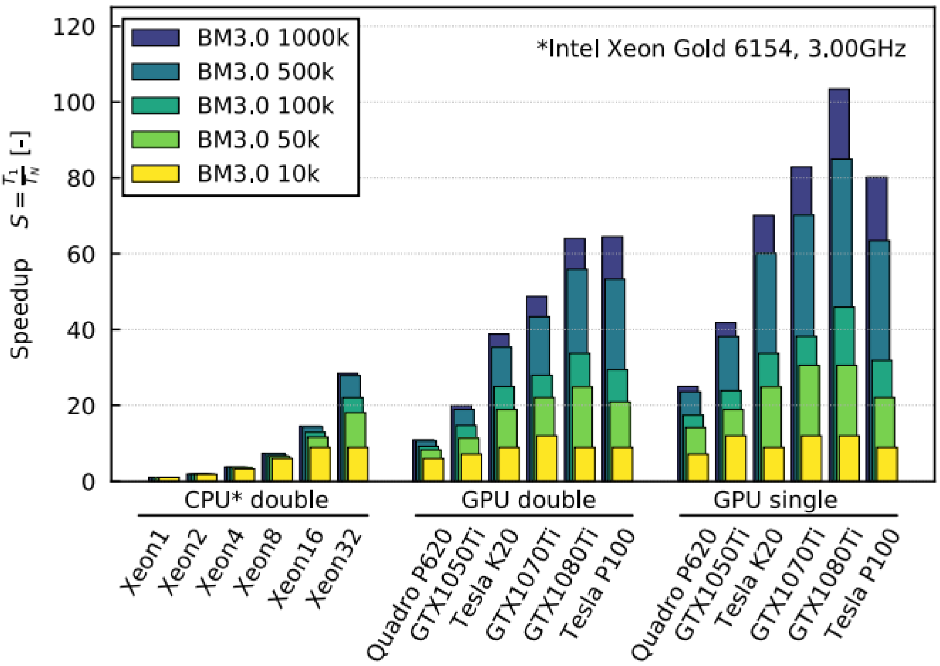


Figure 5 : Speed Up for different mesh size, 10K to 1000K elements, CPU vs GPU

The figure above describes the computational time required to perform a hydraulic simulation with Basement 3 (BM3) with different resolution of the computational mesh and different processors type. The increasing of the mesh resolution from 10k to 1000k increases the computational times, also the types of processor influenced this variable; however regardless of the mesh size and the processor type, running a parallel CPU simulation, option allowed only in the newest version of the solver, is much faster of the traditional single solver, used in the previous versions.

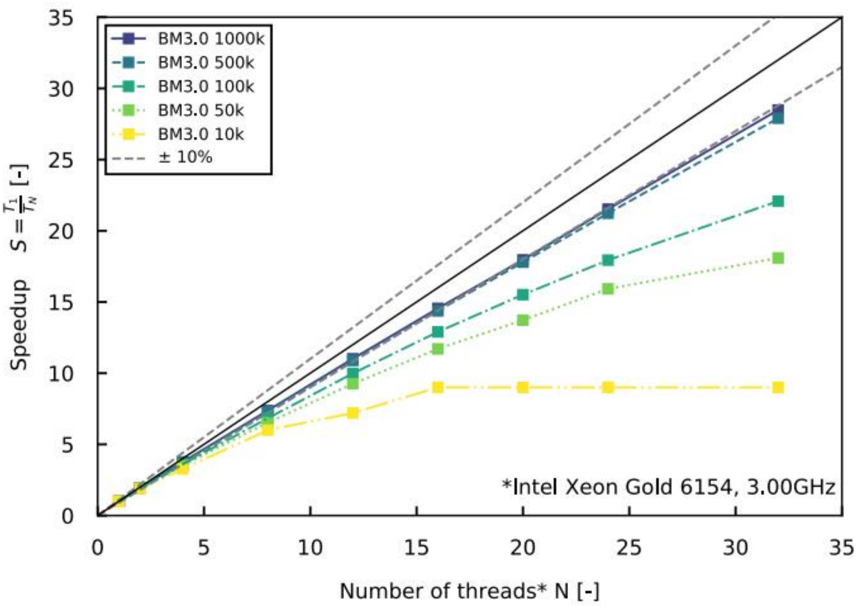


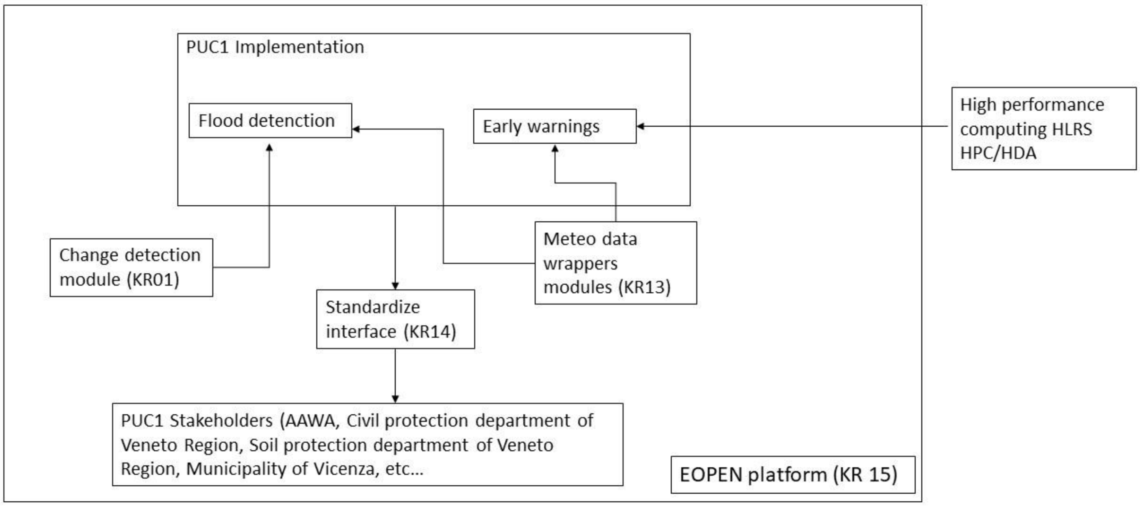
Figure 6: Speed Up for different mesh size, 10K to 1000K elements for a parallel computation with increasing numbers of cores of a fixed type (Intel Xeon 6154 3.00GHz).

The figure above describes the computational time required to perform a hydraulic simulation with Basement 3 (BM3) with different resolution of the computational mesh in case of a parallel CPU simulation. Raising the number of parallel cores used for the simulation increases the speed up factor; in particular for high resolution mesh (1000k) the speedup factor is almost linear increasing the number of cores; while for low resolution mesh there is a maximum numbers of cores for gaining significant time reduction. It is clear from this figure that the updating of the hydraulic solver from a serial to a parallel computing reduces drastically the calculation and elaboration time.

Datasets

1- The DEM DSM will be derived from AIRBUS TERRASAR-X/TAN-DEM HR Spotlight products
2- Snow maps at different resolution (snow cover < 20m, Snow Water Equivalent > 20m)
3- Resolution > 10 m
4- Maps under development (hydraulic model under update)
5- Access to EMS acquired in may 2019
6- Access to EMS acquired in may 2019
7- HR imagery from WorldView satellite (quota 2018)
8- Provided many different index (LAI, NDVI etc.)
9- Corine 2018
10- Weather forecast from FMI (HIRLAM)

5.2.5 PUC1 implementation workflow



5.2.6 Potential Test Scenarios

Scenario 1

The first scenario is about the flood prediction and the early warning products, involving the flood detection product and the hydraulic model; in particular, EOPEN platform identifies areas at risk of imminent flooding based upon satellite and in-situ data and provides automatic early warnings of the likelihood of a flood event based on the result of the hydraulic model.

In this scenario, it will be evaluated i) the accuracy of the detection of the imminent flooded areas from the satellite images, ii) if these results can be merged or layered to the administrative database and formats integrated in the EOPEN platform, iii) if the data can be integrated by information from social media.

In regard to the early warning product, we will evaluate the time required for the platform to provide new or update forecasts, as compared to with the current system used, focusing also on the speed up factor of the hydraulic model due to the parallel GPU's and CPU's computation, in comparison to the current serial computation.

Moreover, it will be evaluated if the prediction and early warning products are easily understandable and if they provide a clear outline of the situation, as these aspects could be translated, during a real flood, in time reduction to make a decision and subsequent actions.

Scenario 2

The second scenario deals with the automatic and real-time identification of flooded areas through the flood detection product.

We will evaluate the accuracy of the detection algorithm; if these results can be merged or layered to the administrative database and formats integrated in the EOPEN platform; if the information is presented in a clear and understandable way that could potential lead to a time reduction to make a decision and subsequent actions during a real flood.

Scenario 3

This scenario will exploit the EOPEN features of providing flood damage based upon the scale of flooded land and user input data.

We will evaluate the precision of the estimated damage, and if the data can be merged or layered to the administrative database and formats integrated in the EOPEN platform.

Finally, we will also assess the time and workload reduction to prepare damage reports.

Scenario 4

This scenario deals with the real time monitoring of hydraulic (i.e. water level) and metrological data (i.e. rainfall), in addition to having access to a database of historical measurements.

We will evaluate the User Experience, in particular regarding the easiness to access to the historical data; the time required for the System to provide the 'real time' measurement and to update it once new data arrives; if the interface is clear and robust; if the data can be merged or layered to satellite data and the administrative database and formats integrated in the EOPEN platform; if the data can be integrated social media derived information.

Scenario 5

This scenario involves the EOPEN Change Detection module applied for road passability assessment in order to provide users with details on current road blockages.

We will evaluate the accuracy of the identification of the road blockage, the time required to provide the 'real time' data and to update it when the situation changes, if the data can be integrated by information from social media; if it is robust; if the data can be merged or layered to satellite data and the administrative database and formats integrated in the EOPEN platform.

Scenario 6

This scenario deals with the platform capability that provides the users with predicted and historical weather forecasts; in addition, EOPEN provides warnings and notifications of potential extreme weather events.

We will evaluate the User Experience, in particular: the easiness to acces the historical data; if the information is clear, easily understandable and if it provides a clear outline of the situation, as these aspects could be translated, during a real flood, in time reduction to make a decision and take subsequent actions.

Finally we will also take into account the time required to provide the 'real time' information and to update it once a new forecast is available.

5.3 PUC 2 – EO datasets to monitor Food Security in South Korea

5.3.1 Use Case story

PUC 2 deals with the food security monitoring in South Korea, focusing on the production of rice. Rice is systematically overproduced in South Korea; resulting in large storage costs, while at the same underproduction of other major grains, making the country largely dependent on imports. Based on the collected user requirements (D2.2) and their subsequent translation into EO based products and services, it was decided to implement a food security monitoring system founded on the accurate mapping of rice and the succeeding estimation of rice yield for the year of inspection (currently 2018). The table below lists the main stakeholders of PUC 2 products and services.

Table 22: PUC 2 main stakeholders

Stakeholder	Description
Korea Rural Economic Institute	A national agricultural policy research institute focused on the development of agriculture, rural areas and the food industry. The institute is responsible for agricultural monitoring, Free Trade Agreements, World agriculture information, agricultural policies, Overseas crop market information, returning to farm support, international cooperation such as the Korean Agricultural Policy Experiences for Food Security(KAPEX).
National Institute of Agricultural Sciences	A research institute under the Rural Development Administration, which is in charge of testing and research to develop agriculture into a sustainable and competitive biological industry. The role of NAAS is generation of new income by exploration of the values of agricultural resources and maintaining a healthy agricultural ecosystem, crop protection and technical development on practical use of biological resources, development of production techniques for safe agricultural products, development of commercialization techniques for globalization of agricultural and Korean foods, and so on.
Korea Rural Community Corporation	A national corporation which focus on rural community development such as stable food production, development and management of agricultural infrastructures. KRC has been contributing to the stable food production for about 50 million people of the nation through the development and management of agricultural infrastructures such as reservoirs and pumping stations, and have also been improving the living standard of rural and fishing villages. Major project of the corporation includes food, water, climate, safety, research and training.

CJ CheilJedang Corporation	The largest general food manufacturer in South Korea, carrying out bio business and manufacture of food, medicine, and feedstuffs and other ingredients such as sugar, flour and oil. Keeping up with market trends CJ cheiljedang analyses social and environmental impacts. They are also dedicated to reinforcing innovative ideas and technological capacities in the fields of food, food ingredient, bio, and feed and livestock. Moreover, they identify the global trend of technologies in the food and bio sectors, explore new promising businesses.
Rural Development Administration (RDA)	A central government organization responsible for extensive agricultural research and services in Korea. The organization has 4 different Research and Development institutes: National Institute of Agricultural Science, National Institute of Crop Science, National Institute of Horticultural and Herbal Science, National Institute of Animal Science. The administration is focused on agendas such as basic agricultural science and technology and development of steady supply of food and state-of-art technology. .
Korea university (prof. Yang)	Seung-Ryong Yang, a professor in department of Food and Resource Economics, Korea university, combats growing problem of food security. To analyse the problem, he releases the special report about world food security situation monthly report using national food security index. The report subscribed to many people who have the interest in food security provides a lot of information and data related to food security. Moreover, it is used to propose policy at national or religion level.

5.3.2 Timeline of PUC 2 implementation

Milestone	Completion month	Status	Comments
Questionnaires	M8	Done	
Requirements gathering	M8	Done	Both user and data requirements are defined. These may be refined during the project based on the feedback of the end users.
Recurrent Neural Networks based rice classification product	M16	Done	National scale application

Random Forest based rice classification product	M17	Done	For the South Korean provinces of Dangjin and Seosan
Change detection – land cover map update methodology (baseline)	M17	Done	
Integration of RNN based rice classification in the EOPEN platform	M18	Done	
End to end integration of the RNN based rice classification product – from data acquisition using the EOPEN Umbrella API to producing rice maps in the front end	M21	In progress	
Integration of change detection algorithm	M23	To be done	
Integration of the RF based rice classification algorithm in USTUTT's HPDA	M23	To be done	National scale application using distributed processing
Development of rice yield estimation product	M23	In progress	
Prototype demonstration	M24	To be done	Held in South-Korea

5.3.3 Monitoring of PUC 2 implementation against the User Requirements

Use case G-id	User of the EOPEN platform:	Current Status
PUC2_GB1	Must be provided with an agriculture monitoring system (based on earth observation data) that will provide more accurate cultivation area mapping and production estimates coupled with statistical data.	In progress
PUC2_GB2	Must be provided with a crop monitoring solution/platform that will integrate spatial and statistical data.	In progress
PUC2_GB3	Should be enabled to overlay the GIS data of "soil atlas" with remote sensing data. **<u>Hint</u>**: *National Institute of Agricultural Sciences provide a system called "Soil Atlas" but this system shows the agricultural land area according to land usage which does not show the actual crop production* *n area and location.*	In progress NOTE: All PUC 2 products will be provided standard formats and will be completely reusable. Both downloadable but also available on the platform that allows users to import their own data as well.
PUC2_GB4	Should be provided with services that account for big data handling, meteorological data coupling with statistical yield estimations and satellite data.	In progress
PUC2_GB5	Must be enabled to integrate different platforms to the governmental statistics system.	
PUC2_GB6	Must be provided with data at farm level (e.g. crop type classification, etc.)	Done
PUC2_GB7	Must be enabled to download agriculture related information through web.	Done
PUC2_GB8	Should be enabled to receive agriculture related information through reports.	Will not be addressed NOTE: Information will be provided in the form of raster

		and vector layers, which is deemed to be far more efficient and useful than report.
PUC2_GB9	Should be enabled to integrate the produced agriculture related information into the national statistics system.	In progress The output products will be provided in appropriate format for integration

Use case D-id	**List of Datasets requested by stakeholders**	**Status**	**Description**
PUC2_DB1	High resolution remote sensing imagery	In progress	Timeseries of vegetation indices, crop growth indicators and rice maps at 10 m spatial resolution
PUC2_DB2	Detailed meteorological observation & forecasting data	In progress	Air temperature, Wind speed and direction, Precipitation, and Solar Radiation
PUC2_DB3	In field inspection data	Done	Land cover maps available with 10 yr interval. These maps are updated through change detection for the year of inspection
PUC2_DB4	Farmers' claims data	Not available	
PUC2_DB5	Accurate yield statistics	To be done	Rice yield estimation product
PUC2_DB6	EO based production status	To be done	Rice yield estimation product and intermediate products (timeseries of vegetation indices)
PUC2_DB7	Statistical data on national fertilizer usage	Not available	
PUC2_DB7	Statistical data on crop yields	To be done	Rice yield estimation

			product

5.3.4 Implementation

The first principal component of the implementation of PUC2 is the accurate, efficient and large scale mapping of rice in South Korea. Korean Ministry of Environment provides land cover maps in a fishnet grid, as shown in the figure below. The grid cell size of the level-3 land cover map is about 6.4 square km.

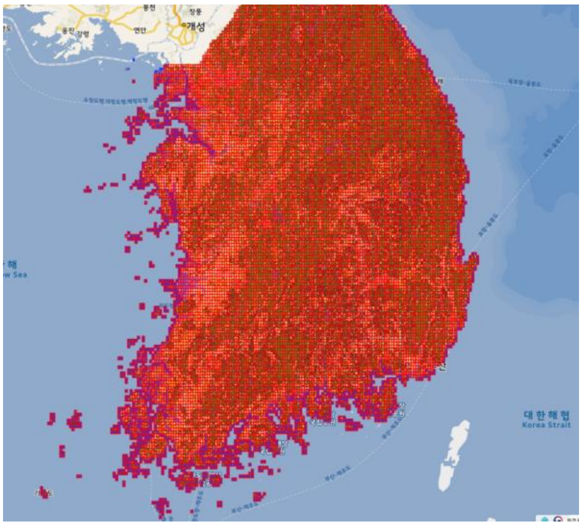


Figure 7: Fishnet grid of land cover maps in South Korea

Below we show the provinces of Dangjin and Seosan, whose land cover maps of 2015 were used to train the classification models.

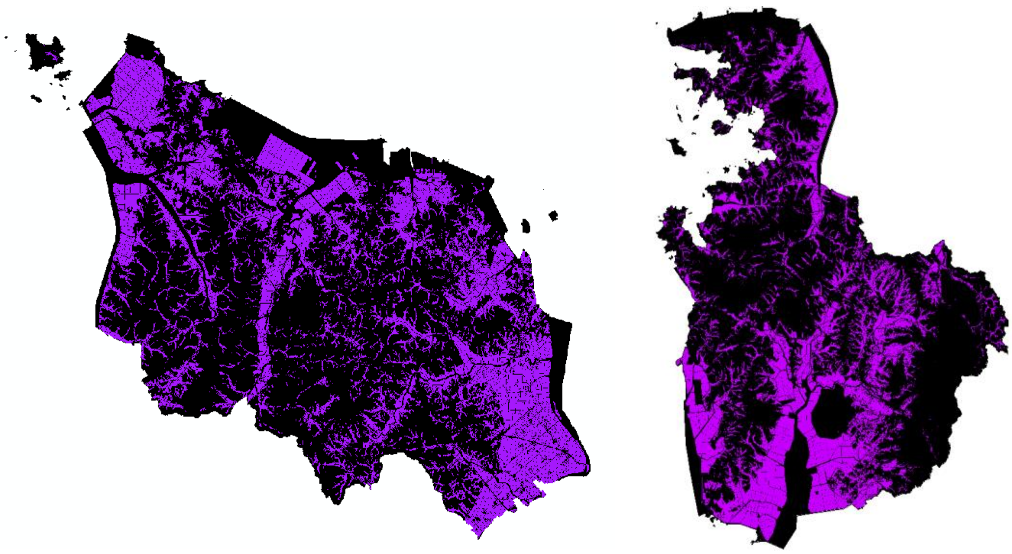


Figure 8: Provinces of Dangjin (left) and Seosan (right) showing rice fields with purple based on the level-3 land cover map distributed in 2015

The rice mapping classification framework for PUC2 involves two different approaches to be compared in terms of effectiveness, efficiency and generalization. The first involves the utilization of a time-series of Sentinel-1 images from 2015 to train a recurrent neural network model using the land cover maps, as presented above, and then fit the model on a Sentinel-1 time-series from 2018. The second method uses a combined input dataset of both Sentinel-1 and Sentinel-2 imagery to train a Random Forest classifier based on a dynamically updated land cover map for the year of inspection (currently 2018).

Method 1 – Recurrent Neural Networks

The first approach for rice paddy classification implements a Recurrent Neural Network (RNN), which is a time-series analysis method. The model was developed using a time series of Sentinel-1 (IW: VV) images, and trained with the dataset of 2015's Dangjin city, which is the greatest rice producer in South Korea. Validation was performed in two ways. First of all, detailed pixel-by-pixel comparison was performed in Dangjin and Seosan cities by using Level-3 land cover product, distributed by Korea Ministry of Environment. The accuracy was recorded as 91.46%.

To verify its general applicability on South Korea, the result was also compared with rice paddy area statistics of 160 cities in South Korea, which is distributed by Korea Statistics. A linear regression between the model result and statistics was "y = 0.8791x –470.5" (y: model result, x: statistics), and the R2 was 0.9691. The results were presented at the SAR Application Session in ISRS 2019.

The algorithm has been successfully tested on the local environment and is being tested on the EOPEN platform. The rice paddy detection algorithm has been integrated in the platform and successfully generated rice paddy maps for Dangjin.

Currently, the rice paddy mask generated in the local computing environment was evaluated with 40 plots of level-3 land cover map in South Korea (several years outdated products were updated to 2018 with visual interpretation using Google Earth Pro, Sentinel-1, Korean domestic street view services) and recorded 96.42% accuracy. The 40 plots are selected include the land use categories of rice agricultural, urban, mountain, and water. Among the 4 classes, agricultural has the lowest mean accuracy with 89.16% in 10 plots while urban, mountain, water recorded 98.02%, 99.77%, and 98.72% respectively. The errors were found mostly at the riverside and shaded areas.

Method 2 – Random Forests

Land cover map update

This development has been already fully reported in deliverable D4.1 EO assisted change detection, as it refers to a change detection methodology. It has been specifically developed to serve the PUC2 purposes via updating the rice map of year 2015 for the South Korean provinces of Dangjin and Seosan to reflect the reality of year 2018 (the selected year of inspection). Detailed land cover maps at the parcel level, distinguishing between rice and non-rice cultivations, are produced every few years in South Korea, as explained earlier. However, in order to systematically produce, for any given year, the envisaged products of rice mapping and rice yield estimation, updated land cover maps are required for the appropriate training and validation of the machine learning algorithms utilised. The output product of this change detection methodology does not attempt to accurately classify rice fields for 2018 but merely delete obvious changes in the land cover map of 2015. Therefore, using only rice pixels from the 2015 land cover map we eliminate outliers for the updated 2018 land cover map. This way the training dataset is refined for the machine learning algorithms to follow for the rice paddy classification product.

Random Forest classification

Using a long time-series of Sentinel-1 VV backscatter products, along with cloud free Sentinel-2 imagery and pertinent vegetation indices, we create a very deep feature space for classification. The figure below illustrates the high level workflow of the methodology. The workflow starts by searching for the appropriate input data using EOPEN's Umbrella API, which was developed as part of task T3.1. EO data acquisition and more information the service can be found in D3.1. The ingested time-series of imagery is then appropriately preprocessed and useful vegetation indices are computed for each Sentinel-2 acquisition (NDVI, PSRI, NDWI and SAVI)[7]. A subset, 10-20% of the updated land cover map is then used to train the Random Forest model, which it thereafter applied on the remaining dataset.

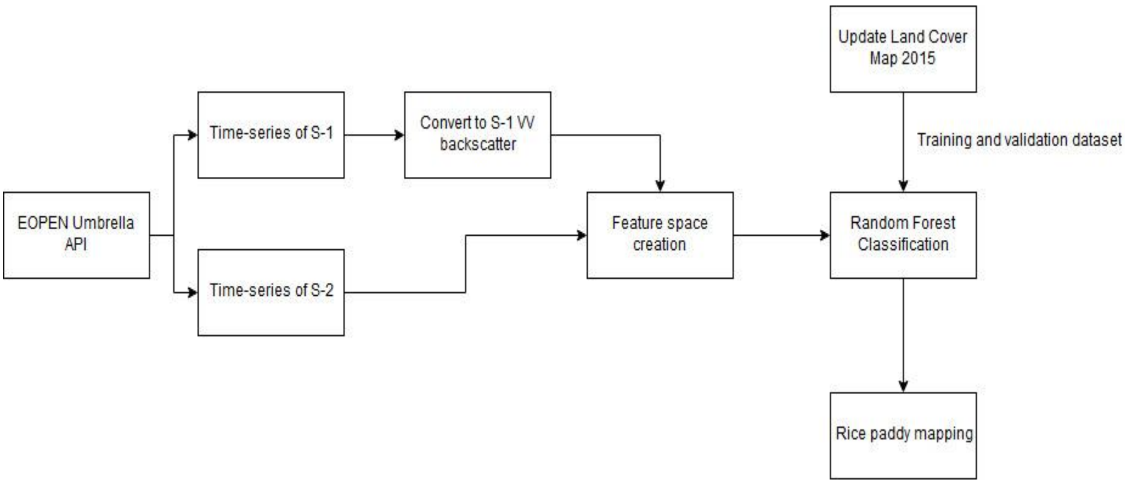


Figure 9: Random Forest rice paddy classification workflow

This rice classification method is currently in the process of being deployed in HLRS's High Performance Data Analytics (HPDA) module, looking into distributed processing solutions (Apache Spark) in order to achieve a national scale application. This approach constitutes a Big Data problem in terms of volume of data (multiple TB of information), processing complexity and scale of application and thus cannot be realized in conventional machines.

Intermediate evaluation and results

Table 23: Overall accuracy, producer's accuracy, user's accuracy for pixel and object based random forest classifications

Feature space	Algorithm	OA	PA	UA

[7] https://www.sentinel-hub.com/develop/documentation/eo_products/Sentinel2EOproducts

Object based	RF	91%	86%	86%
Pixel based	RF	92%	90%	85%

The object based classification for 2018 was performed using the parcels as provided in the 2015 land cover map (81351 parcels of agricultural land). 20% of those parcels were used for training using the labels of 2015. The parcels have changed significantly over the years and it was decided that a pixel-based approach would be more appropriate

Both training and validation were performed using the labels from the 2015 land cover map; therefore accuracies do not reflect the truth completely. Photointerpretation was performed to evaluate the instances when the classifier showed rice pixels, while the 2015 land cover map showed non rice, justifying the slightly low user's accuracy. We inspected the remaining 15% from rice user's accuracy and found that more than 90% of those instances were indeed rice-suggesting a user's accuracy more than 90-95%.

Rice yield estimation

The rice yield estimation methodology has not been yet finalized. However, since the acquisition of pertinent non EO data (weather data, fertilizer usage, cultivation practices etc.) is rather challenging, we hereby present an almost exclusively EO based approach.

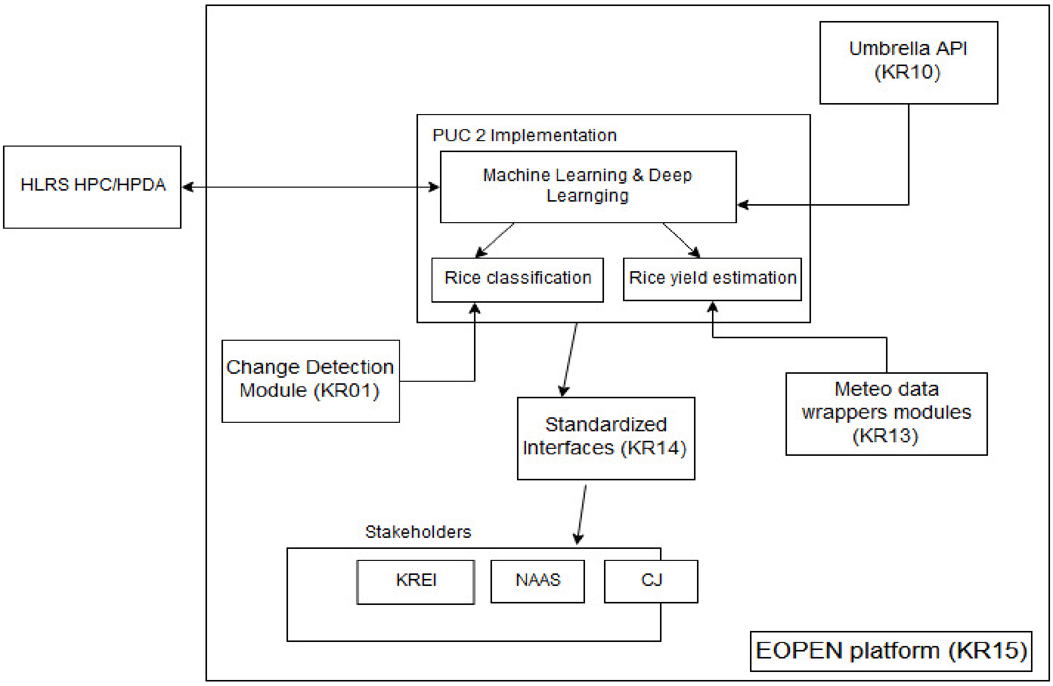
A gradual trend in yields has been identified from year, which is attributed to new technological developments, fertilizer application and improved rice cultivation management (Huang et al., 2013).

$$Y = Y_t + Y_{RS}$$

Y_t represents the component that is regulated by agricultural technology, both chemical and mechanical, but also management practices. Y_{RS} is the component defined by natural environment conditions, which in this case will be simulated by a time-series of vegetation indices, such as NDVI, NDWI etc. Historical yield statistics, available at the province level, will be used to estimate the correlation between natural environment conditions and yield and thereafter model the Y_t to correct/calibrate the NDVI v. yield relationship.

If data availability allows it, look into more sophisticated rice yield estimation models such as ORYZA requiring weather data (solar radiation, wind speed, vapor pressure), soil information, rice variety, water depth etc.

5.3.5 PUC 2 Implementation workflow



5.3.6 Potential Test Scenarios

Scenario 1

The first scenario involves the RNN rice paddy detection algorithm that trains based on the land cover map of 2015. The trained model will be applied to time-series of Sentinel 1 images, at national scale, for the years 2015-2018. This way the generalization of the model will be evaluated, identifying the significance of the accuracy decrease in the succeeding years of 2015. In addition correlations with economic data will be explored and analytics will be extracted on the spatiotemporal evolution of rice over the years.

Scenario 2

The second test scenario involves the Random Forest rice classification approach that utilizes the dynamic land cover map update methodology to create a training dataset for the year of inspection. Train the model using the updated land cover maps of Seosan and Dangjin and apply it on the rest of the country. The generalization of the model will be thereby evaluated.

Scenario 3

Following the example of Scenario 2, this third one will exploit updated land cover maps of other South Korean provinces (if available), and evaluate the differences in accuracies between the two scenarios. Namely, the Seosan-Dangjin trained model will be tested in other provinces against locally trained models. This way we can identify an operational framework for rice mapping, evaluation the generalization and transferability of the scheme.

Scenario 4

This scenario involved the future implementation of rice yield estimation on whose outputs a correlation analysis will be performed with any available economic metrics and data. The accuracy of yield estimation will be evaluated from early on in the year, namely with fewer features, up until the end of the cultivation season. This way we can determine the feasibility of providing accurate estimations on yield early in the year, allowing the stakeholders to take appropriate actions and make timely decisions in advance.

5.4 PUC 3 – Monitoring Climate Change through Earth Observation

5.4.1 Use Case story

In the present and future climate change environment, the average temperature in Finland is rising / will rise more (2°C by 2040), and precipitation will increase faster (5–10% by 2040) than the global average. The changes are affecting winters more than summers with the largest changes in the northern part of the country: Finnish Lapland. Since the early 2000s, Finland has taken a pro-active role in managing the Climate Change situation nationally, with mitigation and adaptation plans. The activities of the Transportation sector runs deeply through the sectors: sustainable industry, land use planning and construction, tourism and recreation as summers become warmer, wetter and longer and snow packed regions shift northward. Our use case begins with historical snow and temperature data, supplemented by EO data, which support Finnish Transportation Agency (FTA)'s current and future road maintenance for the Finnish drivers and riders. Our Use Case continues with temperature and snow data support for the Finnish Lapland communities who are, and will be, experiencing the greatest climate change consequences. The reindeer herders are the frontline people experiencing Climate Change.

Table 24: PUC 3 main stakeholders

Stakeholder	Description
Reindeer Herders	Region: Mostly Northern Finland, Service to be delivered: decades-long time-series of observations (snow depth and air temperature). Short-term and long-term adaptation priorities affected by climate change: • warmer winters → additional feeding • ice under or in the snow pack → digging problems for reindeer • more rain in summers → moulding → reindeer health issues • more extreme weather events • changes in pastures → extra work for herders o Know-how is from experience on responses to short-term changes, therefore, no strategies for long-term changes. o Lack of information for planning which pastures to use in near future (1mo – next year), daily weather observations only cover for the next couple of days o Reindeer Herders need factual data over 10 years for filing financial compensation for 'natural accidents', e.g. extreme climate events. o Reindeer Herders compete for their land-use with foresters, wind farms, miners, transport developers and the Arctic Railway concept. Resolutions between competing parties will require deep and complex impact assessments which include future environmental conditions.
Reindeer	They are more experienced at data usage than reindeer herders

Researchers	Region: Mostly Northern Finland, service to be delivered: decades-long time-series of observations (snow depth and air temperature): • Due to the fine-resolution of reindeer herders' snow-conditions intuitive-assessments, more and longer-term snow-temperature data is needed. • With such data, the reindeer researchers can use snow observations and ground ice observations and simulations, to form a characterization / assessment for the herders for problematic snow: deep snow, late snow melt, icy snow, and ground ice.
Finnish Transportation Infrastructure Agency (FTIA)	Region: All of Finland. Service to be delivered: decades-long time-series of observations (snow depth and air temperature): • Near-zero temperatures are becoming more common due to climate change • Continuous cycle of freezing and thawing increase stress on the road surfaces → increased need for de-icing and road maintenance • Increase in river flow rates increase wear on bridge structures • Need to adapt and refine strategies and prepare sufficient maintenance funding • Insufficient monitoring of snow conditions on lower-class roads

The PUC3 Stakeholders provide the on-the-ground view and the national view for mitigating Climate Change in Finland.

Reindeer Herders: Herding is small-scale livelihood in Northern Finland with about 5000 herders. Directly exposed to changes in land use and climate and have an intuitive, high fidelity knowledge of climate conditions to support reindeer care. Reindeer activities drive tourism in Northern Finland → 90% of tourists are from foreign countries.

Reindeer Researchers: Due to reindeer herding's large Finnish social, cultural and economic importance → multiple research institutes provide financial support to studies. The Reindeer Researchers study herding's economic development and sustainability, the herders' adaptation to / mitigation of Climate Change, and the herders' effects on land-use and environment.

Finnish Transportation Infrastructure Agency (FTIA): Responsible for transportation and road maintenance in Finland. Finland has a total of 78 000 km of public roads, many in sparsely populated regions. Transportation in general accounts for 20% of all greenhouse gas emissions in Finland, out of which road transport in particular accounts for 90%. Climate change is expected to impact all facets of the transport system: the infrastructure, modes of transport and operations and, especially, the maintenance. The FITA must make decisions based on yearly trends in weather and climate.

5.4.2 Timeline of PUC 3 implementation

Milestone	Completion month	Status	Comments

Questionnaires	M8	Done	
Requirements gathering	M8	Done	Both user and data requirements are defined. These may be refined during the project based on the feedback of the end users.
Phone interviews with stakeholder representatives	M10, M18	Done	FMI conducted phone interviews in order to refine the user requirements
Road passability product	M21	To be done	
Integration of change detection algorithm	M24	To be done	
Data gathering	M21	In progress	
Data integration with platform	M22	In progress	
Prototype demonstration	M24	To be done	Held in South-Korea
User workshop	M28	To be done	Event may be held as a webinar instead

5.4.3 Monitoring of PUC 3 implementation against the User Requirements

Requirement ID	Description	Current status
PUC3_GC1	Should be provided with a user friendly interface.	In progress
PUC3_GC2	Should be provided with easy access and management of datasets.	In progress
PUC3_GC3	Must be provided with data integration capabilities.	In progress
PUC3_GC4	Must be enabled to overlay satellite and ground observation data over a map background.	In progress
PUC3_GC5	Should be provided with the capability to select a user-defined area of interest.	In progress
PUC3_GC6	Should be enabled to visualize real time	In

	situation (most recent data).	progress
PUC3_GC7	Must be enabled to browse historical observations.	In progress
PUC3_GC8	Should be provided with tools for statistical analyses of selected area, time period and dataset(s).	To be done
PUC3_GC9	Should be provided with tools for visualizing time series of datasets and statistical analyses.	In progress
PUC3_GC10	Should be provided with time series trend analysis and future projections capabilities.	To be done
PUC3_GC11	Must be enabled to compare multi-platform data (e.g. satellite vs satellite, satellite vs. ground observation).	To be done
PUC3_GC12	Should be provided with access to relevant tweets with specific key words through the platform.	In progress

Dataset ID	Dataset Name	Current status	Description
PUC3_DC1	GlobSnow SWE	Done	Gridded daily, weekly or monthly data for snow cover with 25 km resolution starting at 1995. Based on satellite and ground weather station data.
PUC3_DC2	SMOS Level 3 Freeze/Thaw	Done	Gridded daily soil state (Thaw/Frozen/Partially frozen) with 25 km resolution starting at 2010. Based on satellite and ground weather station data.
	Sentinel LST Product	To be done	Land Surface Temperature data from Sentinel satellites.
PUC3_DC3	FMI ClimGrid	Done	Grid-interpolated daily averages of air temperature with 10 km resolution starting at 1961 until 2018. Based on ground station data. Pre-computed maps (no API).
PUC3_DC4	FMI ClimGrid	Done	Grid-interpolated daily averages of snow cover maps with 10 km resolution starting at 1961 until 2018. Based on ground station data. Pre-computed maps (no API).

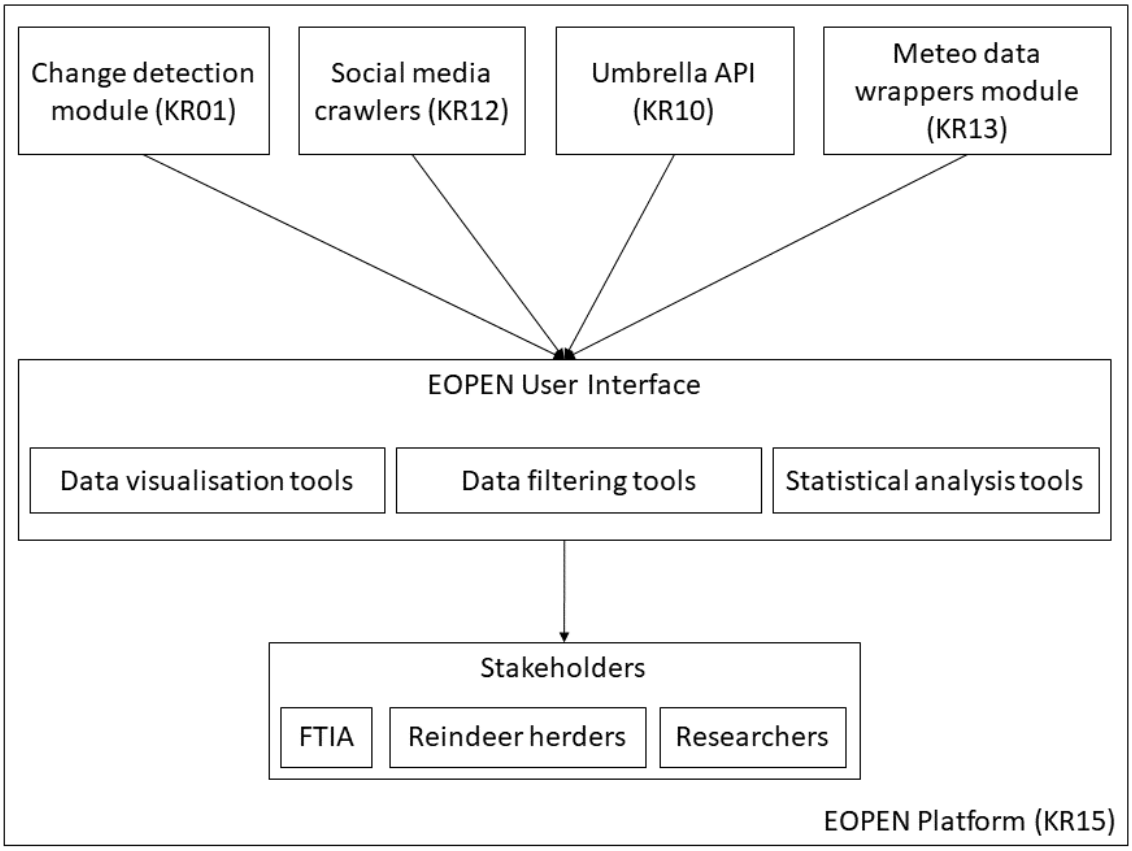
PUC3_DC5	FMI Open Data, Climatological values	In progress	Long-term (30-year) statistics of meteorological variables
PUC3_DC6	Social media data (twitter)	In progress	Annotated set of tweets of snow. FMI provides annotations, CERTH handles the rest.
PUC3_DC7	FMI Open Data, Climate model output	In progress	Estimated changes in climatological values over decades, based on multiple climate models and future climate scenarios.
PUC3_DC8	FMI Open Data, Weather observations	In progress	Instantaneous weather observations and daily, monthly, seasonal, and yearly values.
PUC3_DC9	FMI Open Data, Weather model output	In progress	Numerical weather prediction forecasts and model reanalysis, e.g. ERA5 from Copernicus Climate Change Service
PUC3_DC10	Regions of Finland	To be done	Region and municipality area borders in Finland. Pre-computed maps.
PUC3_DC11	Reindeer herding areas	To be done	Herding area borders in Lapland. Pre-computed map.
PUC3_DC12	FTIA Road maintenance classes	To be done	Road maintenance classification in Finland for snow removal and sanding.

User requirements PUC3_GC1 – PUC3_GC7 and PUC3_GC9 are user interface (UI) requirements, and thusly are provided by the user interface capabilities. FMI provides the Use Case with the meteorological and climatological data (datasets PUC3_DC1 – PUC3_DC5, PUC3_DC7 – PUC3_DC9), which is offered through the EOPEN UI. The UI implementation is currently in progress by SPACEAPPS. These requirements will be evaluated in the future via stakeholder questionnaires, with user satisfaction as the main performance indicator.

Of these requirements, PUC3_GC5 is given special consideration. In addition to giving the user an option to select their own area of interest, EOPEN will also provide predefined areas. The planned areas are Finnish region and municipality borders (PUC3_DC10), and reindeer herding districts (PUC3_DC11). Moreover, our stakeholders have indicated that road maintenance classification dataset would be useful, and so we plan to offer the data in EOPEN (PUC3_DC12).

Lastly, FMI provides the human annotations for Finnish tweets about snow coverage (see section Social media crawlers). The annotation work is required in order to satisfy user requirement PUC3_GC12 and to generate the dataset PUC3_DC6.

5.4.4 PUC 3 Implementation workflow



5.4.5 Potential test scenarios

Test scenario 1

Description: The winter 1972-1973 was exceptionally bad for reindeer herders due to 'deep icy snow' and 'deep snow' as reported by them (Vuojala-Magga et al., 2011). Due to these difficult snow conditions, more herding and additional feeding were required. Visualise the trends for snow depth over past winters. KPI: quantitative difference in snow depth, length of time to get the results

Data: PUC3_DC3, PUC_DC4, reindeer herding areas

Tools: Statistical measures, data filters

Criteria: User is able to compare snow depth maps between the good and the bad winters for particular reindeer herding areas in Lapland, and potentially find a statistically significant difference in snow depth.

Test scenario 2

Description: A recent paper (Anttila et al., 2018) showed that in the European Arctic zone the long-term trends for the onset of melting snow is getting earlier and the melt season is getting longer. Visualize the same trend in EOPEN. Compare the trends between satellite-based and ground-based data.

KPI: qualitative agreement with the trends, number of data sources/tools used, length of time to get the results

Data: PUC3_DC1, PUC3_DC3, PUC2_DC4, PUC3_DC5

Tools: Statistical and time series analysis tools, data filters

Criteria: User is able to show the long-term trend over time for the beginning of snow melting and its length in Northern Finland.

Test Scenario 3

Description: Climate change will increase flooding and cause permafrost to thaw in localized areas. FTIA wants to minimize gaps in road maintenance at low-priority regions and to improve their road maintenance planning. Visualize the trend of thawing on remote roads in Northern Finland

KPI: quantitative difference in thawing trends, number of data sources/tools used, length of time to get the results

Data: PUC3_DC2

Tools: Statistical and time series analysis tools, data filters, map filters

Criteria: User is able to show the trend in thawing over the winters in remote road areas in Northern Finland.

6 EVALUATORS

There are four different evaluator types (see below), who will be responsible for evaluating the different components of the EOPEN platform, but also the PUC end products and services. The different evaluator types add to the trustworthiness of the evaluation process but also provide different levels and perspectives of assessment.

a. Self-assessment
b. Cross-partner assessment (peer)
c. EOPEN user ecosystem.
d. 3rd party users (externals + PO + hackathon)

The *self-assessment* refers to the developer's evaluation and involved both quantitative metrics and qualitative assessment of the EOPEN component in question. The quantitative, boolean or qualitative performance indicators are specified for each of the project's KRs, in Chapters 3 and 4 of this document. One example of such indicators is the measurement of percentage difference in mean download speed between the Umbrella Sentinel Hub of EOPEN (KR10) and the Copernicus Open Access Hub.

The *peer assessment* evaluator refers to EOPEN partners that have not been involved in the development of a particular platform component but have used it in their own workflows. One example for this kind of evaluator would be the usage of the Umbrella Sentinels API by CERTH, within the change detection module (KR04) and specifically the flood detection workflow. CERTH will report on the number of broken links, for instance, during the module's utilization during the execution of PUC1 pilot scenarios.

The *EOPEN user ecosystem* evaluator includes the fully engaged end-users of the PUC products and services and end-users partners of the EOPEN consortium like AAWA. Finally 3rd party users will also evaluate the EOPEN platform components through the questionnaires, workshops (ACRS EOPEN workshop in South Korea), Infodays and hackathons that have been already organized.

7 EVALUATION RESULTS

In this first iteration of the PUC Implementation and Evaluation deliverable, the reporting was more skewed towards the implementation rather than the evaluation of the pilots. Nonetheless, the individual components of the platform, utilized within the PUC workflows, have been reported along with their evaluation metrics. Some preliminary evaluation results for the three PUCs have also been reported. In the second iteration of this deliverable (M28) the evaluation of the PUCs will be completed, including the interpretation of the KR specific KPIs from the perspective of the PUCs, but also including user feedback, cross-partner evaluation (peer), 3rd party users evaluation and EOPEN user ecosystem evaluation.

8 USER FEEDBACK

All PUCs will be evaluated by the end users that will have a hands-on experience with the relevant outputs of the Use Cases. The evaluation will be performed during dedicated events such as the EOPEN session in the ACRS conference in South Korea, or the Info Day in Vicenza, Italy. Updated questionnaires will be shared with stakeholders, after the platform and the pertinent products and services have been demonstrated to them. Finally, Q&A sessions will be another way of collecting user feedback during such sessions.

9 RECOMMENDATIONS, LESSONS LEARNT AND FUTURE PLANS

Recommendations, lessons learnt and future plans will be fully addressed in the second iteration of this deliverable (D7.2). The KR- and PUC-specific evaluations will be interpreted at a higher level relating back to the call requirements and thus assessing EOPEN's level of alignment with them. Collecting all the different individual evaluations we will assess how the EOPEN platform i) enables value adding services on generic data, ii) enables easy and user-friendly access to Copernicus data, iii) is interoperable, iv) optimizes the use of Copernicus data by non-traditional user communities to meet societal challenges. Finally, based on the overall EOPEN experience we will record lessons and learn and best practices for future EOPEN users.

10 REFERENCES

Anttila K., Manninen T., Jääskeläinen E., Riihelä A., and Lahtinen P. (2018). “The Role of Climate and Land Use in the Changes in Surface Albedo Prior to Snow Melt and the Timing of Melt Season of Seasonal Snow in Northern Land Areas of 40◦ N–80◦ N during 1982–2015”, Remote Sens. 10, 1619, DOI: 10.3390/rs10101619

Baevski, A., Edunov, S., Liu, Y., Zettlemoyer, L., & Auli, M. (2019). Cloze-driven Pretraining of Self-attention Networks. arXiv preprint arXiv:1903.07785.

Blondel, V., Jean-Loup, G., Renaud, L., Etienne, L., Journal of Statistical Mechanics: Theory and Experiment 2008 (10), P10008 (12pp) doi: 10.1088/1742-5468/2008/10/P10008. ArXiv: http://arxiv.org/abs/0803.0476

Florian, R., Ittycheriah, A., Jing, H., & Zhang, T. (2003, May). Named entity recognition through classifier combination. In Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4 (pp. 168-171). Association for Computational Linguistics.

Grubinger M, Clough P, Müller H, Deselaers T (2006) The iapr tc-12 benchmark: a new evaluation resource for visual information systems. In: International workshop OntoImage, pp 13–23

Huang, Jingfeng, et al. “Remotely Sensed Rice Yield Prediction Using Multi-Temporal NDVI Data Derived from NOAA's-AVHRR.” PLoS ONE, vol. 8, no. 8, 2013, doi:10.1371/journal.pone.0070816.

Ionescu B, Popescu A, Lupu M, Gînsca AL, Müller H (2014) Retrieving diverse social images at mediaeval 2014: challenge, dataset and evaluation. In: MediaEval

Nguyen, T. V. T., & Moschitti, A. (2012). Structural reranking models for named entity recognition. Intelligenza Artificiale, 6(2), 177-190.

P. Over, G. Awad, J. Fiscus, B. Antonishek, M. Michel, A. Smeaton, W. Kraaij, G. Quénot, "Trecvid 2011-an overview of the goals tasks data evaluation mechanisms and metrics", *TRECVID 2011-TREC Video Retrieval Evaluation Online*, 2011.

Tsikrika T, Kludas J (2010) The wikipedia image retrieval task. In: ImageCLEF. Springer, pp 163–183

Vuojala-Magga T., Turunen M., Ryyppö T., and Tennberg M. (2011). “Resonance strategies of Sami Reindeer Herders in Northernmost Finland during Climatically Extreme Years”, Arctic 64 (2), 227-241

Zanoli, R., Pianta, E., & Giuliano, C. (2009). Named entity recognition through redundancy driven classifiers. Proc. of Evalita, 9.