

D5.2

Project Acronym: PRECIMED

Project full Name: Precision Irrigation Management to Improve Water and Nutrient Use Efficiency in the Mediterranean Region

Data Management Plan

Due date	01/04/2020
Actual submission date	31/07/2020
Project start date	01/10/2019
Duration	36 months
Action(s) concerned	Project Management
Nature	PU
Author	CSIC
Contributor	ODIN

Disclaimer

This deliverable **a.** Reflects only the author's view; and **b.** Exempts the Agency for any use that may be made of the information it contains.



Document history

Date	Author	Action	Status
03/04/2020	M.F. Ortuño (CSIC)	First draft deliverable	Draft version
08/06/2020	A. Skarmeta, J.A. Martínez (ODIN)	Revision data management	Draft version
01/07/2020	M.F. Ortuño (CSIC)	Revision final draft	Final version
14/07/2020	M. Semiani (INRAA)	Minor revision	Approved for release
29/07/2020	N. Katsoulas (UTH)	Minor revisions	Approved
30/07/2020	M.F. Ortuño (CSIC)	Release of the revised final version	Approved by PC



Content

Document history	1
1. Applicable Documents.....	3
1.1. General References	3
1.2. Specific References.....	3
2. Data Summary	4
3. Fair Data	6
3.1- Making data findable, including provisions for metadata.....	6
3.2- Making data openly accessible.....	10
3.3- Making data interoperable.....	11
3.4- Increase data re-use (through clarifying licences).....	11
4. Allocation of resources.....	12
5. Data security.....	12
6. Ethical aspects	12
7. Other.....	12

1. Applicable Documents

1.1. General References

Type of document	Document title	Edition	Date
Contractual	Grant Agreement	*	*
Guide	AMGA	*	*
Guide	Guidelines on FAIR Data Management in Horizon 2020	*	*

*Last edition will be applicable

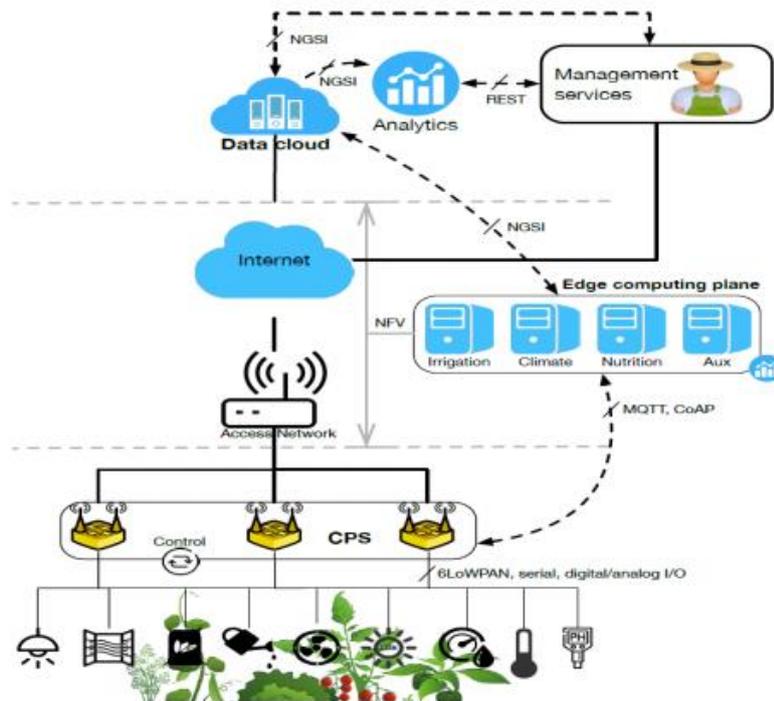
1.2. Specific References

Code	Type of document	Document title	Edition	Date
NA	Contractual	Consortium Agreement	*	*
Precimed-D1.2-Quality Assurance Plan–CO-FV	Procedure	Quality Assurance Plan	*	*

* Last edition will be applicable

2. Data Summary

The main objective of PRECIMED will be the development, validation and transfer of a data driven irrigation/fertilization management system, in order to improve Water and Nutrient Use Efficiency in the Mediterranean Region, by integrating the knowledge about fertilizers and irrigation water management with Information Communication Technologies (ICTs).



The formats to generate, disclose and store PRECIMED’s data will be both digital and printed. As to the main types of produced data, we may highlight the following ones:

- Project deliverables, which will collect the progress of PRECIMED.
- Data that Consortium partners bring to PRECIMED as background IP, that have been described in the Consortium Agreement.
- External communication documents: all the public documents to communicate PRECIMED and the farms results, such as technology and business analysis, presentations, guidelines and manuals, articles, papers, newsletters, etc.
- Internal communication documents: They will be aimed at keeping all data generated during the project, mainly those related to communication among Consortium partners, such as agendas and minutes of meetings, emails, proceedings, agreements, etc.

There are existing data that will feed PRECIMED, both at technical and strategic level. Among them, we find those relevant national and international activities linked to the project, whose main outcomes will be helpful to be used as PRECIMED’s inputs. Furthermore, the project will make use of any existing guidebook produced by the European Commission or any other policy-maker focused on clusters, social economy and entrepreneurship.

The project will NOT opt-out of the Open Research Data Pilot; therefore, in order to manage the data in an effective way, the following measures will be implemented:

- Obtained information and research data, including associated metadata that are needed to validate the results published in scientific publications, will be deposited in a research data repository and will be made accessible for third parties to access, mine and reuse free of charge. In addition to this, the PRECIMED partners will provide information on the tools and services that were used to extract insights from the data and could be used by third parties to validate them, whenever possible, these tools and instruments will be provided.
- For other data and metadata that is not supporting scientific publications the goal will be to make it publicly available and free of charge.

PRECIMED's data gathered will be mainly stored in a software tool with restricted access and a back-up feature. All the partners will have access to the system according to their roles in the project, and the Coordinator will be in charge of managing access permissions, usage procedures and basic training.

A brief summary of the data sets, listed by number that have been identified (June 2020) is presented below:

Dataset 1: Sustainability of agricultural productivity in Mediterranean agrosystems (scarcity of water, low quality water, low water and nutrients use efficiency, inappropriate agricultural practices, etc.) minimizing the use of water and fertilizers by optimizing fertirrigation management which will include climate (temperature, solar radiation, wind speed, rain and air humidity), data on soil water storage, physical and chemical characteristics of soil (soil texture, soil water retention, hydraulic conductivity, total availability water, surface runoff, soil evaporation, ph, nutrients and electrical conductivity), quality/chemical composition of irrigation water, multispectral data and remote sensing of environmental variables.

Dataset 4: Integration of Agro-photovoltaic applications with respect to the reduction of irrigation needs to avoid or mitigate soil salinization and to improve water productivity. It includes data on isolated rural areas related to benefits for aquifer protection, water treatment and irrigation systems using photovoltaic energy and reduction of water consumption by the shading (for protected crops) and by the deficit irrigation strategies. The integration of photovoltaic applications under the large scale irrigation systems is, also; interesting to increase the economic efficiency of the agricultural exploitations.

Dataset 7: Strategies to optimize the use of pollutants sources (such as nitrates, phosphates, ammonium, pesticides "plant protection products") in the pilot zones, which will contain data on surface and ground water used for irrigation and soil pollutants linked to inappropriate agricultural practices as fertilization, the reuse of waste water and the waste sludge, inappropriate irrigation scheduling etc.

Dataset 8: Optimization of crop types in pilot areas in relation to their tolerance/sensibility to water stress and salinity, pollution and type of soil. This will include data on crops physio-phenology, soil and water quality and agricultural productivity.

Dataset 9: Evaluation of benefits and risks that will contain data related to the benefits and risks of the PRECIMED solutions, including data of predictive modeling of the quality and consumption of water in agriculture according to the different types of crops and areas, on the concentrations of beneficial and harmful nutrients in relation to the crop and lastly, on the economic impact in agricultural exploitations.

Dataset 10: Market opportunities and technological demonstration, which will contain data related to the needs of end users of PRECIMED solutions.

Dataset 11: Stakeholder Engagement, which will contain personal data relating to stakeholders who might benefit from PRECIMED solutions.

3. Fair Data

3.1- Making data findable, including provisions for metadata

Efficient data management is vital for the success of projects. This is even more important for large-scale integrating projects, like PRECIMED, with a significant number of farms participants and tasks demanding adequate reporting procedures and collaboration support.

Basic methodology will comprise the considerations on how to manage the data and information created during the project, in order to precisely describe all procedures for keeping and disseminating PRECIMED results, through the various stages of execution, and according to their usages and the way that the audience accesses information. The data related to the project deliverables, farms, external and internal communication documents will follow specific procedures which will be developed by PRECIMED's Quality Manager and defined in the Quality Assurance Plan. These types of data will be stored in a secured, software platform, which will be accessible to all partners and external entities to the Consortium when needed.

Regarding the representation of the information that PRECIMED is going to use, there are different technologies that can be used. Among these alternatives, there are open and standard initiatives, such as NGSI, which is promoted by FIWARE and which allows for a homogeneous way of representation heterogeneous information based on JSON documents. NGSI, in its current v2 version, represents contextual entities following a really straightforward approach. Contextual entities are composed of the following properties: an identifier, a type and a series of attributes, as we can see in the following example.

```
{
  "id": "April012019",
  "type": "MurciaCrop",
  "humidity": {
    "value": 0.6},
  "temperature": {
    "value": 24.6},
  "location": {
    "type": "geo-json",
    "value": {
      "type": "Point",
      "coordinates": [38.4409, -1,3342]}
    }
}
```

This NGSI representation allows also for expressing meta-information by using a specific metadata property.

Based on the different information that must be represented in the frame of this project, concerning to agricultural plots, we have identified the following entities that must be managed:

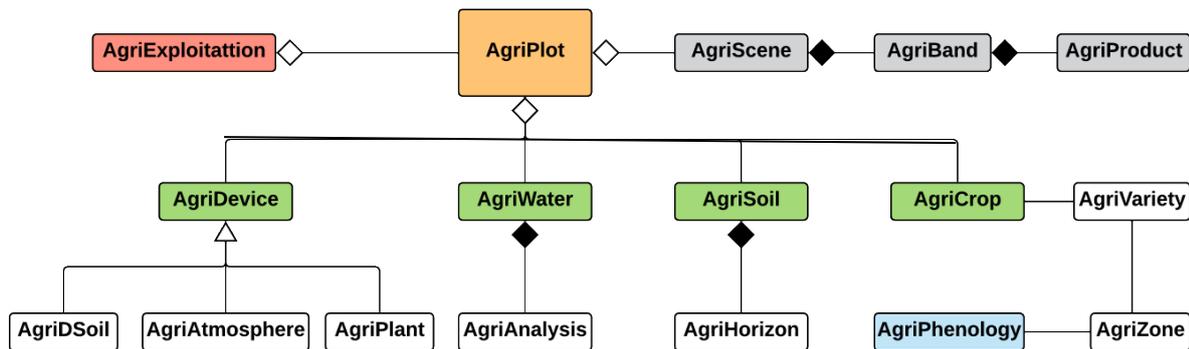


Figure 1: Data model identified for agricultural plot management

Each AgriPlot comprises AgriDevices, AgriWater, AgriSoil and AgriCrop:

- AgriDevice: Represents the IoT devices which can be used for monitoring soil parameters, weather conditions and the plant production.
- AgriWater: Represents the information related to water quality parameters.
- AgriSoil: By its horizons, we can represent the morphological features such as colour, texture, consistence, structure and the likes.
- AgriCrop: Entity representing the sort of crop, its variety, and also its phenological status in relation to the agroecological zone.

Finally, an AgriPlot is linked with an AgriScene, where we can represent this way, the resulting product of the agricultural plot. Finally, AgriPlots are linked with AgriExploitation, allowing this way for linking an agricultural exploitation to several plots.

Following the guidelines, also, provided by FIWARE Data models, and in other cases inspired by them, we have specified the following information per each of the entities defined above. Thus, starting with AgriDevice, this would be the sort of information that the platform would manage:

```

{
  "id": "urn:ngsi:AgriDevice:AgriDSoil:01LaPuebla",
  "type": "AgriDevice",
  "category": {
    "value": ["Soil"]
  },
},
  "batteryLevel": {
    "value": 0.75
  },
},
  "dateFirstUsed": {
    "type": "DateTime",
    "value": "2014-09-11T11:00:00Z"
  },
},
  "controlledAsset": {
    "value": ["soil-RegMurcia-1"]
  },
},
  "serialNumber": {

```



```

    "value": "9845A"
  },
  "mcc": {
    "value": "214"
  },
  "value": {
    "value": "1%3D0.22%3Bt%3D21.2"
  },
  "refDeviceModel": {
    "type": "Relationship",
    "value": "myDevice-wastecontainer-sensor-345"
  },
  "rssi": {
    "value": 0.86
  },
  "controlledProperty": {
    "value": ["moisture", "temperature"]
  },
  "owner": {
    "value": ["urn:owner:comunidadregantes"]
  },
  "mnc": {
    "value": "07"
  },
  "ipAddress": {
    "value": ["192.14.56.78"]
  },
  "deviceState": {
    "value": "ok"
  }
}

```

As we can see, there is an attribute called “controlledProperty” by the means of which we can specify the sort of measurements that the device has to perform. Therefore, if we are now representing an AgriAtmosphere or AgriPlant, we will use this attribute to specify other types of measurements related to these domains such as wind speed, temperature, rainfall, humidity, solar radiation and the likes of the former, or leaf humidity, dendrometer, and so on for the latter.

Regarding AgriWater, there are parameters such as the concentration of chlorides, ammonium, nitrates, phosphates and water turbidity, among others, that can be also incorporated to the platform.

A crop would contain information similar to the following one:



```
{
  "id": "urn:ngsi:AgriCrop:Apricot",
  "type": "AgriCrop",
  "commonName": {
    "value": "Apricot"
  },
  "scientificName": {
    "value": "Prunusarmeniaca"
  },
  "plant": {
    "value": "Fruit tree"
  },
  "demandIrrigation": {
    "value": "Yes"
  }
}
```

Finally, an AgriPlot could be defined as follows:

```
{
  "id": "urn:ngsi:AgriPlot:CRM0001",
  "type": "AgriPlot",
  "name": {
    "value": "La Puebla 1"
  },
  "location": {
    "type": "geo:json",
    "value": {
      "type": "Polygon",
      "coordinates": [
        [
          [-3.164485591715449, 40.62785133667262],
          [-3.164445130316209, 40.62787156737224],
          [-3.164394553567159, 40.62777209976578],
          [-3.164424899616589, 40.62775018317452],
          [-3.164485591715449, 40.62785133667262]
        ]
      ]
    }
  },
  "density": {
    "value": "5x4"
  },
  "flowRate": {
    "value": "4"
  }
}
```



```

    },
    "diameter": {
      "value": "5"
    },
    },
    "emitters": {
      "value": "4"
    },
    },
    "varietychoice": {
      "type": "Relationship",
      "value": "urn:ngsi:AgriCrop:XXX"
    },
    },
    "hasConnection": {
      "type": "Relationship",
      "value": "urn:ngsi:AgriDevice:AgriDSoil:01LaPuebla"
    }
  }
}

```

3.2- Making data openly accessible

As described in section 2 “DATA SUMMARY”, there are several kinds of data that will be generated during PRECIMED, whose access rules will be:

- Project deliverables: according to section 3.1. Project Description.
- Data from surveys, questionnaires and interviews, according to the applicable data protection laws.
- Data that Consortium partners bring to PRECIMED as background IP, according to the stated in Attachment 2: “Background included” of the Consortium Agreement.
- External communication and dissemination material: several criteria may be taken into account depending on the data to be disclosed. According to their applicability, potential criteria may be:
 - a. Data protection laws,
 - b. Section 6.3. “Obligation to disseminate results” of the Consortium Agreement,
 - c. Section 6.6.1. “Open access to scientific publications” of the Consortium Agreement,
 - d. Section 6.6.2. “Open access to research data” of the Consortium Agreement
- Internal communication documents: according to applicable data protection regulations laws.

PRECIMED’s data will be mainly stored in a secured, software platform accessible to the entire Consortium according to the access permissions previously agreed by the parties. The system consists of pre-configured software and SQL Database Modules for the implementation in project specific Internet and Intra/Extranets. The user will be able to access the platform by signing in with his username and password. The Coordinator will be in charge of managing access permissions to the platform.

All the procedures needed to use the software tools in an efficient way will be described in the user manual, which is going to be available to all the members of the Consortium.

The software tool will be configured including all the related databases. The Consortium will be responsible for all project specific content and content management. For this task, the Consortium will nominate a project secretary (a member of the coordination team), who will be responsible for it. The planned system availability

per year is 99,9% and it will include a backup server providing replacement of the provider's server within less than 1 working day, when needed.

For the duration of the project and also after that, the provider is committed to use all of the data included in the database only for contractual related issues. The Consortium will receive a copy of all of the databases (SQL format) and a copy of all the static webpages on a CD-ROM or DVD, one month after the termination of the contract. In addition to that, the provider will delete all project data after having handed over copies of the specific data to the Consortium members, after the termination of the contract.

Initially, it is not expected to count on a data access committee for PRECIMED and the person in charge of being the focal point for any data management issue that may arise during the project will be the Quality Manager, belonging to the coordination team.

3.3- Making data interoperable

Given that PRECIMED is not a research project itself, data produced will not be subject to be exchanged and re-used by researchers, institutions, organizations, etc. Therefore, project data will be only interoperable at a Consortium level by means of the utilisation of the software tool and according to the features described in section 3.1. (datasets, metadata vocabularies, standards, methodologies, etc.).

Nevertheless, there will be three main sources of the project's public data that will be accessible to any third party interested in the following subjects:

- PRECIMED's platform,
- PRECIMED's website and
- social networks

3.4- Increase data re-use (through clarifying licences)

PRECIMED does not envisage providing any data licensing, and those disclosable data, according to the agreed disclosure rules, will be freely accessible. It is expected to keep PRECIMED's data for five years after the project's completion, which is a period while which they may be reused.

Project's Quality Assurance Plan establishes how documentation requirements, procedures, records and other documents are maintained and controlled, including retention periods, during PRECIMED's lifecycle.

4. Allocation of resources

Costs for making data FAIR in PRECIMED are those allocated to the software tool utilisation under partners "other implementation and maintenance is 75.000,00€, and its allocation among Consortium's partners will be direct costs" with the concept "Cost of Software Licenses". The budget for the platform will be calculated proportionally to the budget of each of them in the project.

The person responsible for coordinating PRECIMED's data management process will be the Quality Manager, and the General Assembly will be the ultimate decision-making body of the Consortium and responsible for taking major strategic decisions with respect to data management if necessary. It will also promote consensus in case of conflict and, if no consensus can be found, it will take decisions according to the procedures and rules defined in the Consortium Agreement.

5. Data security

For applications distributed over the internet the risk is high to download foreign code. Therefore, PRECIMED's software tool is digitally signed. The first time the user starts the application he will be asked whether he trusts this signature. If some parts are exchanged by foreign code, the application will not start and an error message will be displayed.

The platform always establishes a secured connection via HTTPS with its web service. In case this is not possible, an unsecured connection via HTTP is used and then, the user will see the following message after login: "Unsecure connection established", besides he will get symbols regarding the connection status at the status bar on the left side: "secured" or "unsecured". In addition to that, the platform has a back-up function to keep project data safe during the project execution.

6. Ethical aspects

There are not any ethical or legal issues that can have an impact on data sharing in PRECIMED.

7. Other

The guideline document "FAIR Data Management in Horizon 2020" has been used for the development of PRECIMED's Data Management Plan.