Big Data to Enable Global Disruption of the Grapevine-powered Industries

D8.5 - Evaluation Report and KPI Assessment

DELIVERABLE NUMBER	D8.5
DELIVERABLE TITLE	Evaluation Report and KPI Assessment
RESPONSIBLE AUTHOR	Aikaterini Kasimati (AUA)

Co-funded by the Horizon 2020
Framework Programme of the European Union

GRANT AGREEMENT N.	780751
PROJECT ACRONYM	BigDataGrapes
PROJECT FULL NAME	Big Data to Enable Global Disruption of the Grapevine-powered industries
STARTING DATE (DUR.)	01/01/2018 (36 months)
ENDING DATE	31/12/2020
PROJECT WEBSITE	http://www.bigdatagrapes.eu/
COORDINATOR	Panagiotis Zervas
ADDRESS	110 Pentelis Str., Marousi, GR15126, Greece
REPLY TO	pzervas@agroknow.com
PHONE	+30 210 6897 905
EU PROJECT OFFICER	Ms. Annamária Nagy
WORKPACKAGE N. | TITLE	WP8 | Grapevine-powered Industry Application Pilots
WORKPACKAGE LEADER	Agricultural University of Athens
DELIVERABLE N. | TITLE	D8.5 | Evaluation Report and KPI Assessment
RESPONSIBLE AUTHOR	Aikaterini Kasimati (AUA)
REPLY TO	akasimati@aua.gr
DOCUMENT URL	http://www.bigdatagrapes.eu/
DATE OF DELIVERY (CONTRACTUAL)	30 June 2019 (M18)
DATE OF DELIVERY (SUBMITTED)	28 June 2019 (M18)
VERSION | STATUS	1.0 | Final
NATURE	Report (R)
DISSEMINATION LEVEL	Public (PU)
AUTHORS (PARTNER)	Aikaterini Kasimati (AUA), Vasileios Psiroukis (AUA)
CONTRIBUTORS	Spyros Fountas (AUA), Maritina Stavrakaki (AUA), Florian Schlenz (GEOCLEDIAN), Simone Parisi (ABACO), Coraline Damasio (INRA), Pantelis Natskoulis (Symbeeosis)
REVIEWER	Coraline Damasio (INRA)

VERSION	MODIFICATION(S)	DATE	AUTHOR(S)
0.1	Initial ToC and document structure	10/03/2019	Aikaterini Kasimati (AUA), Vasileios Psiroukis (AUA)
0.4	Evaluation Report and KPIs - Template	07/06/2019	Aikaterini Kasimati (AUA), Vasileios Psiroukis (AUA)
0.8	Individual Piloting Evaluation Reports and KPIs lists	17/06/2019	Aikaterini Kasimati (AUA), Vasileios Psiroukis (AUA), Spyros Fountas (AUA), Maritina Stavrakaki (AUA), Florian Schlenz (GEOCLEDIAN), Simone Parisi (ABACO), Coraline Damasio (INRA), Pantelis Natskoulis (Symbeeosis)
0.9	Internal Review	26/6/2019	Coraline Damasio (INRA)
1.0	Partners review, final comments and edits	28/06/2019	Aikaterini Kasimati (AUA), Vasileios Psiroukis (AUA)

PARTICIPANTS	CONTACT
Agroknow IKE (Agroknow, Greece)	Panagiotis Zervas Email: pzervas@agroknow.com
Ontotext AD (ONTOTEXT, Bulgaria)	Todor Primov Email: todor.primov@ontotext.com
Consiglio Nazionale DelleRicherche (CNR, Italy)	Raffaele Perego Email: raffaele.perego@isti.cnr.it
Katholieke Universiteit Leuven (KULeuven, Belgium)	Katrien Verbert Email: katrien.verbert@cs.kuleuven.be
Geocledian GmbH (GEOCLEDIAN Germany)	Stefan Scherer Email: stefan.scherer@geocledian.com
Institut National de la Recherché Agronomique (INRA, France)	Pascal Neveu Email: pascal.neveu@inra.fr
Agricultural University of Athens (AUA, Greece)	Katerina Biniari Email: kbiniari@aua.gr
Abaco SpA (ABACO, Italy)	Simone Parisi Email: s.parisi@abacogroup.eu
SYMBEEOSIS LONG LIVE LIFE S.A. (Symbeeosis, Greece)	Kostas Gardikis Email: c.gardikis@gmail.com

ACRONYMS LIST

AA1 (DPPH)	Antioxidant Activity 1 (2,2-DiPhenyl-1-PicrylHydrazyl)
AA2 (ABTS)	Antioxidant Activity 2 (2, 2'-Azino-Bis-3-ethylbenzoThiazoline-6-Sulfonic acid)
BA	Biological Activity
BDG	BigDataGrapes
D	Deliverable
DSS	Decision Support System
ECa	Electrical Conductivity
FAIR	Findable, Accessible, Interoperable, Reusable, as set of principles acting as an international guideline for high quality data stewardship
FRAP	Ferric Reducing Antioxidant Power
L8	Landsat-8
LAI	Leaf Area Index
MAC	Maceration
MTT	3-(4,5-diMethylThiazol-2-yl)-2,5-diphenylTetrazolium bromidefor
NDRE	Normalized Difference Red Edge Index
NDVI	Normalized Difference Vegetation Index
pH	Potential of Hydrogen
RI	Refractive Index
S2	Sentinel-2
SIRT1	Sirtuin 1
SVIs	Spectral Vegetation Indices
TFC	Total Flavonoids Content
TMC	Total Microbial Count
TPC	Total Phenolic Content
UAE	Ultrasound Assisted Extraction
WP	Work Package
Y&M	Yeast and Moulds count
KPIs	Key Performance Indicators

EXECUTIVE SUMMARY

The deliverable D8.5, "Evaluation Report and KPI Assessment", aims to provide a report on the results of the application piloting sessions, in line with the defined experimental protocols and in accordance with the evaluation methodology. Based on the experimental and evaluation protocols defined in the context of T8.2, pilot partners monitor the execution of the application pilots and evaluate their results qualitatively and quantitatively against the Key Performance Indicators (KPIs) defined by the relevant protocols. The objective of this deliverable is to provide, in the first part, an overview and a first evaluation regarding each of the four pilots' progress, stating and explaining the current status of development, the successfulness of implementation and the achieved performance of the BigDataGrapes Pilots. An updated version of this deliverable, including pilots' final evaluation, is due to M36 of the project lifetime.

Deliverable D8.5, "Evaluation Report and KPI Assessment", is based on the individual Qualitative and Quantitative Evaluations of the following pilots: Table and Wine Grapes Pilot (AUA), Wine Making Pilot (INRA), Farm Management Pilot (ABACO & GEOCLEDIAN) and Natural Cosmetics Pilot (Symbeeosis). This document reports an overview of the application piloting sessions' results, both qualitatively and quantitatively against the defined KPIs, for each of the four pilots. Information was directly provided by the pilot leaders to ensure the specificity of the guidelines and the smooth progress and evaluation of all piloting sessions.

The document is structured as follows: Chapter 1 serves as an introduction to the deliverable whereas Chapter 2 provides an overview of the four pilots' Qualitative and Quantitative evaluation. Each of pilots' Qualitative Evaluation Summary is divided in five sections: the pilot's evaluation summary, plan progress, status of implementation, impact and potential pilot modification or required plans for improvement. For the Quantitative evaluation a KPIs list along with baseline values have been defined, including the Domain Specific KPIs only for those pilots that this applies and the Technological KPIs for all pilots. This part of the deliverable provides tangible results. Chapter 3 presents the structure of the BigDataGrapes Survey, the planning of the questionnaire, which is to be distributed to the end-user a few months after the completion of the Intermediate phase (M22) and two months before the end of the Summative phase (M34). The survey is divided in the Pilot Basis Evaluation and the Technological Basis Evaluation, in a similar way to Chapter 2. Finally, the last chapter of the deliverable consists of the overall conclusions regarding the evaluation of the pilots, including discussion and suggestions that could improve their impact.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Data-driven approaches have the potential to improve decision making in different industries and settings. This is also the case for the grapevine-powered industries, where it is evident that a rich, large-scale and diverse data pool is needed for carrying out the foreseen research and industry-centred activities. Therefore, throughout the lifecycle of the project, BigDataGrapes will continuously collect and monitor sensor data derived from all experimental sites accessible by the pilot partners, generating an expansive and diverse collection of datasets. These streams of data and datasets serve as the basis for carrying out research and technical work and are being used as the testbed for enabling the implemented technical components to efficiently handle the volume and intricacies of these data, acquired from realistic in-field conditions. A data marketplace demonstrator using these data assets harvested, serves as the project's experimentation environment where testing and adjustments of the proposed technical solutions can be carried out in a realistic setting. As the project progresses, the data pool is continuously enriched in volume and range, in accordance with the needs and requirements of the covered use cases.

Moving from testing in laboratory conditions to testing in real-world settings, BigDataGrapes has designed and is executing application pilots pertaining to the defined Use Cases, under WP8, "Grapevine-powered Industry Application Pilots". This work package is responsible for the pilots' planning and preparation, the definition of the experimental and evaluation protocols to be followed, the execution of the pilots and ultimately, the collection and evaluation of the pilot results and their assessment over indicators defined by the end users. In this context, this document provides a pilots' evaluation, a report on the results of the executing application piloting sessions during the BigDataGrapes project lifetime, in line with the defined experimental protocols and in accordance with the evaluation methodology.

The development of fully defined demonstrators for each of the grapevine-powered industry use cases allow to showcase and evaluate the BigDataGrapes platform and components in the context of specific end-user requirements from the different pilots. This evaluation is centered on realistic, strictly defined experiments that reflect real-life operations of the related industries. A detailed evaluation plan has been produced and is being followed during pilot execution, following an iterative approach of assessment that will be performed according to the proposed three-phase human-centred activities, organised in the following phases:

Formative phase, leading up to the "Use Case Definition & Assessment Planning" (M9): Industry-centred requirements and the concrete use cases where the BigDataGrapes solution will be applied and tested against these requirements will be defined through WP2 and WP8. During this phase, suitable data and processes for fulfilling the requirements of the specific use cases will be identified and relevant piloting activities will be defined.

Intermediate phase, leading up to "Functional Assessment Sessions" (M18): The first round of controlled pilot trials, will implement a first version of the pilots, using the first versions of newly developed BigDataGrapes components. These will be restricted piloting trials in terms of scale and complexity. The objectives of these trials are (a) to provide data for the assessment of early BigDataGrapes components and (b) to refine the pilots themselves into their subsequent iterations.

Summative phase, leading up to "Operational Assessment" (M36): The final phase entails the validation of the BigDataGrapes components in real-life conditions and with realistic complexity. The components will be used throughout the timespan, with developments in the technologies incorporated opaquely in the operational platform. A summary of the operation of the system and the respective pilot observations will be delivered, followed by a final Evaluation report where the performance of the system will be assessed against the established evaluation criteria and the appropriate Key Performance Indicators.

2 QUALITATIVE AND QUANTITATIVE EVALUATION

Evaluation within Task 8.5 is to be both formative and summative. The former is essentially self-assessment and will be carried out by all partners through filling the "Qualitative and Quantitative Evaluation", a part of the evaluation that provides tangible results via this deliverable. The summative evaluation will involve external as well as internal evaluation in the form of "BigDataGrapes Pilots' Survey".

This part of the deliverable is dedicated to the qualitative and quantitative evaluation of the BDG pilots. All four pilot partners (Table and Wine Grapes Pilot- AUA, Wine Making Pilot- INRA, Farm Management Pilot- ABACO- GEOCLEDIAN, Natural Cosmetics Pilot- SYMBEEOSIS) provided a report on the results of the application piloting sessions, in line with the defined experimental protocols and in accordance with the evaluation methodology in the context of T8.2, in order to evaluate their progress, stating and explaining the current status of development, the successfulness of implementation and their achieved performance.

The Evaluation consists of two parts the "Qualitative Evaluation Summary" and the "Quantitative Evaluation against the KPIs". Each of pilots' Qualitative Evaluation Summary is divided in five sections:

- ✓ Pilot's Evaluation Summary, including specific objectives, achievements/results and problems/challenges
- ✓ Plan Progress, providing details on the conducted tasks in the reporting period
- ✓ Status of Implementation, including the actors involved, methodology, deployed components and gathered data and format
- ✓ Impact of each piloting session
- ✓ Potential Pilot Modifications or required plans for improvement.

For the Quantitative evaluation a KPIs list along with baseline values have been defined, including the Domain Specific KPIs and the Technological KPIs for all pilots. The overall goal of the Technological KPIs is to assess the data generation process within the pilots and specific Use Case Scenarios. These KPIs include big data metrics such as the 3Vs (Volume, Variety, Velocity). The pilots are in a start-up phase. Therefore, the baseline values for the KPIs have been defined by the pilot partners, using various sources, such as collected data (both from the first year of the application pilot's lifetime and/or historical data owned by the piloting partner), available literature, statistical data, sector average, and expert knowledge. The baseline values will be used to calculate the effect of the pilots.

The aim of the evaluation process is to clearly distinguish methodological and functional testing from real-world evaluation by actual users, to ensure that technical progress is directed towards improving realistic applications of the technologies. A detailed evaluation for the BDG technical outcomes under the specified use cases, are recorded in the form of an evaluation report for each pilot execution, briefly describing the involvement of the pilot and presenting the obtained results against the evaluation criteria and the appropriate Key Performance Indicators (KPIs).

2.1 TABLE AND WINE GRAPES PILOT INDIVIDUAL EVALUATION (AUA)

2.1.1 Qualitative Evaluation Summary

In order to report the Table and Wine Grapes Pilot's progress, AUA has completed the following table with the necessary information, regarding the current status of development, the successfulness of implementation, its impact and potential modifications.

AUA	Table and Wine Grapes Pilot Qualitative Evaluation Summary

Pilot Evaluation Summary

Specific Objectives

Deriving meaningful knowledge from many relevant, yet heterogeneous data sources is very important and will act as the basis for future decision-making processes. Throughout the lifecycle of the project, AUA will continuously collect and monitor sensor, farming and phenological data derived from all test sites located in Greece.

Soil properties, climate conditions and cultivation techniques constitute significant variables, which affect the quality of the final product. In particular, soil data (soil texture, soil electrical conductivity etc.) and weather data (average temperature, humidity etc.) affect both crop quality data (sugar content, anthocyanins content, phenolic compounds concentrations etc.) and crop quantity data (crop yield, berry weight and size etc.).

Some of the goals to be achieved through this sensor and farming data collection, is to denote associations and correlations between precision agriculture information and phenological data and grape and wine chemical analysis. Location-specific data will be used as auxiliary sources and will lead to the supply of vegetation indexes corrected for vineyard cultivation practices, more accurately determined vegetation stages and input to plant performance and grape quality indicators among others. Finally, the ultimate goal is to correlate the aforementioned data with earth observation data to examine the effectiveness of applying machine learning techniques and eventually train the relevant machine learning components.

The Table and Wine Grapes pilot is linked to the Prediction and Farm Management Use Cases and more specifically to four Scenario Hypothesis, namely B1. Yield Prediction, B3. Crop Quality Prediction for Optimizing Post Harvest Treatments of Table Grapes, C1. Optimisation of Farm Practices in the Vineyard, C2. Management Zones Delineation for Vineyards, as shown in Figure 1.

Table 1: Scenario Hypothesis served by the Table and Wine Grapes Pilot

Use cases	Use Case Scenarios	Partner
A. Data Anomaly Detection & Classification	A. Earth Observation Data Anomaly Detection & Classification	A. ABACO-GEOCLEDIAN
B. Prediction	B1. Yield Prediction B2. Predicting Biological Efficacy B3. Crop Quality Prediction • for Optimizing Post Harvest Treatments of Table Grapes (B3-1) • for Optimizing Winemaking (B3-2)	B1. AUA B2. Symbeeosis (APIGEA) B3-1. AUA B3-2 INRA
C. Farm Management	C1. Optimization of Farm Practices in the Vineyard C2. Management Zones Delineation for Vineyards	C1. ABACO-GEOCLEDIAN AUA C2. AUA

Achievements/Results

The first step to the successful implementation of the piloting activities for

the Table and Wine Grapes pilot was to define the appropriate "Data and Datasets" to be collected throughout the project's lifetime. The second step was to update and refine the Scenario Hypothesis and create a link to the data. Thus, during the first crop season/experimental period, which belongs to the Formative phase, suitable data and processes for fulfilling the requirements of the specific Use Cases were identified and relevant piloting activities took place.

This includes the collection of the following data:

- High quality topographic data have been collected from all three (3) test sites and topographic and elevation maps have effectively been generated.
- Georeferenced apparent soil Electrical Conductivity (ECa) data has been collected from two (2) test sites (Kontogiannis Estate and Palivou Estate) and could effectively be divided into management zones.
- Vegetation Indices-based maps have been generated from all six measurements and the different phenological growth stages of grapevines have been successfully monitored for the first season.
- Weather and soil data have been recorded effectively since the installation of the IoT stations on the fields.
- Yield data have been collected for the first harvesting period.
- Qualitative and quantitative data have been gathered from all three sites, which will be further statistically analysed and correlated with the data coming from the field for the first season.
- Access to Landsat-8 and Sentinel-2 satellite data for the test sites under the Table and Wine Grapes Pilot has been provided by GEOCLEDIAN, in order to enable temporal and spatial observations and analysis through their field monitoring service ag|knowledge.

After the completion of the data acquisition procedures by the end of the first crop season/experimental period data management activities took place, including Data Normalisation, Data Modelling, Semantics Annotation and transformation to RDF among others.

Problems/Challenges

Soil electrical conductivity mapping and in general ECa non-touch sensing systems are extremely sensitive to interferences from nearby metallic objects. Out of the three (3) experimental sites of the Table and Wine Grapes pilot, Palivou estate and Kontogiannis estate are open-field vineyards, allowing for successful ECa data collection. In Fasoulis estate, however, vines are covered with a light nest-like fabric approximately 0.5 meters from the top of the canopy, placed on metallic pillars across the entire vineyard. For this reason, no accurate ECa data could be collected from this site.

Another hardship we encountered was a hardware malfunction of the EM-38 MK2 sensor on the first data collection visit to the experimental sites. The handheld logger, used to wirelessly store data and control the sensor's functions, lost Bluetooth contact with the sensor during data collection on Palivou estate, demanding a direct cable connection between logger and sensor. This way, the sensor was no longer mount-able to the tractor, forcing the AUA team to perform the data collection manually.

Pilot Plan Progress	Measurements related to the Table and Wine Grapes pilot successfully took place throughout the first crop season of the pilot on the three (3) experimental sites. As it is natural, emphasis was given during the summer months (May through September) when grapevines grow and produce grapes and therefore allowing for crop monitoring data collection, which is the procedure that generates the greatest volume of data of the pilot. More specifically, an overall plan of the experimental methodology of the pilot is presented below: At first, the boundaries of the vineyards were geo-referenced using GPS technology at the very beginning of the project, as soon as the experimental fields were chosen. Time-stable zones have been formed using soil ECa mapping, assisted by elevation mapping using the RTK-GPS. These data related to the boundaries, management zones and elevation will be used throughout the course of the project. These measurements mentioned above have been performed by the AUA Precision Agriculture lab. - *Dataset Name:* Topographic data and elevation maps - *Data Description:* Spatial data (boundaries and elevation data) - *Data collection frequency: a single measurement per crop season* Soil and weather data have been continuously collected since Day 1 of the project. The setup of the IoT stations and the data acquisition have been performed by the AUA Precision Agriculture lab. - *Dataset Name: IoT Stationary data* - *Data Description: Soil moisture and meteorological parameters* - *Data collection frequency: Constant data stream throughout the entire year* Canopy reflectance data and vegetation indices have been recorded with the use of Crop Circle ACS-470, SpectroSense2 and Crop Circle RapidSCAN CS-45 sensors six times per season/summer starting at the beginning/middle of May, so that the phenological development of the grapevine, which is divided into nine (9) principal growth stages, will be followed in the best way. The measurements mentioned above have been performed by the AUA Precision Agriculture lab. - *Dataset Name: Canopy sensing and vegetation indices* - *Data Description: Canopy reflectance and vegetation indices data* - *Data collection frequency: six (6) per crop season, with a minimum of two (2) sensors in each data collection* Landsat-8 and Sentinel-2 satellite data have been collected for the same periods with the canopy reflectance measurements, again six times per season/summer. The satellite images have been provided by GEOCLEDIAN. - *Dataset Name: Satellite data* - *Data Description: Landsat-8 and Sentine-2 imagery of the test sites* - *Data collection frequency: six (6) per crop season* Some of the qualitative and quantitative characters of the grapevines, such as pH, soluble solids, total titratable acidity, antioxidant capacity by DPPH, FRAP assay, and aminoacids, have been tested three times over a season. Finally, the rest of the qualitative and quantitative characters have been

	assessed at the end of each season, when harvesting. The qualitative and quantitative measurements have been performed by the AUA Vitis lab. - *Dataset Name:* Classical analytical techniques (HPLC) - *Data Description:* Phenolic composition data - *Data Frequency: Laboratory analyses take place once every year, after harvest* The grape and berry mechanical properties (weight, length, width, density etc.), berry deformation, berry detachment, density, grape volume, berries diameter, berries weight for table grapes will be measured at post-harvest. - *Dataset Name: Grape and Berry Mechanical Properties* - *Data Description: Lab Measurements of berries' properties* - *Data Frequency: Laboratory analysis' take place once every year, after harvest* Finally, yield data has been collected and yield mapping has been performed during the harvest period, by both AUA labs. - *Dataset Name: Yield Mapping* - *Data Description: Yield data* - *Data collection frequency: 1/crop season*
Status of Implementation	*Actors Involved* AUA PA lab and Vitis Lab have successfully collected data throughout the first crop season from the three (3) experimental sites (Palivou estate, Kontogiannis estate and Fasoulis estate), while GEOCLEDIAN provided the satellite imaging along with their data analysis service for the same sites. Agroknow, KULeuven and Ontotext participated as Tech providers, with Agroknow and Ontotext acting as model providers, while KULeuven was responsible for data visualisation. Finally, all pilot partners and tech partners provided insights in order to formulate and define the scenario hypothesis. *Methodology* The first step to the successful implementation of the piloting activities for the Table and Wine Grapes pilot was to define the appropriate "Data and Datasets" to be collected throughout the project's lifetime. The second step was to update and refine the Scenario Hypothesis and create a link to the data. Thus, during the first crop season/experimental period, which belongs to the Formative phase, suitable data and processes for fulfilling the requirements of the specific Use Cases were identified and relevant piloting activities took place. For the data collection, the methodology used is as described in D8.2 - Experimental Protocols and Evaluation Methodology. After the completion of the data acquisition procedures, by the end of the first crop season/experimental period, data management activities took place, including Data Normalisation, Data Modelling, Semantics Annotation and transformation to RDF among others. *Deployed Components* <u>Precision Agriculture Lab</u> • EM38-MK2 probe (Geonics LTD, Mississauga, ON, Canada) (Figure 1). Data collection is supported by the DAS70-AR Data Acquisition System (Archer Data logger). The EM38 measures apparent soil ECa in

millisiemens per metre (mS/m) in the root zone at 0.5 and 1.0 m depth and the in-phase ratio of the secondary to primary magnetic field in parts per thousand (ppt).

Figure1: EM38-MK2 (left) and Archer Data logger (right)

- HiPer V RTK GPS (Topcon Positioning Systems Inc., Livermore, CA, United States) (Figure 2). Records topographical data, such as field boundary points, and elevation data. The final output can be a KML, KMZ file. This measurement has been performed a single time throughout the course of the project, at the beginning of the table and wine grapes pilot, prior to all other measurements.

Figure 2: Topcon HiPer V RTK GPS

- Crop Circle ACS-470 (Holland Scientific Inc., Lincoln, NE, United States) (Figure 3). This remote sensing tool is measuring the radioactive transfer and the biophysical characteristics of plant canopies. It is an active crop canopy sensor that provides basic reflectance information from plant canopies and soil as well as classic spectral vegetative index data (NDVI, NDRE etc.). Vegetative indices measurements have been done in two different canopy parts, by the side and at upper canopy of the vines,

after mounting the equipment to a winegrowing tractor.

Figure 3: Crop Circle ACS-470

- SpectroSense2+ GPS (Skye Instruments Ltd, Landrindod Wells, UK) (Figure 4). Used to estimate the Leaf Area Index (LAI) and NDVI vegetation indices.

Figure 4. SpectroSense2+

- Crop Circle RapidSCAN CS-45 (Holland Scientific Inc., Lincoln, NE, United States). Used to estimate vegetation indices such as NDVI and NDRE indices (Figure 5).

Figure 5: Crop Circle RapidSCAN AC-45 handheld reflectance sensor

- Software such as Surfer 11 (Golden Software), ArcGIS (ESRI, Redlands, CA, USA), Global Mapper for the generation of thematic maps.
- Two Vantage Pro 2 weather stations (Davis Instruments Corp., Hayward, CA, United States) (Figure 6) with rain sensor, to detect rainfall, anemometer to measure wind speed and direction, air temperature sensor, air humidity sensor, barometer to monitor atmospheric pressure. The basic equipment can be supplemented with sensors for UV and solar radiation. The automatic weather station will be installed inside the vineyard. Weather information is being recorded throughout the growing season.

Figure 6: Vantage Pro2 Weather Station (left) and Decagon EC-5 soil moisture sensor (right)

- Four Decagon EC-5 soil moisture sensors (METER Group, Inc., Pullman, WA, USA) (Figure 6) recording throughout the growing season the humidity and temperature of the soil.

Laboratory of Viticulture

- Soluble solids will be determined using an ATAGO N1-a refractometer with a 0-32 Brix measurement range at 0.28 Brix increments.
- Total titratable acidity will be measured by titration with a 0.1 N NaOH solution and will be expressed as tartaric acid.
- The quantitative and qualitative analysis of the substances, which exist in berries, must and wines such as, organic acids, sugars, phenolic compounds, amino acids, anthocyanins, volatile compounds, etc. will be performed using an HPLC Shimadzu Nexera comprising a gradient pump Shimadzu Nexera X2, a ProStar model 410 AutoSampler, and a ProStar model 330 Photodiode Array Detector on a reversed-phase Waters C18 x select (250 mm x 4.6 mm, 5 mm) column.
- Antioxidant activity (2,2-diphenyl-1-picrylhydrazyl, DPPH) will be evaluated by the free radical scavenging activity of DPPH using a modified colorimetric method, while the reduction of the DPPH radical will be determined by measuring the absorption at 517 nm in a UV/Vis spectrophotometer (Perkin Elmer, Lambda 25, Beaconsfield, Bucks, U.K.). The absorption of the antioxidant activity (Ferric Reducing Antioxidant Power, FRAP) will be measured at 593 nm.
- Assimilation rate (photosynthesis) and stomatal conductance will be obtained by measurement of inlet and outlet CO_2 and H_2O relative concentration using a portable photosynthesis system (Li-6400XT, Li-Cor, Lincoln Nebraska, USA).
- Chlorophyll concentration will be measured on the leaves using a SPAD 502 (Konica Minolta, Europe).

Site Description

Three test sites have been chosen for data collection for BigDataGrapes in Greece. These are situated in the regional unit of Corinthia, in the north-eastern part of Peloponnese. The following have been selected: for winemaking Palivou Estate and Kontogiannis Estate and for table grapes

Fasoulis Estate.

Palivou Estate: is located in Nemea, planted with *Vitis vinifera* L. cv. 'Agiorgitiko' and 'Merlot' for winemaking. The row orientation is northeast-southwest, and the training/trellis system is VSP (vertical shoot positioned)- cane pruning, double Guyot.

Figure 7: Palivou Estate test site (Google Earth Pro)

Kontogiannis Estate: in Ancient Corinth having the same VSP -double Guyot or double Royat- training/trellis system planted with 'Roditis', 'Savatiano', 'Mavroudi' and 'Soultanina' for winemaking. Its row orientation is north to south.

Figure 8: Kontogiannis Estate test site (Google Earth Pro)

Fasoulis Estate: situated in Nemea, cultivated with 22 different table grape varieties, where each line has a different variety. The orientation is southeast to northwest.

Figure 9: Fasoulis Estate test site (Google Earth Pro)

Gathered Data and Formats

- Spatial data (boundaries and elevation data)
  - Dataset using for the pilot: Topographical and elevation maps
  - Data collection frequency: Once at the start of the pilot
  - Associated data model/format: csv, xls, xml
  - Data size: MB

- Geo-referenced apparent soil electrical conductivity (ECa)
  - Dataset using for the pilot: ECa maps
  - Data collection frequency: Once at the start of each year
  - Associated data model/format: csv, xls, xml
  - Data size: MB

- Canopy reflectance and vegetation indices using proximal sensors
  - Dataset using for the pilot: Vegetation Indices maps
  - Data collection frequency: Six (6) times during each crop season
  - Associated data model/format: csv, xls, xml
  - Data size: MB

- Canopy reflectance and vegetation indices using satellite imagery
  - Dataset using for the pilot: Vegetation Indices maps
  - Data collection frequency: Six (6) times during each crop season
  - Associated data model/format: GEOTIFF, png
  - Data size: GB

- Qualitative and quantitative characters; Grape and berry mechanical properties (weight, length, width, density etc.), berry deformation, berry detachment, density, grape volume, berries diameter, berries weight.
  For wine and table grapes: soluble solids, pH, total titratable acidity, total phenols and anthocyanins, total flavonoid content, total flavanol, flavonol, flavone content, tannins, antioxidant capacity (trans-resveratrol, piceid, ε-viniferin) by DPPH, FRAP assay, aminoacids
  - Dataset using for the pilot: Grape and Berry Mechanical Properties
  - Data collection frequency: Once at the end of each crop season
  - Associated data model/format: csv, xls, xml

- Data size: MB

- Full phenolic profile of grapevine varieties in correlation with the phenological stages to improve the quality of viticultural products
  For table grapes: leaf analysis, foliar chlorophyll contents photosynthetic pigment content of the leaves, water potential correlated to the proline content
  - Dataset using for the pilot: Phenolic composition data
  - Data collection frequency: Once at the end of each crop season
  - Associated data model/format: csv, xls, xml
  - Data size: MB

- Yield data
  - Dataset using for the pilot: Yield variation maps
  - Data collection frequency: Once every year, during harvesting
  - Associated data model/format: csv, xls, xml
  - Data size: MB

- IoT stationary data
  - Dataset using for the pilot: Soil moisture data, meteorological parameters
  - Data collection frequency: Constant data stream throughout the year
  - Associated data model/format: csv, xls, xml
  - Data size: MB

Table 2: Table and Wine Grapes Pilot Data and Datasets

Name	DataSet Description	Priority	Provenance	Data Type Format	Data size
Yield Mapping	Yield data	Essential	Laboratory equipment	csv, xls	MB
Grape and berry mechanical properties	Measurements	Essential	Laboratory equipment	csv, xls	MB
Classical analytical techniques (HPLC)	Phenolic composition data	Essential	Laboratory equipment	csv, xls	MB
Topographic da ta and elevation maps	Spatial data (boundaries and elevation data)	Essential	Remote sensing	csv, xls, xml	MB
Canopy sensing and vegetation indices	Canopy sensing data	Essential	Proximal sensors	csv, xls	MB
IoT stationary data	Soil moisture data, meteorological parameters	Essential	IoT data	csv, xls	MB
Drone imagery	Drone images	Essential	Multispectral and thermal cameras	GEOTIFF	GB
Crop Calendar	Records of crop groth stages and agricultiral operations	Essential	Log files	doc, xls	MB
Eca sensing	Geo-referenced soil electrical conductivity data	Essential	Proximal sensors	csv, xls	MB

Impact	Some of the goals to be achieved through this sensor and farming data collection, is to denote associations and correlations between precision agriculture information and phenological data and grape and wine chemical analysis. Location-specific data will be used as auxiliary sources and will lead to the supply of vegetation indexes corrected for vineyard cultivation practices, more accurately determined vegetation stages and input to plant performance and grape quality indicators among others. Finally, the ultimate goal is to correlate the aforementioned data with earth observation data to examine the effectiveness of applying machine learning techniques and eventually train the relevant machine learning components. The Table Grape and Wine pilot aims to achieve the following: • Allow for data-based predictions to be made with accuracy levels higher than ever before. Combination of improved resolutions (spectral, spatial and temporal) of remotely sensed images, coupled with more precise on-the-ground multiple data sources, such as soil, weather and vegetation indices derived from sensors can create powerful models able to deliver predictions taking into consideration a plethora of factors and variables, drastically increasing the overall accuracy of the estimations. • Optimize the timing of table grapes harvest and storage based on data from multiple sources. They will also receive production data and assure them whether the production covers the specific standards that are set by the supermarkets. A powerful system that allows growers to efficiently plan the harvest and storage of their table grapes will greatly benefit them and increase the overall production quality of the sector. • Assess soil, weather and vegetation data as recorded from the pilots by unlocking the value of big data for irrigation and fertilization in agriculture and promote potential real-time decision support tools. • Develop tools that automatically generate a number of options for delineating management zones, taking into consideration two or more variables and optimizing the process.
Pilot Modifications	✓During the first year of piloting activities no measurements with the Laser Scanner LMS100 LiDAR sensor were performed. ✓Addition of drone imagery during the second crop season: Two (2) Phantom 4 Pro drones (Dà-Jiāng Innovations, Shenzhen, Guangdong, China) (Figure 10) equipped with a multispectral Parrot Sequoia+ camera (Parrot SA, Paris, France) and a Flir Vue Pro thermal camera (FLIR Systems Inc., Wilsonville, Oregon, United States) to collect aerial imagery data and generate Vegetation Indices and Irrigation/Water Activity maps respectively.

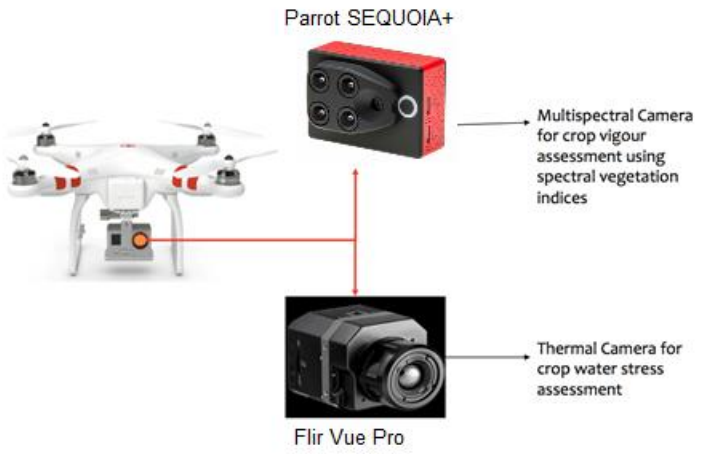

Figure 10: Phantom 4 Pro drone w/ multispectral and thermal camera

2.1.2 Quantitative Evaluation Against KPIs

Domain Specific KPIs

AUA has created the list of domain specific KPIs for the Table and Wine Grapes Pilot and has defined their baseline values, which are presented in the following table.

Table 3: Table and Wine Grapes Pilot Domain Specific KPIs Catalogue

Variable	Definition	Units	2018 Baseline	2019	2020
Total soluble solids	The minimum sugar content of the must at harvest for the production of red dry wine	Brix	20		
Total titratable acidity	The minimum total titratable concentration of the must at harvest for the production of red dry wine	g tartaric acid/L of must	3.5		
Total anthocyanin content	Minimum total anthocyanin content for the production of red dry wine	mg malvidin/g of fresh skin	3.00		
Selective harvesting	The purpose is to achieve different harvest dates depending on the grape quality characters per plot/cell instead of harvesting the entire vineyard on the same date	Number of harvesting dates per plots per vineyard	1		

Technological KPIs

Additionally, in order to perform a complete quantitative evaluation for the Table and Wine Grapes Pilot, a Technological KPIs list along with baseline values have been defined by AUA.

Table 4: Table and Wine Grapes Pilot Technological KPIs Catalogue

Variable	Definition	Units	2018 Baseline	2019	2020
Focusing Big Data					
Volume	Variation in raw data volume – Proximal sensor data	MB	26		
Volume	Variation in raw data volume – Weather data	MB	5		
Volume	Variation in raw data volume – Earth observation data	GB	300		
Volume	Variation in raw data volume – Drone Imagery	GB	*		
Volume	Variation in raw data volume – Yield and Quality	MB	2		
Variety in Data Source Types	Number of different data source types	Data sources	15		
Variety in Data	Number of different types of data (in different resolutions)	Datasets	10		
Velocity	Speed of data generated – Proximal sensor data	MB/crop season	26		
Velocity	Speed of data generated – Weather data	MB/year	5		
Velocity	Speed of data generated – Earth observation data	GB/crop season	125		
Velocity	Speed of data generated – Drone Imagery	GB/crop season	-		
Velocity	Speed of data generated – Yield and Quality	MB/crop season	2		

*Drone imagery is expected to add up to another 200GB from 2019, when the data collection starts.

2.2 WINE MAKING PILOT INDIVIDUAL EVALUATION (INRA)

2.2.1 Qualitative Evaluation Summary

In order to report the Wine Making Pilot's progress, INRA has completed the following table with the necessary information, regarding the current status of development, the successfulness of implementation, its impact and potential modifications.

INRA	Wine Making Pilot Qualitative Evaluation Summary
Pilot Evaluation Summary	*Specific Objectives* Our specific objectives are related to the following topics: - to have a device to improve data quality (correction) and make FAIR data - to have a better understanding of 'How data from the field can affect the wine quality?' and 'How vine water status can affect the wine quality?' - to discover knowledge in order to design new viticulture / winemaking systems The Wine Making pilot is linked to the Prediction Use Cases and more specifically to the Scenario Hypothesis B3-2. Crop Quality Prediction for Optimizing Wine Making. *Achievements/Results* Climatic database • Results: precipitation, water height, evapotranspiration, humidity, insolation, wind. French Network of Grapevine Repositories • Results: morphological description, genetic profile, accessions, location, aptitudes, details. PHIS • Results: leaf area, plant height, biomass, plant width and images/ plant trait extractions. SilexVitioeno • Results: soil characteristics, vineyards, plots, sub-plots, grapevines, different treatments of experiment, cropping management, grape/ berry properties, yield, observations regarding the grape characteristics during the growing season. Monitoring of winemaking operations • Different stages of the winemaking process, duration of winemaking steps, dates etc. ALFIS • Alcoholic fermentation kinetics. Laboratory analysis • Chemical and physical analysis of must and wine: pH, alcohol, total acidity, volatile acidity, residual sugars, etc. Sensory analysis • Wine flavour profiles, scores regarding aromas presence. *Problems/Challenges*

	Data "linkability": The main challenge is related to traceability from field to wine. Indeed, diverse persons are involved in data collection at different scales and it is not always easy to make links between all operations (especially between winemaking and field data). Moreover, data are not always in our information system. Our challenge is to make connections between all data to be able to have a transversal approach from the vineyard to the final product.
Pilot Plan Progress	Data are coming from diverse sources. The main data sources identified are: Climatic database - *Climatic data available from 1989 to 2018.* - *Partners involved: INRA.* - *The goal is to connect environmental or climatic data to information related to the winemaking process or vines.* French Network of Grapevine Repositories - *The French Network of Grapevine Repositories (RFCV) includes 36 regional partners involved in the preservation of grapevine genetic resources and selection. More than 180 repositories are distributed to the grape producing regions in France.* - *The network partners are the regional stakeholders involved in grapevine conservation and selection in France. They play a pivotal role in the conservation and valorisation of our viticultural heritage.* PHIS - *Three experiments have been achieved on the phenotypic platform in 2012, 2013 and 2014 (the PhenoArch phenotyping platform).* - *Partners involved: the researches at the LEPSE unit (INRA).* - *Measurement goal: context of water scarcity and global climate changes, the researches at the LEPSE aim at analysing and modelling the responses of plants to drought and high temperatures as well as their genetic variability at the intra- and inter-specific levels.* SilexVitioeno - *SilexVitioeno Pech Rouge is an information system filled by people from Pech Rouge and about plots, vine stock, and berries all along grapevine lifecycle.* - *Partners involved: INRA, experimental unit of PechRouge.* - *The measurement goal is to characterize vines and their environment to optimize crop management.* Monitoring of winemaking operations - *The monitoring of winemaking operations is obtained with a form filled after harvest to record all information linked with the must studied.* - *Partners involved: Workers from Pech Rouge.* - *The goal of this monitoring is the traceability to compare and follow every steps of the winemaking process.* ALFIS - *Alfis is a SQL database started in 2004; this is an automatic monitoring connected to an information system.* - *Partners involved: INRA unit called Science for Oenology ("SPO").* - *The goal is to make fermentations under controlled conditions with an online acquisition of fermentation kinetics to understand must and*

	yeast behaviour and optimise the process of alcoholic fermentation. Laboratory analysis - *Laboratory analysis is done on each sample of must and wine.* - *Partners involved: experimental unit of Pech Rouge and another INRA unit "Science for oenology" SPO.* - *The purpose is to have chemical and physical analysis of must and wine and look at their evolution* Sensory analysis - *Sensory analysis is done by wine experts on some wine samples before and after bottling.* - *Partners involved: Science for Oenology unit.* - *Measurement goal: to obtain wine flavour profiles.* All this data gathered will help us to build hypothesis and create semantic models. All data are recorded but they still need to put in order and adapted to the Big Data Grapes project.
Status of Implementation	*Actors Involved* INRA has successfully collected data throughout the first crop season. More specifically, in Viticulture from Pech Rouge Nicolas Saurin (Team leader) and the Team Viticulture / Quality Grapes. In Winemaking from Pech Rouge Jean-Michel, Alain Samson (Team leader) and the Team Innovative Technology/ Oenology. Agroknow, CNR, KULeuven and Ontotext participated as Tech providers, with Agroknow and Ontotext acting as model providers, CNR on the data analytics, while KULeuven was responsible for data visualisation. Finally, all pilot partners and tech partners provided insights in order to formulate and define the scenario hypothesis. *Methodology* Viticulture: Research topics are mainly related to the agro-climatic and social context of the South of France. It can be summarised as follows: hot climate, dryness, climate change, pH and K+ elevation, grapes and wine acidity diminution, irrigation, varieties selection, vineyard cultural practices, diversification, sensors. The following is an example of the main steps for vineyard experiments: • Phenology: budding, flowering, veraison (observation, visual counting). • Hydric status: water potential (pressure chamber): 1 time a week from June to harvest. • Apex: apex growth: 1 or 2 times a week (from mai/ June until harvest). • Exposed leaf area (calculation with width, height and foliage porosity): once after growth stop (in July, August). • Maturity: sample of 200 berries with berry weight analysis, refractometric indexes (sugar content), total acidity, pH, assimilable nitrogen, and for red wines anthocyanins and total polyphenol index. 3 sampling dates (at 2 or 3 dates and at harvest). • Aroma precursors, depending on the project. • Yield components: bunch number, weight per grapevine at harvest. • Cut wood weight: number of shoots, shoots weight per grapevine in winter (during pruning, after harvest).

Winemaking:
First, information related to harvest is recorded. Then, different operations are monitored during the winemaking process. For example, some observational data are done concerning the different product features and also observational results of some attributes for a particular product stage such as grape, initial must, must after alcoholic fermentation, and finished wine (such as sensorial analysis achieved by judges).
Here, Figure 11, is an example of the different steps of winemaking done at Pech Rouge for red wines.

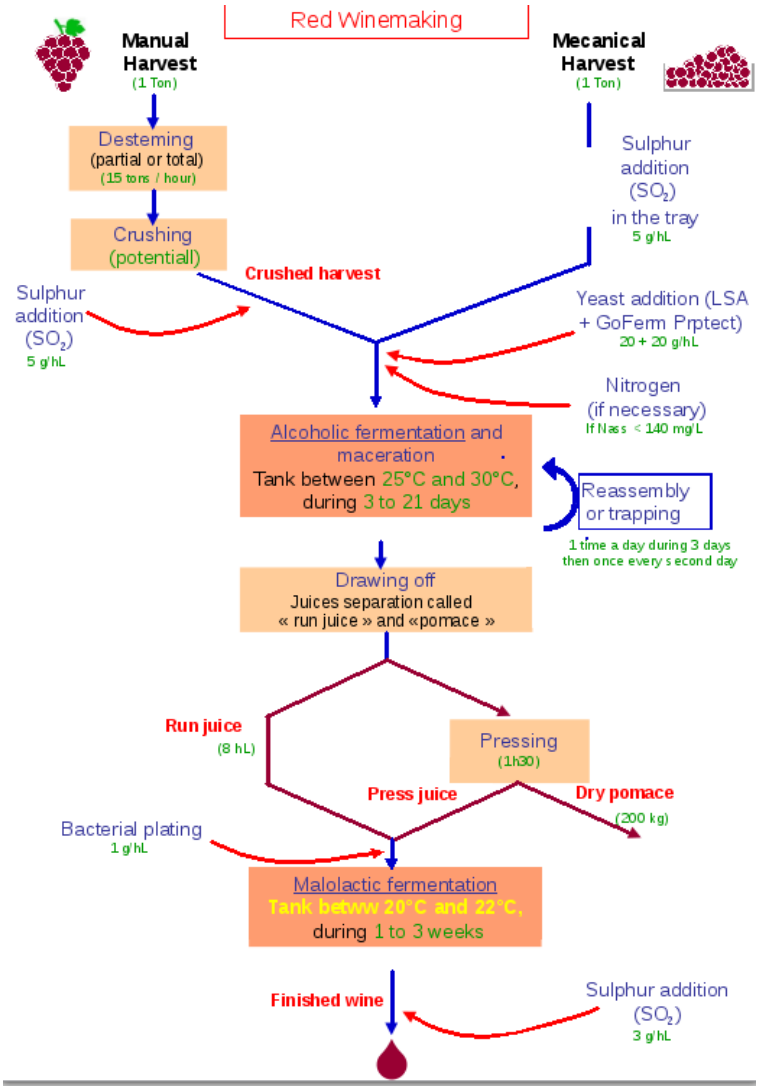

Figure 11: Steps for winemaking taking place in Pech Rouge

Deployed Components
Viticulture:
The land field of Pech Rouge includes a total area of 170 ha of land planted with 38 hectares of vines, distributed in three areas. The INRA Pech Rouge Experimental Unit also contains analytical laboratories, technological tools.

Winemaking:
The land field of Pech Rouge includes a total area of 170 ha of land planted with 38 hectares of vines, distributed in three areas. The INRA Pech Rouge

Experimental Unit also contains analytical laboratories, technological tools and finally a Sensory Analysis Laboratory, which enables the tasting of different wines.

- ✓ Laboratory equipment
- ✓ Expert analysis
- ✓ Experimental technological facilities: A technological facility dedicated to grape extraction, grape processing and winemaking. A technological facility for delayed fermentations allowing alcoholic fermentations under controlled conditions with on-line acquisition of fermentation kinetics.
- ✓ Winery
- ✓ Packaging facility
- ✓ Barrel cellar
- ✓ Wine bar
- ✓ Laboratory analysis
- ✓ Sensorial analysis

Gathered Data and Formats

Viticulture:

- ✓ The gathered data using this measurement technique: climatic data, genetic data, and information about vineyards, plots, sub-plots, grapevines, and different treatments of experiment.
- ✓ Defined frequency of data collection: every years but variables assessed depend on projects goals.
- ✓ Database: SilexVitiOeno, CLIMATIK, PHIS, French Network of Grapevine Repositories, format: xls/ csv, size: MB.

Winemaking:

- ✓ The gathered data using this measurement techniques: different operations during the winemaking process: recorded observational data concerning the different product features. Observational results of some attributes for a particular product stage such as grape, initial must, must after alcoholic fermentation, and finished wine.
- ✓ Defined frequency of data collection: every years but variables assessed depend on projects goals.
- ✓ Database: files coming from Pech Rouge experimenters/ laboratories science for oenology unit, Alfis.
- ✓ Format: xls, hand-written, size: MB.

Table 5: Wine Making Pilot Data and Datasets

Name	DataSet Description	Priority	Provenance	Data Type Format	Data size
Genetic Data	Genetic profile, Morphological description, origin, etc.	Essential	French Network of Grapevine Repositories (Database of the collections)	csv or api	MB
Soil characteristics	Texture, pH etc.	Essential	Field measurement	xls	MB
Plot management	Treatments/fertilizing (when, what, how much), ground handling, or tasks related to the culture management, pest control, water status, yield etc	Essential	Field measurement	Pdf,doc, xls	MB
Climatic data	Rainfall, temperature, radiation etc.	Essential	Field measurement	xls	MB
Grape and berry mechanical and chemical properties	Anthocyanin content, weight, length, width, density etc.	Essential	Field measurement	xls	MB
Qualitative and quantitative characteristics of must	Sugar content, alcohol, pH etc.	Essential	Laboratory equipment		MB
Winemaking activities	Bioconversion of sugar into ethanol and CO_2, Monitoring of alcoholic fermentation and sugar content, yeast characteristics etc.	Essential	Laboratory equipment	xls, pdf	MB
Sensory analysis	Expert panel of tasters' sensory analysis (wine bitterness, astringency, phenol content, aroma etc.)	Essential	Expert analysis	xls	MB
Wine commercial information	Number of bottles produced, number of bottles sold	Additional Data	Selling point	xls	MB

Examples:

- Data collection Phenological stages (Budding, flowering and veraison).
  - Sampling: 15 grapevine per repetition.
  - Frequency: 2 dates per stages.
- Maturity (weight of 200 berries, sugar, total acidity, pH, assimilable nitrogen, anthocyanins, polyphenols).
  - Sampling: 200 berries per treatment + 200 berries for anthocyanins and polyphenols.
  - Frequency: 14 days before harvest, 7 days before harvest and at harvest.
- Aroma precursors

	○ Sampling: 20 bunch of grape/treatment. • Yield ○ Sampling: 15 grapevines / repetition at harvest. • Leaf Area Index (LAI) ○ Sampling: 4 transepts per treatment.
Impact	The main impact of the Wine Making piloting sessions is related to data gathering at its experimental site. Links have been identified but also lack of data connection. A direct impact will be noticeable on our way of working, storing and managing data.
Pilot Modifications	No need for modifications has appeared yet for the Winemaking Pilot.

2.2.2 Quantitative Evaluation Against KPIs

Domain Specific KPIs

INRA has generated the list of domain specific KPIs for the Wine Making Pilot and has defined their baseline values, which are presented in the following table.

Table 6: Wine Making Pilot Domain Specific KPIs Catalogue

Variable	Definition	Units	2018 Baseline	2019	2020
Product yield per plot	Kg per plot of grapes harvested	Kg/plot	1637		
Product yield	Kg per ha of grapes harvested	Kg/ha	4107		
Irrigation	m^3 of water per plot to irrigate vineyards	m^3 /plot	187		
Wine volume per kilogram harvested Red wines	Red wines It corresponds to the volume of wine produced per kilogram of harvested grape	L/kg	0.37		
Wine volume per kilogram harvested White and rosé wines	White and rosé wines. It corresponds to the volume of wine produced per kilogram of harvested grape	L/kg	0.19		
Marc quantity extracted for red wines after fermentation	Red wines. 100 x marc weight / harvest weight	%	13.92		
Marc quantity extracted for white	White and rosé wines. 100 x marc weight / harvest	%	28.39		

and rosé wines after fermentation	weight				
Residual sugar content in wine – red wines	Sugar content in wine. We have to check that the value of residual sugar is below 2 g/L. Red wines Calculation: 100 * nber of conformed wine/total wine number	%	100		
Residual sugar content in wine – white and rosé wines	Sugar content in wine. We have to check that the value of residual sugar is below 2 g/L. White and rosé wines Calculation: 100 * nber of conformed wine / total wine number	%	95		
Volatile acidity after alcoholic fermentation for red wines	It must be 0,10 <x < 0,98. Red wines Calculation: 100 * nber of conformed wine/total wine number	g/L H_2SO_4	100		
Volatile acidity after alcoholic fermentation for white and rosé wines	It must be 0,10 < x < 0,88 White and rosé wines Calculation: 100 * nber of conformed wine/total wine number	%	100		
Malic acid concentration in red wines after malolactic fermentation	It must be < 0,1 g/L (to be sure that the fermentation process is finished). Red wines. Calculation: 100 * nber of conformed wine / total wine number	%	57		
Volatile acidity (post malolactic fermentation) for red wines	Volatile acidity for red wines after malolactic fermentation. This parameter is regularly checked	g/L H_2SO_4	0,35		
Color intensity (darkness) for red wines – visual analysis Before bottling	The purpose is to have a dark color for red wines. The ratio calculated corresponds to the number of judges who found the wine dark / total number of judges who are able to detect the characteristic	0 < Ratio < 1	0.71		

Color intensity (clearness) for white and rosé wines – visual analysis Before bottling	The purpose is to have a clear wine for white and rosé wines. The ratio corresponds to the number of judges who found the wine clear / total number of judges who are able to detect the characteristic	0 < Ratio < 1	0.70		
Fruity flavor Before bottling	The fruity flavor is well desired for all wine types. The ratio corresponds to the number of judges who detected this aroma / total number of judges	Ratio	0.59		

Technological KPIs

Additionally, in order to perform a complete quantitative evaluation for the Wine Making Pilot, a Technological KPIs list along with baseline values have been defined by INRA.

Table 7: Wine Making Pilot Technological KPIs Catalogue

Variable	Definition	Units	2018 Baseline First year of the project	2019	2020
Focusing Big Data					
Volume	Variation in raw data volume – Plot Management	MB	9.4		
Volume	Variation in raw data volume – Climatic data	KB	80		
Volume	Variation in raw data volume – Grape and berry mechanical and chemical properties	KB	93		
Volume	Variation in raw data volume – Qualitative and quantitative characteristics of must and wine	KB	380		
Volume	Variation in raw data volume – Winemaking activities	KB	102		

Volume	Variation in raw data volume – Sensory Analysis	KB	530		
Volume	Variation in raw data volume – Satellite Data	GB	47		
Velocity	Speed of data generated during harvesting period	MB/ harvesting period, 4 months	14.9		
Velocity	Speed of data generated – Satellite data	GB/month	2.63 S2 1.25 L8		
Variety in Data Source Types	Number of different data source types	Data sources	18		
Variety in Data	Number of different types of data (in different resolutions)	Datasets	9		
Data transformation	Number of rdf triplets, from raw data	Number	0		
Data linked	% of data linked, data connection – dataset linked divided by the total number of dataset	%	11%		
Level of FAIR-ness	Fair data assessment tool especially for winemaking activities	RDA SHARK evaluation (David et al., 2019)	16/18 Never 2/18 If Mandatory 0/18 Sometimes 0/18 Always		
Big Data Process Metrics					
Data Normalization (Homogenization)	Steps number needed for data to be available for analysis and processing Winemaking activities	Number	7		
Existing ontology AFEO enrichment	Number of classes	Number	68		
Existing ontology AFEO enrichment	Number of properties	Number	8		

It is important to underline that these variables make sense if they are well described in ontologies using semantic web to be able to do machine learning on them.

2.3 FARM MANAGEMENT PILOT INDIVIDUAL EVALUATION (ABACO-GEOCLEDIAN)

2.3.1 Qualitative Evaluation Summary

In order to report the Farm Management Pilot's progress, ABACO, with the help of GEOCLEDIAN, has completed the following table with the necessary information, regarding the current status of development, the successfulness of implementation, its impact and potential modifications.

ABACO	Farm Management Pilot Qualitative Evaluation Summary
Pilot Evaluation Summary	*Specific Objectives* The ABACO and GEOCLEDIAN Farm Management Pilot is focused on developing a unique system that satisfies these needs: • Farm Management with all the functionalities to support the farmer in his day by day activities and gather data from the field • Hosting data from different sources with proper tools and functionalities for comparisons and easy data management • Data exchange. A "day by day" data producer, to feed the generated data into the other BDG components and make use of the incoming information from the other BDG components. • Data visualization. The data relevant for the farmer should be displayed in a way that provides an added value and new insights to the farmer for his activities. The ultimate goal for this pilot is to provide precise monitoring of the vegetation – atmosphere – soil system and manage best practice activities, through the improvements of SITI4farmer, with the support of specific DSS tools and precision farming features. The Farm Management pilot is linked to the Data Anomaly Detection & Classification and the Farm Management Use Cases and more specifically to three Scenario Hypothesis, namely A. Earth Observation Data Anomaly Detection & Classification, C1. Optimisation of Farm Practices in the Vineyard, C2. Management Zones Delineation for Vineyards. *Achievements/Results* The pilot has compiled best practice activities and has started to read properly information provided by sensors installed in fields. GEOCLEDIAN has acquired and processed Copernicus Sentinel-2 and USGS Landsat-8 images for all sites during the first 18 months of the pilot run time and made them available to all partners. GEOCLEDIAN's Processing Platform provides the service ag|knowledge that allows the automatic crop monitoring for fields with advanced products based on all spectral bands of these satellites. The processing platform has been substantially improved and the development into a Big Data Processing platform has started. New products are available. *Problems/Challenges* Challenges have been encountered on training the DSS and precision farming tools.

Pilot Plan Progress

Abaco's main Tasks and Operations, that have been performed to achieve the goals of the pilot are:

- Formal Engagement of the winery companies
- Collecting information of fields, terrain, product quality
- Analysis for the sensors set up on the right spot and configuration
- Setup of SITI4farmer for the 2 companies
- Development of the system sensors station interfacing
- Measurements and monitoring of field activities

Piloting activating (in particular the last one) will be part of the entire project and is considered as an activity in continuous improvement, where data and results will be presented to the other partners regularly.

Macro Activity	Time
GEOCLEDIAN: Data acquisition, processing & provision	Q3-Q4.2018
Winery Company formal engagement	Q3.2018
Abaco's Hardware & Software supplying	Q3.2018
Deploying of SITI4farmer	Q4.2018
Abaco's Development & Configuration for sensors integration	Q4.2018
Training of user on the system	Q4.2018
GEOCLEDIAN: Integration of new data sources (new vegetation indexes)	Q1-Q2.2019
Field Measurements & monitoring	Q1.2019 to Q4.2020
GEOCLEDIAN: Development of Management Zones & data anomaly detection	Q1-Q4.2019
GEOCLEDIAN: Improvement of vineyard specific products with feedback from users	Q1-Q4.2020

GEOCLEDIAN has acquired and processed the described satellite data of all sites up to M18. Visible images and Vegetation Index Maps can be produced in our Processing platform and the data is available to all project partners in near real-time. The satellite data processing platform has been substantially improved.

The pilot aims at providing a precise monitoring of the vegetation – atmosphere – soil system and manage best practice activities with the support of specific DSS tools and precision farming features. It has compiled best practice activities and started to read properly information provided by sensors installed in fields.

Status of Implementation

Actors Involved

Farm Management Pilot (ABACO-GEOCLEDIAN) and farm management data collected from the test sites in Italy, with GEOCLEDIAN providing the satellite and SVIs datasets and KU Leuven for the visualisation of the dashboard.

Methodology

Abaco is going to release a version of its product; SITI4farmer is ready to be used in the field by 2 different winemakers, owners of the test sites in

Italy and all the project partners.
SITI4farmer is able to (with just an internet connection trough browser):

- Prepare the graphical crop plan
- Manage farming practices and phenology phases
- Analyse indices and dashboards to support decisions (agro-meteorology and vegetation)
- Keeping farm data organized and accessible
- Recording field data with the SITI4l and app
- Printing and export data

Furthermore, it's able to integrate weather data and services from different sources, also, it can use open databases and local land registries made available by everyone that has an exposed service.

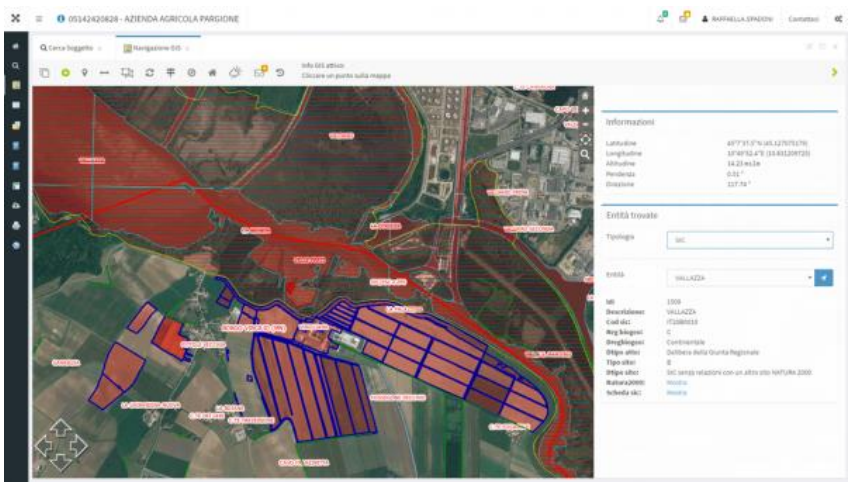

Figure 12: SITI4farmer screen view

GEOCLEDIAN has acquired and processed the described satellite data of all sites up to M18. Visible images and Vegetation Index Maps can be produced in our Processing platform and the data is available to all project partners in near real-time. The satellite data processing platform has been substantially improved. A series of developments have been implemented and deployed successfully to improve data download & processing, performance monitoring, scalability and data visualization and to enable the delivery of the new data products and vegetation indexes that were developed for Abaco. In the frame of this project the field monitoring service Ag|knowledge is further developed from a basic image delivery service into an Agricultural Big Data Processing Platform that allows the scalable production, provision & analysis of large-scale data. Developments on Management Zones & data anomaly detection have been started.

Deployed Components
In order to make full and comprehensive measurements in the fields, automatizing as much as possible, Abaco has acquired and integrated within the system 2 sensors stations dedicated to this purpose. Sensor Stations were purchased directly from Abaco and installed within two farms.

Sensors and weather station are set to working via radio with a central server, and transmit data directly to SITI4farmer. They are equipped with:

- Modem, aerial, battery, solar panel;
- Rain Gauge Module
- Temperature and humidity sensors
- Wind direction system
- Wind speed measurement sensors
- Solar Radiation sensor
- Single Leaf Temperature Sensor
- Infrared Temperature Module
- IR Temperature sensor
- Instruments Leaf Wetness Sensor Module with 5 meter of cable
- Drill & Drop Sensor (Temperature and soil moisture sensors)

Figure 13: Sensor & Weather Station

Figure 14: Rain Gauge Module

Concerning GEOCLEDIAN, these components have been deployed successfully in the frame of the improved satellite data processing platform Ag|knowledge:

- Extended data download and processing components
- New processing performance monitoring tools
- New vegetation indexes and data products component
- New API endpoints for data delivery

- New data visualization tools for data review and analysis

Site Description
The approach consist on the involvement of 2 wineries, making them an active part of the project, collecting data from the field, in automatic and manual manners, and therefore contribute to the results.

Company Name: CASATO PRIME DONNE CIRCA
Address: Località Casato – Montalcino, Tuscany, IT
GPS Coordinates : 43.088196° N 11.464319° E
Internet Site: www.cinellicolombini.it

12 HA of Vineyards of Brunello of Montalcino

Figure 15: 12 HA of wineyards of Brunello of Montalcino

Company Name: CANTINA IL PALAZZO
Address: Loc. Antria, Arezzo, Tuscany, IT
GPS Coordinates: 43.502773, 11.904402
Internet Site: www.tenutailpalazzo.it

35 HA of Vineyards of CHIANTI D.O.C.

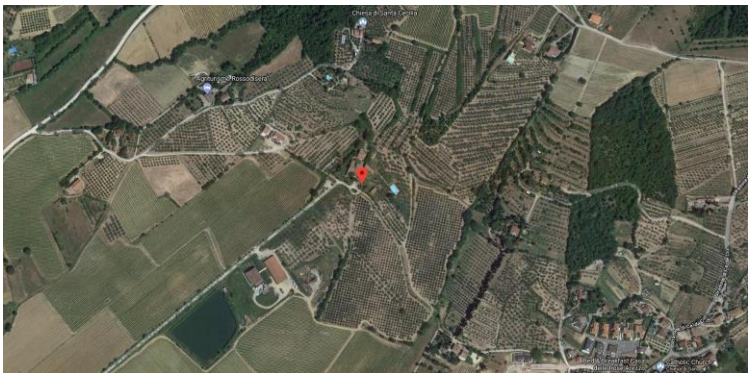

Figure 16: 35 HA of Wineyards of CHIANTI D.O.C.

Gathered Data and Formats

- Several Satellite vegetation indices based on Sentinel 2 and Landsat 8 data (NDVI, NDRE1, NDRE2, NDWI, EVI2, SAVI, CIRE)
- Pessl Instruments Weather and Soil sensors: Air temperature and humidity, wind speed and direction, precipitation, soil

temperature and water content at different deepness (from 0 to 50 cm), leaf and bunches IR Temperature, Leaf moisture.

- Weather and soil data are monitored every hour, Satellite data depends on data availability in function of coverage with an average of 6 days frequency,
- Rest-API JSON format for both Satellite and Weather-Soil station

Table 8: Farm Management Pilot Data and Datasets

Name	DataSet Description	Priority	Provenance	Data Type Format	Data size
Sentinel-2	Sentinel-2A/B MSI visible & NIR bands, NDVI time series & advanced products	Essential	Copernicus EO Programme, ESA	JSON, GEOTIFF, PNG	150 GB/year*site
Landsat-8	Landsat-8 OLI visible & NIR bands & advanced products	Essential	USGS, NASA	JSON, GEOTIFF, PNG	6 GB/year*site
VHR data	VHR satellite data, e.g. TripleSat VHR optical bands	Additional	TBD, e.g. 21AT	TBD, e.g. GEOTIFF	TBD
Chemical and physical info on grapes	Antocyanins, Ph, Brix values during maturation	Additional	Excel table file	XLS	TBD
Day by Day Activities in term of treatments, fertilization, field operation	Diary where farmer or operators can record and /or plan all the activities on their fields	Essential	SITI4farmer	Text file	TBD
Plot and Fields information georeferenced	Information of Plots position, shaping, cultures, type of seed, dates, and everything related on the culture and the farm itself (form official and not official point of view)	Essential	Form the field through SITI4farmer	Test files	TBD
Relative Humidity	Relative humidity (RH) is the ratio of the partial pressure of water vapor to the equilibrium vapor pressure of water at a given temperature	Essential	Field Sensors	Decimal Data	TBD
Air Temperature		Essential	Field Sensors	Decimal Data	TBD
Global Solar Radiation	It's the power per unit area received from the Sun in the form of electromagnetic radiation in the wavelength range of the measuring instrument	Additional	Field Sensors	Decimal Data	TBD
Wind Speed and Direction		Essential	Field Sensors	Decimal Data	TBD
Soil Temperature		Additional	Field Sensors	Decimal Data	TBD
Soil Moisture	Measurement of the water in the large and intermediate size pores that can move about in the soil and be easily used by plants	Essential	Field Sensors	Decimal Data	TBD
Precipitation	Rainfall measurements	Essential	Field Sensors	Decimal Data	TBD
Infrared Surface Temperature	Temperature Surface calculated with infrared measurements	Essential	Field Sensors	Decimal Data	TBD

Impact	High impact on DSS provided to the pilots

Pilot Modifications	We decided to not purchase VHR data for the time being as the open data used currently yields more information about vineyards than expected at project start. We first want to concentrate on the exploitation of these information like water stress and phenological parameter derivation from dense open data time series.

2.3.2 Quantitative Evaluation Against KPIs

Domain Specific KPIs

ABACO has acquired the list of domain specific KPIs and their baseline values for the Farm Management Pilot from the IL Palazzo test site in Italy, which are presented in the following table.

Table 9: Farm Management Pilot Domain Specific KPIs

Variable	Definition	Units	2018 Baseline	2019	2020
Harvested Area	ha of harvested area	ha	35	-	
Product Yield	Kg per ha of grapes harvested, wine produced, raisins produced	Kg/ha	100	-	
Grape Product Quality			High	-	
Production Costs	Costs in Euros per year	Euros/year	6800	-	
Organic Fertilizer Use	Kg fertilizer used per kg grapes harvested per year	Kg/kg y	600	-	
Organic Pesticides Use	Kg pesticides used per kg grapes harvested per year	Kg/kg y	0	-	
Irrigation Cost	Euros per ha	Euros/ha	0	-	
Fertilization Cost	Euros per kg	Euros/kg	0	-	
Pesticides Cost	Euros per kg	Euros/kg	720	-	
Labour Cost	Euros per hour	Euros/hr	2500	-	

Technological KPIs

Additionally, in order to perform a complete quantitative evaluation for the Farm Management Pilot, a Technological KPIs list along with baseline values have been defined by ABACO and GEOCLEDIAN.

Table 10: Farm Management Pilot Technological KPIs Catalogue

Variable	Definition	Units	2018 Baseline	2019 (Up to M18)	2020
Focusing Big Data					
Volume	Variation in raw data volume – Sentinel2	GB	130.50	57.00	
Volume	Variation in raw data volume – Landsat8	GB	54.00	20.00	
Volume	Variation in raw data volume – Pessl Instrumens	MB	2.5	3.3	
Variety in data	Sentinel 2	Number of scenes	174	76	
Variety in data	Landsat 8	Number of scenes	54	20	
Variety in data	Pessl Instrumens	Hours	24h * 90 days	24h * 365 days	
Variety in Data Source Types	Data sources (Sentinel 2, Landsat 8, Pessl Instruments)	-	3	3	
Variety inter Data	All variables measured (Satellite vegetation Indices, Soil data , Weather data, Canopy data)	Datasets	47	47+7(new satellite indices)	
Velocity	Speed of data generated – Sentinel 2	GB / month	10.88	4.75	
Velocity	Speed of data generated – Landsat 8	GB / month	4.50	1.67	
Velocity	Speed of data generated – Pessl Instrument	MB / month	0.25	0.3	

2.4 NATURAL COSMETICS PILOT INDIVIDUAL EVALUATION (SYMBEEOSIS)

2.4.1 Qualitative Evaluation Summary

In order to report the Natural Cosmetics Pilot's progress, SYMBEEOSIS has completed the following table with the necessary information, regarding the current status of development, the successfulness of implementation, its impact and potential modifications.

SYMBEEOSIS	Natural Cosmetics Pilot Qualitative Evaluation Summary
Pilot Evaluation Summary	*Specific Objectives* There is a need in extracting the most out of pharmaceutical plants for both economic and environmental reasons. A real challenge is to add high value to by-products. Wine making produces a lot of by-products that may have a significant biological value if there are adequate data concerning farm management. These data can lead to decisions concerning the processing of by-products in order to produce high added value active ingredients for cosmetics and food supplements. The scenario hypothesis presumes that precision farming and control of parameters linked to the quality of wine may provide by-products of superior quality. In particular, the pilot intends to gather samples of vineyard by-products across the Greek territory and more specifically vine leaves of two different grape varieties (Agiorgitiko and Mandilaria) and test their phytochemical profile and biological value after extraction. The Natural Cosmetics pilot is linked to the Prediction Use Cases and more specifically to B2. Predicting Biological Efficacy. *Achievements/Results* The preparation (i.e. maceration and ultrasound assisted extraction) of vine leaf extracts and testing of their biological efficacy for each sample took place at the laboratory of the collaborative to Symbeeosis Company APIVITA S.A. – Natural Cosmetics, located in Industrial Park of Markopoulo Mesogaias in Greece. At the laboratory conducted extractions under the two different methods and the following measurements of biological activity (BA): pH, RI, TPC, TFC, Total Microbial Count, Yeasts & Moulds, DPPH & ABTS assay. The measurement of toxicity (MTT assay) and of gene expression (target SIRT1 mRNA transcripts) on skin cells are still pending and will be contacted together with the second-year samples analyses. The collected data from the natural cosmetics pilot provided the necessary information for the evaluation of the quality of each sample, linked with the special characteristics of the vineyard of origin (SVIs data). Finally, the correlation analyses of BA parameters with the vineyards characteristics pointed out which information should be taken into account for building the models that will support the Decision Support System (DSS). *Problems/Challenges* The BA parameters data intended to be correlated with weather data but due to limited input it was preferred to do the correlation with satellite data for vegetation indices (SVIs). Another shortcoming was the relatively low by-products samples collected for BA analysis, which is scheduled to be overcame by the next seasons sampling.
Pilot Plan Progress	• Start month: M1, duration: 18M • Partners involved: SYMBEEOSIS (BA parameters Data), GEOCLEDIAN (SVIs datasets), CNR (data correlation analysis), Ontotext (data modelling), Agroknow (data management, transformation, uploading to the software stack), AUA (WP leader) Third parties involved: APIVITA S.A. (laboratory analyses of samples) • Pilot's objective: The collected data from the natural cosmetics pilot will

provide the necessary information for the evaluation of the quality of each sample, linked with the special characteristics of the vineyard of origin. The goal is to face the challenge: "how data from the field can be linked to the biological efficacy of final products - an application on wine making by-products".

- Progress of activities: The collected data from the natural cosmetics pilot provided the necessary information for the evaluation of the quality of each sample, linked with the special characteristics of the vineyard of origin (SVIs data). The correlation analyses of BA parameters with the vineyard's characteristics pointed out which information should be taken into account for building the models that will support the DSS.

Status of Implementation

Actors Involved

Natural Cosmetics Pilot (SYMBEEOSIS) and the BA data collected from samples all around Greece, with GEOCLEDIAN providing the SVIs datasets, CNR for the data correlation analysis, Ontotext for data modeling, Agroknow for data management and their appropriate transformation for uploading to the software stack, and KU Leuven for the visualisation of the dashboard.

Methodology

A. Sample Collection

For the first year of the project, sixteen regions of the Greek territory have been chosen for sample collection, i.e. dried vine leaves of two different grape varieties (Agiorgitiko and Mandilaria). Also, samples of both grape varieties from the vineyard of Hellenic Agricultural Organization "DIMITRA" located in Attica will be tested. The dispersion and origin of the samples is shown in the following map, where the samples of Agiorgitiko are pictured in green and the samples of Mandilaria in red.

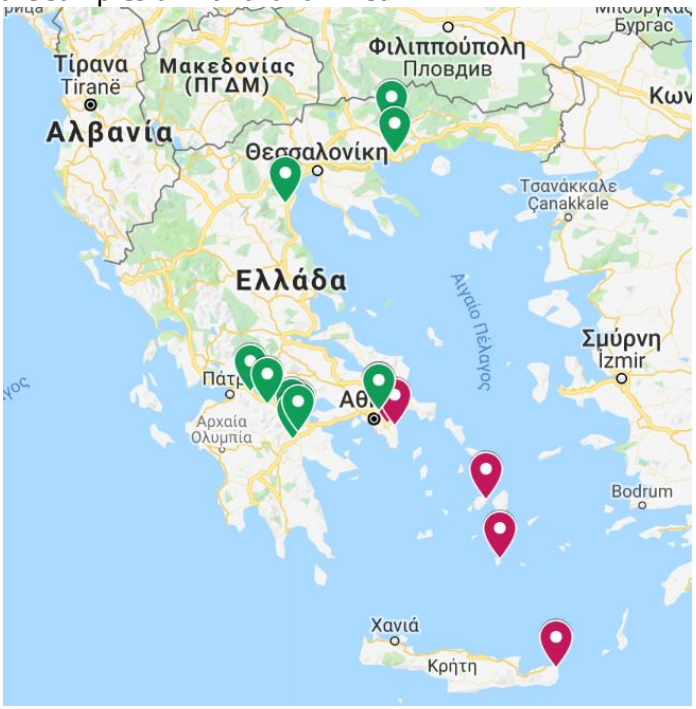

Figure 17: Dispersion of samples across the Greek territory

In the following table there is a list of the vineyards chosen for sample

collection and their location.

Table 11: Vineyards chosen for sample collection

	Vineyard	Grape Variety	Region	City
1	Semeli Wines	Agiorgitiko	Peloponnese	Nemea
2	Pavlidis Estate	Agiorgitiko	Northern Greece	Drama
3	RIRA Vineyards	Agiorgitiko	Peloponnese	Aigio
4	Vassaltis Vineyards	Mandilaria	Aegean	Santorini
5	Strofilia Estate Winery	Agiorgitiko	Peloponnese	Stimfalia
6	Papagiannoulis Winery	Agiorgitiko	Northern Greece	Katerini
7	Tetramythos Wines	Agiorgitiko	Peloponnese	Ano Diakopto
8	Skouras Domaine	Agiorgitiko	Peloponnese	Argos
9	Moraitis Winery	Mandilaria	Aegean	Paros
10	Toplou Winery	Mandilaria	Crete	Sitia
11	Aoton Winery	Mandilaria	Attica	Peania
12	Biblia Chora Estate	Agiorgitiko	Northern Greece	Kavala
13	Papagiannakos Domaine	Mandilaria	Attica	Markopoulo
14	Hellenic Agricultural Organization "DIMITRA"	Mandilaria	Attica	Lykovrisi
15	Hellenic Agricultural Organization "DIMITRA"	Agiorgitiko	Attica	Lykovrisi
16	Agricultural University of Athens	Agiorgitiko	Peloponnese	Nemea

B. Laboratory testing

The preparation of vine leaf extracts and testing of biological efficacy of each sample will take place at the laboratory of collaborating Company APIVITA S.A. – Natural Cosmetics, located in Industrial Park of Markopoulo Mesogaias in Greece.

Figure 18: Collaborating Company's (APIVITA) laboratory

Deployed Components

- Extractions will be conducted using Elma S60H Elmasonic Ultrasonic Bath.

Figure 19: Elma S60H Elmasonic

- The measurement of pH will be conducted with a seven compact pH meter, METTLER-TOLEDO.

Figure 20: pHmeter, METTLER-TOLEDO

- The measurement of refractive index will be conducted with a Digital Refractometer RX-a- series ATAGO

Figure 21: Digital Refractometer ATAGO

- A NUVE Incubator and a Laminar Telstar BO-II-A will be used for the measurement of total microbial count with classic development of micro-organism in petri-dishes.

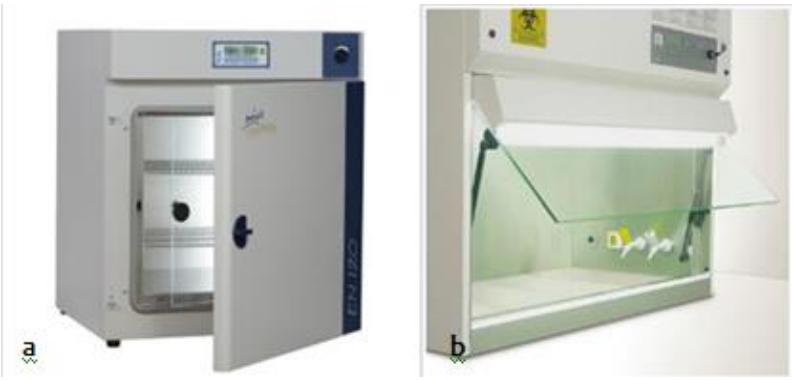

Figure 22: (a) NUVE Incubator, (b) Laminar Telstar BO-II-A

- A Memmert will be used for the measurement of yeasts and moulds with classic development of micro-organism in petri dishes.

Figure 23: Memmert Universal Oven 055 UN/UNm

- A UV 1800 SPECTROPHOTOMETER, SHIMADZU EUROPA will be used for the measurement of antioxidant activity (DPPH & ABTS assay), total phenolic content and total flavonoid content.

Figure 24: UV Spectrophotometer

- A Nanoquant, infinite M200 Pro, TECAN will be used for the measurement of toxicity on skin cells (MTT assay).

Figure 25: Nanoquant, infinite M_{200} Pro

- A CFX connect Real time System, BIO-RAD will be used for the measurement of gene expression on skin cells (Target SIRT1 mRNA transcripts using real time PCR).

Figure 26: CFX connect Real time System

Table 12: Laboratory components; Testing frequency one sampling per year; Analysis duration 3 months

Name	Experimentation Controlled Variables Description
Elma S 60 H Elmasonic Ultrasonic Bath	Ultrasonic bath for extraction
SevenCompact pH meter METTLER-TOLEDO	Bench pH-meter for pH measurement
Digital Refractometer RX-a- series ATAGO	Digital Refractometer for measurement of Refractive Index
NÜVE EN 055 Incubator	Incubation chamber for classic development of bacteria
Memmert Universal Oven 055 UN/UNm	Incubation chamber for classic development of yeast & fungi

Laminar Telstar BO-II-Advance	Class II Biological Safety laminar air flow cabinet for microbiological analyses
Elma S 60 H Elmasonic Ultrasonic Bath	Ultrasonic bath for extraction

Gathered Data and Formats

Table 13: Natural Cosmetics Pilot Data and Datasets

Name	DataSet Description	Priority	Provenance	Data Type Format	Data size
SVIs Data	Sentinel-2A/B MSI visible & NIR bands, NDVI time series	Essential	Copernicus EO Programme ESA	json, geotiff, png	TB
Agiorgitiko Samples UAE (11 samples)	Data on biological efficacy of Agiorgitiko dried vine leave samples, Ultrasound Assisted Extraction	Essential	Laboratory testing	csv, xls	MB
Agiorgitiko Samples MAC (11 samples)	Data on biological efficacy of Agiorgitiko dried vine leave samples, Maceration	Essential	Laboratory testing	csv, xls	MB
Mandilaria Samples UAE (5 samples)	Data on biological efficacy Mandilaria dried vine leave samples, Ultrasound Assisted Extraction	Essential	Laboratory testing	csv, xls	MB
Mandilaria Samples MAC (5 samples)	Data on biological efficacy Mandilaria dried vine leave samples, Maceration	Essential	Laboratory testing	csv, xls	MB
Weather Data	Weather data on the regions selected for sample gathering	Essential	Open source data		

Impact	The collected data from the natural cosmetics pilot provided the necessary information for the evaluation of the quality of each sample, linked with the special characteristics of the vineyard of origin (SVIs data). The correlation analyses of BA parameters with the vineyard's characteristics pointed out which information should be taken into account for building the models that will support the DSS.
Pilot Modifications	The BA parameters data intended to be correlated with weather data but due to limited input it was preferred to do the correlation with satellite data for

vegetation indices (SVIs). Another shortcoming was the relatively low by-products samples collected for BA analysis, which is scheduled to be overcame by the next seasons sampling.

2.4.2 Quantitative Evaluation Against KPIs

Domain Specific KPIs

SYMBEEOSIS has generated the list of domain specific KPIs for the Natural Cosmetics Pilot and has defined their baseline values, which are presented in the following table.

Table 14: Natural Cosmetics Pilot Domain Specific KPIs Catalogue

Variable	Definition	Units	2018 Baseline	2019	2020
Agiorgitiko Samples/ parcel	Number of samples per vineyard (parcel)	Number	1		
Mandilaria Samples/ parcel	Number of samples per vineyard (parcel)	Number	1		
Agiorgitiko Samples	Samples of vine leaves to be analysed	Number	16		
Mandilaria Samples	Samples of vine leaves to be analysed	Number	16		
UAE and MAC efficiency	Percentage of extract from incoming raw material	%	>60		
Extract pH	Ranges for acceptable pH	pH	> 3.5		
Extract RI	Ranges for acceptable % for RI	%	22±4		
Extract TMC	Ranges for acceptable Total Microbial Count	CFU	< 10		
Extract Y&M	Ranges for acceptable Yeats and Moulds counts	CFU	< 10		
Processing Time	Overall time for extraction, required analysis, and assessment of new product	Months	3		

Technological KPIs

Additionally, in order to perform a complete quantitative evaluation for the Natural Cosmetics Pilot, a Technological KPIs list along with baseline values have been defined by SYMBEEOSIS.

Table 15: Natural Cosmetics Pilot Technological KPIs Catalogue

Variable	Definition	Units	2018 Baseline	2019	2020
Focusing Big Data					
SVIs Volume Data	Sentinel-2 A/B MSI visible & NIR bands, NDVI time series	GB	301		
SVIs Volume Data	Landsat 8 A/B MSI visible & NIR bands, NDVI time series	GB	160		
Agiorgitiko and Mandilaria Samples UAE BA parameters Volume	Data on biological efficacy of samples of Agiorgitiko and Mandilaria dried vine leaves, developed with Ultrasound Assisted Extraction	KB	58		
Agiorgitiko and Mandilaria Samples MAC parameters Volume	Data on biological efficacy of samples of Agiorgitiko and Mandilaria dried vine leaves, developed with Maceration	KB	58		
Variety in Data Source Types	BA parameters, SVIs	Data sources	10		
Variety in Data	Number of different types of data (in different resolutions)	Datasets	6		
SVIs Velocity Data	Sentinel-2 A/B MSI visible & NIR bands, NDVI time series	GB/month	25.13		
SVIs Velocity Data	Landsat 8 A/B MSI visible & NIR bands, NDVI time series	GB/month	13.33		
BA Parameters Velocity	Speed of data generated – BA Parameters	KB/season	58		
Big Data Process Metrics					
Data Normalization (Homogenization)	Time needed for data to be available for analysis and processing	Months	3		

3 BIGDATAGRAPES PILOT'S SURVEY

Each BigDataGrapes partner should complete the Pilots' Survey as a part of the Human-centred Evaluation report. Based on the Experimental and Evaluation Protocols and the Use Case Scenarios, as they have been defined and identified, partners should provide feedback on the execution of the application pilots in order to qualitatively evaluate their results (user satisfaction and involvement).

This part describes a user-centered assessment for the BigDataGrapes pilots. It includes the survey which is divided in two levels of assessment, the "Pilot Basis Evaluation" and the "Technological Basis Evaluation", of heterogeneous sets of quantitative and qualitative indicators, metrics used to measure the effectiveness of the pilots. In order for this to be completed, an iterative approach of assessment will be performed according to the proposed three-phase human-centred assessment activities. More specifically, the first distribution of the survey will take place a few months after the completion of the Intermediate phase (M22), which involves the first round of controlled pilot trials and implementations of the first versions of the newly developed BigDataGrapes components. The second distribution will happen at two months before the end of the Summative phase (M34), which entails the validation of the BigDataGrapes components in real-life conditions and with realistic complexity.

Following the concept of gradual extension of functionality, intended audience and assessment of this scheme, the pilots will interact with the community and the pilot evaluators accordingly. Thus, this survey will be distributed to all relevant stakeholders involved in the BigDataGrapes piloting activities. Feedback will be asked from the end-user, including farmers, producers, owners of the reference vineyards, agronomists, lab testers, researchers, stakeholders and the BigDataGrapes pilot and technological partners.

Table 16: BigDataGrapes Pilots' Survey Structure

BigDataGrapes Pilots' Survey	Section (Question Numbers)	Assessment Group	Designated Phase
Pilot Basis Evaluation	**End-User Survey** (Q1 – Q11)	End-user and BDG pilot partner	Surveys will be completed twice: • After the completion of the Intermediate phase (M22) and • Before the end of the Summative phase (M34)
	Vineyard Information and Demographics (Q12 – Q20)	End-user	
Technological Basis Evaluation	**Global Technological Evaluation** (Q21 – Q31)	BDG pilot and tech partner	
	Data Management, Analytics, Modelling and Visualisation (Q32 – Q35)	BDG tech partner	

3.1 PILOTS BASIS EVALUATION

3.1.1 End-user Survey

In order to document how grapevine-powered business problems correspond to (big) data challenges, a set of detailed Use Case specifications, highlighting the data challenges that decision makers face, have been identified under WP2 - Grapevine-Powered Industries Big Data Challenges. The "Pilot Basis Evaluation" includes the "End-user survey" that is directly related to these Use Cases, covering Data Anomaly Detection and Classification, Prediction and Farm Management. Indicators, such as, ease of use, usefulness, collaboration, increased efficiency, effort reduction, satisfaction and visualisation-interaction, are included in this section. The "Pilot Basis Evaluation" will include (but not be limited to) indicators from the following table:

Table 17: Indicators for Pilot Basis Evaluation for Assessment Group "End-user, BDG pilot partner", after the completion of the Intermediate phase and at the end of the Summative phase.

Indicator	Examples
Ease of Use	How simple is it to perform… How easy is it to predict…
Usefulness	How can the use of the BigDataGrapes components and tools developed for the pilots be compared to the current use of hardware and software on…
Collaboration	How often do stakeholders with different skills work together on the Grapevine-powered Industries Big Data Challenges?
Increased Efficiency	Do the BigDataGrapes components and tools developed for the pilots enable increased efficiency on…
Time Reduction	Do the BigDataGrapes components and tools developed for the pilots allow for the reduction in time on…
Effort Reduction	Do the BigDataGrapes components and tools developed for the pilots contribute to effort reduction on…
Satisfaction	What is your overall satisfaction for the BigDataGrapes components and tools developed for the pilots regarding…
Visualisation-Interaction	What is your overall satisfaction for the BigDataGrapes components and tools developed for the pilots regarding… Do the BigDataGrapes components and tools developed for the pilots encourage User Interaction?

3.1.2 Vineyard Information and Demographics

The second part of this level of assessment is a complementary section, the "Vineyard Information and Demographics" that incorporates questions related to the vineyard of interest, final product and end-user of each pilot. This part of the survey is intended to be completed by the end-users and the BDG pilot partners.

Prior to each survey distribution, an introductory presentation on the piloting activities and objectives will be given to the end-users (farmers, producers, owners of the reference vineyards, agronomists, and other stakeholders) to ensure improved survey validity.

3.2 TECHNOLOGICAL BASIS EVALUATION

The "Technological Basis Evaluation" focuses on the evaluation of the technical solutions proposed within the pilots, regarding their effectiveness. Real-world settings allow showcasing and evaluating the BigDataGrapes platform and components in the context of specific end-user requirements from different areas. The developed BigDataGrapes methods and tools go beyond the state-of-the-art in Big Data management, processing and leveraging data value. Therefore, this section includes questions on Data Management, Data Analytics and Processing, Data Modelling and Semantics and Data Visualisation and User Interaction. The "Technological Basis Evaluation" section includes the "Global Technological Evaluation", which is directed to both BDG pilot and technological partners, and the "Data Management, Analytics, Modelling and Visualisation" section, exclusively addressed to the BDG tech partners.

3.2.1 Global Technological Evaluation

The questions in the "Global Technological Evaluation" are formed in such a way in order to examine the FAIR-ness, scalability, resource optimization, openness-flexibility, consistency, reliability, conformity and collaboration within the pilots. Additionally, the size of the datasets that the pilots enable to deal with is another quantitative indicator.

Table 18: Indicators for Technological Basis Evaluation for Assessment Group "BDG pilot and tech partner", after the completion of the Intermediate phase and at the end of the Summative phase.

Global Technological Evaluation	
Indicator	Example Questions
FAIR-ness	Do the BigDataGrapes pilots help in making research data and algorithms FAIR (Findable, Accessible, Interoperable, Reusable)? Do they support access and integrate data from various data sources and of different types (textual, SQL, RDF, images, location data, etc.)? Are there advantages compared to current research environments and data management practices?
Scalability	Do the BigDataGrapes pilots allow for improved scalability in data?
Resource Optimisation	Do the BigDataGrapes pilots enable management capabilities and resource optimisation (such as for security, compute resource management, governance and reuse)?

Openness- Flexibility	Do the BigDataGrapes pilots permit the addition of new data and enhanced functionality? Do they promote the transfer of knowledge by sharing components and data? Do they promote flexibility by the integration of open-source data and libraries?
Consistency	Are the BigDataGrapes pilots consistent and integrated in order to support an entire data analytics pipeline? Do they provide a seamless end-to-end experience, to make users more productive across the whole data and analytics pipeline?
Reliability	Are the BigDataGrapes components and tools developed for the pilots reliable enough for day-to-day use?
Conformity	Are results from the pilots conforming to user expectations and quality standards?
Collaboration	How do stakeholders with different skills work together on the Grapevine-powered Industries Big Data Challenges?

3.2.2 Data Management, Analytics, Modelling and Visualisation

The main goal of the last section of the survey will be to provide an evaluation of the technological progress of each of the four pilots in terms of time and effort reduction, enhanced performance, satisfaction and user interaction. The "Technological Pilot Evaluation" will also contribute to the development and improvement of the following:

- Resource management: management of data and tools.
- Data access: access and integration of data from various data sources with a focus on semantic issues.
- Machine learning: access to machine-learning approaches and also include support for modern machine-learning approaches like ensemble techniques (boosting, bagging and random forests) and deep learning.
- Flexibility, extensibility and openness: Integration of open-source data and libraries to the piloting activities.
- Data exploration and visualization: to provide interactive visualization.
- User interface: BigDataGrapes must provide coherent "look and feel" and support for the visual components for the application pilots.
- Collaboration: promote the collaboration between users with different skills.

Table 19: Evaluation for Data Management, Analytics, Modelling and Visualisation for Assessment Group "BDG tech partner", after the completion of the Intermediate phase and at the end of the Summative phase.

Data Management, Analytics, Modelling and Visualisation	
Evaluation regarding	Example Questions

Data Management	What is your overall **satisfaction** for the BigDataGrapes pilots regarding data **volume**? Data **variety**? Data **velocity**?
Data Analytics and Processing	In terms of Data Analytics and Processing, do the BigDataGrapes pilots allow the **reduction in time**? Contribute to **effort reduction**? Enable **enhanced performance**?
Data Modelling and Semantics	Do the BigDataGrapes pilots allow the **reduction in time** for Semantic Annotation? Initiate the Semantic Interoperability?
Data Visualisation and end-user interaction	Do the BigDataGrapes pilots encourage **User Interaction**? Encourage interactive **visual analytics** of multiple scale data?

4 DISCUSSION AND CONCLUSIONS

This deliverable, the "Evaluation Report and KPI Assessment", belongs to WP8, "Grapevine-powered Industry Application Pilots". This work package is responsible for the planning and preparation of pilot, the definition of the experimental and evaluation protocols to be followed, the execution of the pilots and ultimately, the collection and evaluation of the pilot results and their assessment over indicators defined by the end users. The deliverable aims to provide a report on the results of the application piloting sessions, inline with the defined experimental protocols and in accordance with the evaluation methodology, providing an overview and a first evaluation regarding each of the four pilots' progress. It states and explains the current status of development, while the implementation and achieved performance of the BDG pilots are assessed.

Evaluation is to be both formative and summative. The former is essentially self-assessment and will be carried out by all partners through filling the "Qualitative and Quantitative Evaluation", which is consisted of a total of four reports that displays the current status of the piloting activities and thus it is providing tangible results. The summative evaluation will involve external as well as internal evaluation in the form of "BigDataGrapes Pilots' Survey". This deliverable describes the structure of the BDG Survey, which is to be distributed to the end-user a few months after the completion of the Intermediate phase (M22) and two months before the end of the Summative phase (M34).

During the first year of piloting sessions, all four pilots successfully gathered data from their respective experimental sites. The individual reports have been analysed and the results have shown that the gross data volume resulted in a total of over ~ 690 GB throughout the entire year, with emphasis given to each site's respective crop season for data acquisition procedures. More specifically, all pilot partners used ~ 43 different data sources to generate ~ 39 unique datasets. One of the most popular characterization methods of Big Data is the "3V", representing Volume, Variety and Velocity of data generated respectively. From the pilots' data, it becomes obvious that out of the 3 "V"s, the Variety aspect has met the sufficient requirements. The first will represent the baseline value for the most KPIs, over which to measure the improvements in the coming two years. The list will be continuously updated during the project's lifetime.

The survey is divided in two levels of assessment, the "Pilot Basis Evaluation" and the "Technological Basis Evaluation", including heterogeneous sets of quantitative and qualitative indicators, metrics used to measure the effectiveness of the pilots. The "Pilot Basis Evaluation" includes the "End-user survey" that is directly related to these Use Cases, covering Data Anomaly Detection and Classification, Prediction and Farm Management. Indicators, such as, ease of use, usefulness, collaboration, increased efficiency, effort reduction, satisfaction and visualisation-interaction, are included in this section. The questions in the "Technological Basis Evaluation" are formed in such a way in order to examine the FAIR-ness, scalability, resource optimization, openness-flexibility, consistency, reliability, conformity and collaboration within the pilots. Additionally, the size of the datasets that the pilots enable to deal with is another quantitative indicator. Finally, the survey includes questions on Data Management, Data Analytics and Processing, Data Modelling and Semantics and Data Visualisation and User Interaction. The main goal of the last section of the survey will be to provide an evaluation of the technological progress of each of the four pilots in terms of time and effort reduction, enhanced performance, satisfaction and user interaction.

To conclude, the evaluation progress report submitted by the four pilots clearly prove that BDG project has a disruptive innovative potential to bring new and market driven ICT technologies into the grapevine-powered

industries and that the majority of the pilots are performing their activities successfully and timely, showing professionalism and technical skills. The project's Evaluation Report and KPI Assessment presented in this report is a constantly updating roadmap to an optimal evaluation of the BDG pilots, tools and components, while being aligned with the project vision and objectives. An updated version of this deliverable, including a potential refined evaluation approach, is due in M36 of the project lifetime.

5 REFERENCES

Abed et al. (2015). In vitro assessment of cytotoxic, antioxidant and antimicrobial activities of leaves from two grape varieties collected from arid and temperate regions in Palestine. QScience Connect 4.pp. 1-9.

Anastasiou, E., Tsiropoulos, Z., Balafoutis, T., Fountas, S., Templalexis, C., Lentzou, D., & Xanthopoulos, G. (2017). Spatiotemporal stability of management zones in a table grapes vineyard in Greece. Advances in Animal Biosciences, 8(2), 510-514.

Balafoutis, A. T., Koundouras, S., Anastasiou, E., Fountas, S., & Arvanitis, K. (2017). Life Cycle Assessment of Two Vineyards after the Application of Precision Viticulture Techniques: A Case Study. Sustainability, 9(11), 1997.

Barreales et al. (2019). Effects of irrigation and collection period on grapevine leaf (Vitis vinifera L. var. Touriga Nacional): Evaluation of the phytochemical composition and antioxidant properties. Scientia Horticulturae 245, pp. 74–81.

Biniari, K., Gerogiannis, O., Daskalakis, I., Bouza, D., &Stavrakaki, M. (2018). Study of Some Qualitative and Quantitative Characters of the Grapes of Indigenous Greek Grapevine Varieties (Vitis vinifera L.) using HPLC and Spectrophotometric Analyses. NotulaeBotanicaeHortiAgrobotanici, 46(1), 97-106.

Bordiga M. (Ed.) (2015). "Valorization of wine making by-products" CRC Press.

Ferhi et al. (2019). Total Phenols from Grape Leaves Counteract Cell Proliferation and Modulate Apoptosis-Related Gene Expression in MCF-7 and HepG2 Human Cancer Cell Lines. Molecules, 24, pp. 612.

Fiume, M.M. (2012). Safety Assessment of Vitis Vinifera (Grape)-Derived Ingredients. Cosmetic Ingredient Review Expert Panel 2012. Pp. 1-35.

Rasimah, C. M. Y., Halimah, B.Z., Azlina, A. (2011). Evaluation of user acceptance of mixed reality technology, Australasian Journal of Educational Technology, 27 (Special issue, 8), 1369-1387.

Romain David, Laurence Mabile, Mohamed Yahia, Anne Cambon-Thomsen, Anne-Sophie Archambeau, et al. (2019). How to assess FAIRness to improve crediting and rewarding processes for data sharing? A step forward towards an extensive assessment grid. *RDA 13th (P13) Plenary Meeting*, Philadelphia, United States. ⟨https://www.rd-alliance.org/rda-13th-plenary-meeting-information⟩. ⟨10.5281/zenodo.2625721⟩. ⟨hal-02094678⟩.

Stavrakaki, M., & Biniari, K. (2016). Genotyping and phenotyping of twenty old traditional Greek grapevine varieties (Vitis vinifera L.) from Eastern and Western Greece. Scientia Horticulturae, 209, 86-95.

Teixeira, A., Baenas, N., Dominguez-Perles, R. et al. (2014). Natural Bioactive Compounds from Winery By-Products as Health Promoters: A Review. International Journal of Molecular Sciences, 15(9), 15638-15678.

APPENDIX- BIGDATAGRAPES PILOTS' SURVEY

PILOT BASIS EVALUATION

Select the BigDataGrapes Pilot you are involved in?

☐*Table and Wine Grapes* ☐*Winemaking* ☐ *Farm Management* ☐*Natural Cosmetics*

End-user Survey

Please answer the following questions if you are a **farmer, producer, owner** or other **BigDataGrapes pilot partner**:

Q1. Please select all indicators on which the BigDataGrapes pilots have an impact on:

☐*Ease of use* ☐*Usefulness* ☐*Collaboration* ☐*Increased efficiency* ☐*Time reduction*

☐*Effort reduction* ☐*Satisfaction* ☐*Visualisation-Interaction*

Ease of Use

Q2. How **simple** is it to perform…

- Earth Observation Data Anomaly Detection & Classification?

☐ *Difficult* ☐*Somewhat difficult* ☐ *Neutral* ☐ *Somewhat simple* ☐*Very simple*

- Optimization of Farm Practices in the Vineyard?

☐ *Difficult* ☐*Somewhat difficult* ☐ *Neutral* ☐ *Somewhat simple* ☐*Very simple*

- Management Zones Delineation for Vineyards?

☐ *Difficult* ☐*Somewhat difficult* ☐ *Neutral* ☐ *Somewhat simple* ☐*Very simple*

Q3. How **easy** is it to predict…

- Total Yield?

☐ *Difficult* ☐*Somewhat difficult* ☐ *Neutral* ☐ *Somewhat easy* ☐*Very easy*

- Biological Efficacy?

☐ *Difficult* ☐*Somewhat difficult* ☐ *Neutral* ☐ *Somewhat easy* ☐*Very easy*

- Crop Quality for Optimizing Post Harvest Treatments?

☐ *Difficult* ☐*Somewhat difficult* ☐ *Neutral* ☐ *Somewhat easy* ☐*Very easy*

- Crop Quality for Optimizing Winemaking?

☐ *Difficult* ☐*Somewhat difficult* ☐ *Neutral* ☐ *Somewhat easy* ☐*Very easy*

Usefulness

Q4. How can the **use** of the BigDataGrapes components and tools developed for the pilots be compared to the current use of hardware and software on…

- Earth Observation Data Anomaly Detection & Classification?

☐ *Not useful* ☐*Not very useful* ☐ *Neutral* ☐ *Somewhat useful* ☐*Very useful*

- Total Yield Prediction?

☐ *Not useful* ☐*Not very useful* ☐ *Neutral* ☐ *Somewhat useful* ☐*Very useful*

- Biological Efficacy Prediction?

☐ *Not useful* ☐*Not very useful* ☐ *Neutral* ☐ *Somewhat useful* ☐*Very useful*

- Crop Quality Prediction for Optimizing Post Harvest Treatments?

☐ *Not useful* ☐*Not very useful* ☐ *Neutral* ☐ *Somewhat useful* ☐*Very useful*

- Crop Quality Prediction for Optimizing Winemaking?

☐ *Not useful* ☐*Not very useful* ☐ *Neutral* ☐ *Somewhat useful* ☐*Very useful*

- Management Zones Delineation for Vineyards?

☐ *Not useful* ☐*Not very useful* ☐ *Neutral* ☐ *Somewhat useful* ☐*Very useful*

- Optimization of Farm Practices in the Vineyard?

☐ *Not useful* ☐*Not very useful* ☐ *Neutral* ☐ *Somewhat useful* ☐*Very useful*

Collaboration

Q5. How often do stakeholders with different skills **work together** on the Grapevine-powered Industries Big Data Challenges?

☐ *Rarely* ☐ *Somewhat rarely* ☐ *Frequently* ☐ *Somewhat regularly* ☐*Regularly*

Increased Efficiency

Q6. Do the BigDataGrapes components and tools developed for the pilots enable **increased efficiency** on…

- Earth Observation Data Anomaly Detection & Classification?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Total Yield Prediction?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Biological Efficacy Prediction?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Crop Quality Prediction for Optimizing Post Harvest Treatments?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Crop Quality Prediction for Optimizing Winemaking?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Management Zones Delineation for Vineyards?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Optimization of Farm Practices in the Vineyard?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Time Reduction

Q7. Do the BigDataGrapes components and tools developed for the pilots allow for the **reduction in time** on…

- Earth Observation Data Anomaly Detection & Classification?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Total Yield Prediction?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Biological Efficacy Prediction?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Crop Quality Prediction for Optimizing Post Harvest Treatments?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Crop Quality Prediction for Optimizing Winemaking?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Management Zones Delineation for Vineyards?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Optimization of Farm Practices in the Vineyard?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Effort Reduction

Q8. Do the BigDataGrapes components and tools developed for the pilots contribute to **effort reduction** on…

- Earth Observation Data Anomaly Detection & Classification?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Total Yield Prediction?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Biological Efficacy Prediction?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Crop Quality Prediction for Optimizing Post Harvest Treatments?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Crop Quality Prediction for Optimizing Winemaking?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Management Zones Delineation for Vineyards?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Optimization of Farm Practices in the Vineyard?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Satisfaction

Q9. What is your **overall satisfaction** for the BigDataGrapes components and tools developed for the pilots regarding…

- Earth Observation Data Anomaly Detection & Classification?

☐*Unsatisfied* ☐*Somewhat unsatisfied* ☐*Neutral* ☐*Somewhat satisfied* ☐*Satisfied*

- Total Yield Prediction?

☐*Unsatisfied* ☐*Somewhat unsatisfied* ☐*Neutral* ☐*Somewhat satisfied* ☐*Satisfied*

- Biological Efficacy Prediction?

☐*Unsatisfied* ☐*Somewhat unsatisfied* ☐*Neutral* ☐*Somewhat satisfied* ☐*Satisfied*

- Crop Quality Prediction for Optimizing Post Harvest Treatments?

☐*Unsatisfied* ☐*Somewhat unsatisfied* ☐*Neutral* ☐*Somewhat satisfied* ☐*Satisfied*

- Crop Quality Prediction for Optimizing Winemaking?

☐*Unsatisfied* ☐*Somewhat unsatisfied* ☐*Neutral* ☐*Somewhat satisfied* ☐*Satisfied*

- Management Zones Delineation for Vineyards?

☐*Unsatisfied* ☐*Somewhat unsatisfied* ☐*Neutral* ☐*Somewhat satisfied* ☐*Satisfied*

- Optimization of Farm Practices in the Vineyard?

☐*Unsatisfied* ☐*Somewhat unsatisfied* ☐*Neutral* ☐*Somewhat satisfied* ☐*Satisfied*

Visualisation and Interaction

Q10. What is your **overall satisfaction** for the BigDataGrapes components and tools developed for the pilots regarding…

- First use experience?

☐*Unsatisfied* ☐*Somewhat unsatisfied* ☐*Neutral* ☐*Somewhat satisfied* ☐*Satisfied*

- Overall usage experience?

☐*Unsatisfied* ☐*Somewhat unsatisfied* ☐*Neutral* ☐*Somewhat satisfied* ☐*Satisfied*

Q11. Do the BigDataGrapes components and tools developed for the pilots encourage User Interaction?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Vineyard Information and Demographics

Please answer the following questions if you are a **farmer, producer, owner** or other **BigDataGrapes pilot partner**:

Vineyard Information

Q12. What are the **final grapes' products**?

☐*Table grapes* ☐*Winemaking grapes* ☐ *Raisins or Currants* ☐*Natural Cosmetics*

Q13. Which are the grape **varieties** in use?

……
……
……
……
……

Q14. What is the **total area** of the vineyard(s) of interest?

☐*Less than 1 ha* ☐*1-10 ha* ☐*10-50 ha* ☐*More than 50 ha*

Q15. What type of **farming system** do you practice?

☐*Conventional* ☐*Organic* ☐*Bio-dynamic*

Demographics

Q16. What is your **country**?

☐*Greece* ☐*France* ☐*Italy* ☐*Other*

Q17. What is your specific **location**?

……
……
……
……
……

Q18. What is your **gender**?

☐*Male* ☐*Female*

Q19. What is your **age group**?

☐*18-34* ☐*35-44* ☐*45-54* ☐*55-64* ☐*65+*

Q20. What is your **education level**?

☐*Primary* ☐*Secondary* ☐*Graduate* ☐*Post graduate*

TECHNOLOGICAL BASIS EVALUATION

Global Technological Evaluation

Please answer the following questions if you are a **BigDataGrapes pilot and technological partner**:

Q21. Please select all indicators on which the BigDataGrapes pilots have an impact on:

☐ *FAIR-ness* ☐*Scalability* ☐*Resource Optimisation* ☐*Openness-Flexibility*

☐ *Consistency* ☐*Reliability* ☐*Conformity* ☐*Collaboration*

FAIR-ness

Q22. Do the BigDataGrapes pilots…

- Help in making research data and algorithms FAIR (Findable, Accessible, Interoperable, Reusable)?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Support access and integrate data from various data sources and of different types (textual, SQL, RDF, images, location data, etc.)?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Q23. Are there advantages compared to current research environments and data management practices?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Scalability

Q24. Do the BigDataGrapes pilots allow for improved scalability in data?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Resource Optimisation

Q25. Do the BigDataGrapes pilots enable management capabilities and resource optimisation (such as for security, compute resource management, governance and reuse)?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Openness - Flexibility

Q26. Do the the BigDataGrapes pilots…

- Permit the addition of new data and enhanced functionality?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Promote the transfer of knowledge by sharing components and data?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Promote flexibility by the integration of open-source data and libraries?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Consistency

Q27. Are the BigDataGrapes pilots consistent and integrated in order to support an entire data analytics pipeline?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Q28. Do they provide a seamless end-to-end experience, to make users more productive across the whole data and analytics pipeline?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Reliability

Q29. Are the BigDataGrapes components and tools developed for the pilots reliable enough for day-to-day use?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Conformity

Q30. Are results from the pilots conforming to user expectations and quality standards?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Collaboration

Q31. How do stakeholders with different skills work together on the Grapevine-powered Industries Big Data Challenges?

☐ *Rarely* ☐ *Somewhat rarely* ☐ *Frequently* ☐ *Somewhat regularly* ☐*Regularly*

Data Management, Analytics, Modelling and Visualisation

Please answer the following questions if you are a **BigDataGrapes technological partner**:

Data Management

Q32. What is your overall satisfaction for the BigDataGrapes pilots regarding…

- Data Volume?

☐*Unsatisfied* ☐*Somewhat unsatisfied* ☐*Neutral* ☐*Somewhat satisfied* ☐*Satisfied*

- Data Variety?

☐*Unsatisfied* ☐*Somewhat unsatisfied* ☐*Neutral* ☐*Somewhat satisfied* ☐*Satisfied*

- Data Velocity?

☐*Unsatisfied* ☐*Somewhat unsatisfied* ☐*Neutral* ☐*Somewhat satisfied* ☐*Satisfied*

Data Analytics and Processing

Q33. In terms of Data Analytics and Processing, do the BigDataGrapes pilots…

- Allow the reduction in time?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Contribute to effort reduction?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Enable enhanced performance?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Data Modelling and Semantics
Q34. Do the the BigDataGrapes pilots…

- Allow the reduction in time for Semantic Annotation?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Initiate the Semantic Interoperability?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

Data Visualisation and end-user Interaction
Q35. Do the the BigDataGrapes pilots…

- Encourage User Interaction?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*

- Encourage interactive visual analytics of multiple scale data?

☐*Definitely not* ☐*Probably not* ☐*Neutral* ☐*Probably yes* ☐*Definitely yes*