Data Management Plan (1)

Project Management

Stylianos Kotsopoulos

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 130709.

Disclaimer

Copyright message

Document Information

Grant Agreement Number	130709	Acronym		DIANA
Full Title	Detection and Integrated Assessment of Non-authorised water Abstractions using EO			
Horizon 2020 Call	EO-1-2016: Downstream applications			
Type of Action	Innovation Action			
Start Date	1st January 2017	**Duration**		36 Months
Project URL	-			
Document URL	-			
EU Project Officer	Iulia SIMON			
Project Coordinator	Polimachi Simeonidou, Anna Osann			
Deliverable	D7.2: Data Management Plan (1)			
Work Package	WP7 – Project Management			
Date of Delivery	**Contractual**	M3	**Actual**	M3
Nature	ORDP – Open Research Data Pilot	**Dissemination Level**		Public
Lead Beneficiary	AGRO APPS			
Lead Author	Stylianos Kotsopoulos	**Email**	skotsopoulos@agroapps.gr	
	AGRO APPS	**Phone**	+30 6947566800	
Other authors	-			
Reviewer(s)	Luciano Mateos (FERAGUA), Elena Navarro Soriano (FERAGUA)			
Keywords	Data Management Plan			

Document History

Version	Issue Date	Stage	Changes	Contributor
0.1	17/3/2017	Draft	-	AGROAPPS
0.2	18/3/2017	Draft	Minor, in Track Change Mode	FERAGUA
0.3	20/3/2017	Final	Changes Accepted	AGROAPPS

Table of Contents

Executive summary

This deliverable aims to collect information about the data sources that the DIANA project consortium will work with. More specifically, for each work package the partners have to define the data sources they will use and create a Data Management Plan (DMP). The DMP includes a description of data management life cycle for all data sets and specifies which data will be findable, openly accessible, interoperable and re-usable. It also includes information about legal issues, privacy and maintenance. The DMP is a live document and it will be regularly updated as more information becomes available and data issues are resolved.

The deliverable is structured in the following chapters:

- Chapter 1 includes a description of the methodology used
- Chapter 2 includes the description of the DMP Components

Methodology

This document, D7.2 – Data Management Plan (1) is a deliverable of the DIANA project, which is funded by the European Union's Horizon 2020 Programme under Grant Agreement 730109. DIANA is aimed at co-designing and openly demonstrating a commercial service platform that will empower water managers and authorities to optimize the identification and inspection of non-authorised water abstractions for irrigation as well as improve their water management policies and practices, especially in extreme conditions such as drought. The scope of the project is to be co-created and defined along with the users and stakeholders so as to be shaped according to their needs and requirements.

In this deliverable the DIANA consortium presents information about the data that will be collected and generated in the context of the project. The DMP has been based on the updated version of the "Guidelines on FAIR Data Management in Horizon 2020"[1] version 3.0 released on 26th of July 2016 by the European Commission Directorate – General for Research & Innovation. The DIANA DMP addresses the following issues:

- Data Summary
- FAIR data
- Allocation of resources
- Data security
- Ethical aspects
- Other issues

Why is a Data Management Plan needed?

The DMP is an essential document for the DIANA project that addresses issues related to data management. By creating an earlier plan for managing data at the beginning of the project, will save to the consortium time and effort later on. Furthermore, the DMP helps clarify at an early stage individual and collaborative roles and responsibilities and elaborates reviewers by demonstrating whether the consortium has carefully considered how data quality, security, confidentiality and sharing will be managed.

1 European Commission, (26 July 2016*), Guidelines on FAIR Data Management in Horizon 2020*, Version 3.0

Who is responsible for the implementation of DMP?

The DMP is formulated by AGRO APPS from the early stages of the project addressing crucial management aspects of the data that will be collected/ generated in the course of DIANA. AGRO APPS has provided on time all the work package leaders and rest of the partners with a template that includes all the issues that the DMP addresses along with instructions to fill the template.

What kind of data will be affected by DMP?

Early in the project, the consortium will identify the related sets of data that will be utilised during the project lifetime. This includes: i) key data required for the provision of DIANA services such as satellite, EO meteorological and auxiliary data from various sources as well as several other data to be collected and utilised throughout the project such as ii) data gathered through the interview-based survey of the project aimed at revealing the needs and requirements of DIANA users and stakeholders, iii) data collected during the pilots of the project necessary for the co-evaluation and validation of the services and business cases, etc.

According to the "Guidelines on FAIR Data Management in Horizon 2020" the DMP applies to the following types of data:

- Making data findable, including provisions for metadata
- Making data openly accessible
- Making data interoperable
- Increase data re-use

The DMP addresses the existing suitable standards including several open standards that are currently available and can be applied in the framework of DIANA, such as the EU Directive 2007/2/EC on Infrastructure for Spatial Information in the European Community (INSPIRE) which addresses spatial data themes required for environmental applications[2]. Moreover, there are various disciplinary metadata standards[3], however the DIANA consortium has identified a number of available best practices and guidelines for working with Open Data, mostly by organisations or institutions that support and promote Open Data initiatives, and will be taken into account. These include:

[2] http://inspire.jrc.ec.europa.eu/

[3] http://www.dcc.ac.uk/resources/metadata-standards

- Open Data Foundation[4]
- Open Knowledge Foundation[5]
- Open Government Standards[6].

For datasets that do not have standards, the consortium will provide metadata to help secondary users understand and reuse them.

Due to the DIANA products and services are principally intended for commercial use, the DMP addresses the open data licenses for any interested party following a certain "embargo period". After this "embargo period" open access and data sharing for verification and re-use will be possible through a suitable repository.

DMP Components in DIANA

DMP Components in WP1 – Analysis and co-creation

DMP Component	Issues to be addressed
Data Summary	The purpose of the data collection is to identify and analyse the requirements of DIANA users and stakeholders of the co-creation process of DIANA tools. The development and the evaluation is one of the central objectives of the project. The following data formats will be produced: • Excel database of extended stakeholder mapping results; • Audio and video documentation of some events (2-3 files per pilot area). Data will be collected from events (mostly during meetings, but also online) of interaction with stakeholders. The data will be used in the ongoing multi-actor stakeholder process and will be confidential, to remain within each pilot area stakeholder community

[4] http://www.opendatafoundation.org/

[5] https://okfn.org/

[6] http://www.opengovstandards.org/

	(with sharing to be provided only with the consensus by the information providers, i.e. each stakeholder).
Making data findable, including provisions for metadata	N/A (Confidential Data)
Making data openly accessible	N/A (Confidential Data)
Making data interoperable	N/A (Confidential Data)
Increase data re-use	N/A (Confidential Data)
Allocation of resources	N/A (Confidential Data)
Data security	N/A (Confidential Data)
Ethical aspects	N/A (Confidential Data)
Other issues	N/A (Confidential Data)

DMP Components in WP2 – Earth observation data products and services

DMP Component	Issues to be addressed
Data Summary	The purpose of WP2 data is to provide Earth Observation products with key information about: • **Non- authorised water abstraction** through the identification of irrigation activities and the estimation of abstracted/consumed water volumes. The non-authorised water abstraction detection and monitoring service will be the core service of the DIANA platform. Satellite images will be processed in the context of WP3. They will be processed to obtain time series of spectral indices to assist the identification of irrigated area. Maps of irrigated areas will be provided to WP3 which is in charge of providing this information to final users.

	• **Seasonal drought forecasting and monitoring,** by using seasonal climate forecast in combination with hydrological models, a six month ahead drought forecast will be provided to the users for their region of interest. This forecast can be used as input to update existing regional or local drought management plans enabling the authorities to take the needed actions in order to cope with the phenomena. • **Support to the implementation and monitoring of the WFD,** for long-term sustainable water management. The data which will be collected and generated in the WP2 will cover the needs to implement and test the DIANA solutions in all the pilots located in Italy, Spain and Romania. The products will be used in WP3 which is in charge of providing this information to final users. Data will be available in raster and vector format, accessible through GeoServer application, Map server application, PostGIS database and FTP. The origin of the data for WP2, will be from: • Copernicus Scientific Hub per Sentinel Data, (https://scihub.copernicus.eu/dhus/#/home) • USGS for Landsat data, (https://earthexplorer.usgs.gov/); • GFS forecasts and analyses data (ftp://ftpprd.ncep.noaa.gov/pub/data/nccf/com/gfs/prod/) • MODIS 16-day vegetation product (https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13a1)

	• CFS seasonal climate forecasts (ftp://ftpprd.ncep.noaa.gov/pub/data/nccf/com/cfs/prod/) • ERA-Interim Reanalysis data (http://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/) • EOCD, a data service platform implemented by the University of Natural Resources and Life Science (BOKU) which provides access to individual Sentinel-2 granules (ortho-rectified image tiles of 100 by 100 km2 in UTM/WGS84 projection). The service runs on the Earth observation Data Centre (EODC), which is a collaborative IT infrastructure for archiving, processing, and distributing Earth observation (EO) data. To use this service, which provides added value products, it will be necessary to have an agreement with BOKU University. Sentinel-2 data are about 4 GB each, while Landsat around 1 GB each, both compressed. Assuming 3 pilot cases, and a need to have at least two more images per month on a yearly basis, this accounts for over 2 TB of image data. Indices and classification products will account for an additional 20%, hence a total of 2.4 TB of data is foreseen to be generated. Data and products will be useful for the Water User Association, Reclamation and Irrigation Consortia, Paying Agencies, Regional and National Government. They will be ingested by the DIANA platform and disseminated to project stakeholders, while their usefulness will be demonstrated during the pilot cases.
Making data findable, including provisions for metadata	N/A

Making data openly accessible	The image data and the processed products will be available to all stakeholders through a DIANA Platform (see WP3 for details). Registered users will have unlimited access to the products for the duration of the project. A unique identifier will be assigned to each EO data following the provider's identification mechanism. For the added value products (Vegetation index, BOA, etc.) a unique identification separated by underscore (_) will be appended to file name. Versioning identifier will be also appended at the end of the name product starting by 0 (_v00, v01, .., vnn) Searching dataset will be guarantee by build-in keyword search tool of Geoserver, MapServer and PostGIS. INSPIRE metadata will be created for all the EO-based geospatial products that will be generated in the lifetime of the project. All EO data, value added products, code and metadata will be stored in central data repository server of DIANA project. Mirror servers for relevant EO data could be hold by project partners.
Making data interoperable	PostGIS, Mapserver and Geoserver tools will be available for a widely accessible management of EO information. INSPIRE protocol will be used for metadata descriptors. INSPIRE provides typical standard for geospatial data. No standard vocabulary will be used and no ontology mapping is foreseen.
Increase data re-use	To access to the EO added value data, partnership or licensing agreement will be required. The EO-based geospatial products that will be generated in DIANA will be made available for re-use for the project's lifetime and beyond.

	All EO-based products will remain usable after the end of the project. No particular data quality assurance process is followed, and no relevant warranties will be provided. EO-based products will remain re-usable at least two years after the project's conclusion. To access to the EO added value data, partnership or licensing agreement will be required.
Allocation of resources	The costs for maintaining a database of the EO-based products that will be generated to serve the pilot demonstrations are: • Hosting cost: estimated expenditure for the project life time is about 2500 EUR (about 800 EUR for each years beyond project end); • Personnel cost: development cost is indicated in the project; maintenance cost after project life time could be foresee in 0.5 PM per year. The cost of data FAIR will be covered by project resources. Responsibility of data is shared among projects partners.
Data security	Servers are managed by the IT department. They are regularly backed up and secured.
Ethical aspects	N/A
Other issues	N/A

DMP Components in WP3 – DIANA service platform design and development

DMP Component	Issues to be addressed
Data_WP3_1_System_arcitecture_and_design	
Data Summary	Integration framework, component descriptions and dependencies, API descriptions, information flow

	diagram, internal and external interfaces, software and hardware requirements, use case scenarios and testing procedures will be described in deliverable D3.1 Overall architecture, system design and integration framework specification. This will be the basis upon which the system will be built.
Making data findable, including provisions for metadata	After the document is delivered to the EU and the consortium decides to do so, it will become both discoverable and accessible to the public. This document will contain a table indicating all versions of the document, along with who contributed to each version, what the changes where as well as the date a new version was created.
Making data openly accessible	The data will be available in D3.1: Overall architecture, system design and integration framework. The dissemination level is confidential. It will be available through the DIANA wiki for the members of the consortium.
Making data interoperable	N/A
Increase data re-use	This deliverable could be used as an example for technical teams who are interested in building this kind of systems.
Allocation of resources	N/A
Data security	All data will be securely saved in the AGRO APPS premises and will be shared with the rest of the partners using the DIANA wiki.
Ethical aspects	There are no ethical aspects related to the described dataset.
Other issues	N/A

DMP Component	Issues to be addressed
Data_WP3_2_Maps_produced_by_the_Diana_services	
Data Summary	The DIANA platform will be heavily based on map controls that will offer to the users the possibility to

	have access to multiple map layers indicating useful information derived by the DIANA services. Tiff files, bearing geolocation metadata, for each type of map will be produced. We have no indication for the size of each file yet. DIANA will use existing data such as: • Land use-land cover maps • Maps of irrigated and non-irrigated areas • Maps of soil physical properties • Crop maps
Making data findable, including provisions for metadata	Registered users will be able to discover maps that offer them useful information for the completion of their duties. As part of every user action we will produce meaningful metadata (time and date of data creation or data amendments, owners of actions that took place, service that produced the map). Metadata will assist the discoverability of the data and related information.
Making data openly accessible	Only signed in users will have access to the produced maps. Given that DIANA will be a paid product, users will have to pay in order to access the maps produced. The maps and the metadata will be made available for use by the DIANA or third party applications through the secure API that we will create. Only the authorized technical team will have access to the hard drive that will host all the produced tiff files.
Making data interoperable	N/A
Increase data re-use	Maps produced during the project, will be offered to anyone who asks for it. After the end of the project, these data will only be available to users who will buy the product. Paying users will be able to download such data and use it for their own purposes.
Allocation of resources	We plan to preserve map tiff files for long time in order to offer the users the opportunity to go back in time and compare current farm conditions with those of the past.

Data security	All data generated by the platform will be saved on the DIANA server. SSL connections will be used to ensure secure exchange of information. In case of necessary updates, the old data will be overwritten and all actions will be audited in detail. A log will be kept, containing the changed text for security reasons. Daily backups for a period of 3 days will be kept. All backups will be hosted on a remote server to avoid disaster scenarios. All servers will be hosted behind firewalls inspecting all incoming requests against known vulnerabilities such as SQL injection, cookie tampering and cross-site scripting. Finally, IP restriction will enforce the secure storage of data.
Ethical aspects	There are no ethical aspects related to the described dataset.
Other issues	N/A

DMP Component	Issues to be addressed
Data_WP3_3_Statistical_Reports	
Data Summary	EO products and service outputs will be further processed in order to produce statistical reports in the form of charts, graphs and tables. DIANA users will use this information to make better decisions.
Making data findable, including provisions for metadata	Registered users will be able to discover reports, charts and graphs corresponding to their area of interest and selected time period. Users will be able to export the reports and use them the way they want. Metadata containing the dates the reports/charts/graphs were created, the user for whom they were produced and number of downloads will also be produced.

Making data openly accessible	Only registered users will have access to reports, graphs and charts. As a private product, users will have to pay in order to access the maps produced. The reports and the metadata will be made available for use by the DIANA and third party applications through the secure API that we will create. The database that will host the above data and any related metadata will not be discoverable to other network machines operating on the same LAN, VLAN with the DB server or other networks. Therefore only users with access to the server (DIANA technical team members) will be able to discover the database.
Making data interoperable	Statistical reports will be downloadable in xml format so that it is easy to be further used for other purposes.
Increase data re-use	Reports produced during the project, will be offered to anyone who asks for it. After the end of the project, these data will only be available to users who buy the product. Paying users will be able to download such data and use it for their own purposes.
Allocation of resources	We plan to preserve these data for long time in order to offer the users the opportunity to go back in time and compare current farm conditions with those of the past.
Data security	All data generated by the platform will be saved on the DIANA server. SSL connections will be used to ensure secure exchange of information. In case of necessary updates, the old data will be overwritten and all actions will be audited in detail. A log will be kept, containing the changed text for security reasons. Daily backups for a period of 3 days will be kept. All backups will be hosted on a remote server to avoid disaster scenarios. All servers will be hosted behind firewalls inspecting all incoming requests against known vulnerabilities such as SQL injection, cookie tampering and cross-site scripting.

	Finally, IP restriction will enforce the secure storage of data.
Ethical aspects	There are no ethical aspects related to the described dataset.
Other issues	N/A

DMP Component	Issues to be addressed
Data_WP3_4_Website content	
Data Summary	DIANA users will generate data via the platform. These data apart from the users' personal information contain the administrative boundaries and settlements of each area, maps of irrigated and non-irrigated areas and soil maps. The data described above will be saved in the DIANA central database or as tiff files on the DIANA server. Detailed log of user actions (login, logout, account creation, visits on specific parts of the app) will be kept in the form of a text file. This log will be useful for debugging purposes. Reports containing information on user devices (which browsers and mobile phones) as well as number of mobile downloads (taken from play store for android downloads and app store for mac downloads) will be useful for marketing and exploitation purposes, as well as decisions regarding the supported browsers and operating systems.
Making data findable, including provisions for metadata	As part of every user action we will produce meaningful metadata (time and date of data creation or data amendments, owners of actions that took place, service that produced the map). Metadata will assist the discoverability of the data and related information. The administrator of the platform will be the only one with the ability to discover all data generated by the platform.

Making data openly accessible	Data will only be available to registered users and administrators. The data produced by the platform cannot be shared with others without the user's permission. No open data will be created as part of DIANA The database is only accessible by the authorized technical team.
Making data interoperable	N/A
Increase data re-use	N/A
Allocation of resources	N/A
Data security	Platform generated data will be saved on the DIANA database server. Sensitive user data (emails, passwords) are encrypted using strong algorithms. SSL connections will be used in order to ensure secure exchange of information. In case of necessary updates, the old data will be overwritten and all actions will be audited in detail. A log will be kept, containing the changed text for security reasons. Daily backups for a period of 3 days will be kept. All backups will be hosted on a remote server to avoid disaster scenarios. All servers will be hosted behind firewalls inspecting all incoming requests against known vulnerabilities such as SQL injection, cookie tampering and cross-site scripting. Finally, IP restriction will enforce the secure storage of data.
Ethical aspects	There are no ethical aspects related to the described dataset.
Other issues	N/A

DMP Components in WP4 – Pilot deployment, monitoring and co-evaluation

DMP Component	Issues to be addressed

Data_WP4_2_Pilot deployment	
Data Summary	In situ data will be collected in order to keep track of the performance and impact of the platform and its services. The data will include time series of precipitation, map of evapotranspiration, land use map, crop phenology, etc. The data will be collected by the pilot partners. The data will be available in *.xls and *.csv formats and will be used to validate the existing models and methodologies.
Making data findable, including provisions for metadata	The data will be locally hosted at DIANA server. For each dataset, related metadata will describe data structure and methodology used to collect the data.
Making data openly accessible	All DIANA in situ data will be locally hosted at DIANA server.
Making data interoperable	The data will be provided in commonly used physical units, BBHC scale of phenological development stages of a crop.
Increase data re-use	The data will be available for re-use as soon as the quality is approved. No time limits for re-use of the data will be imposed.
Allocation of resources	N/A
Data security	N/A
Ethical aspects	No personal data will be distributed within the described datasets.
Other issues	N/A

DMP Component	Issues to be addressed
Data_WP4_3_Co-evaluation and validation of DIANA	
Data Summary	The aim of the data collection is to effectively manage ambiguity during the pilot deployment period and lead to meaningful, demand-driven improvements. The accuracy of DIANA services will be validated based on its own validation methodology. The following data formats will be produced:

	• *.csv • *.xlsx These data files will be useful mainly to the DIANA project partners and DIANA Advisory Board members for evaluation purposes.
Making data findable, including provisions for metadata	Task 4.3 will produce results of a blend of quantitative and qualitative techniques for validation of DIANA services.
Making data openly accessible	The validation results will be documented in reports along with recommendations about adjustments and will be made publicly available on the DIANA website; the results of the validation will be possibly published in peer-reviewed scientific journals. Public access of the raw data files will be determined later in the project life according to the business model which will be decided. DIANA partners will have full access over the raw data by authenticating themselves into the platform.
Making data interoperable	N/A
Increase data re-use	N/A
Allocation of resources	N/A
Data security	The private data (raw data) will be placed in a password area on the website or on DIANA Wiki, where only the project members can have access. The reports will be public, so there is no need for security measurements.
Ethical aspects	N/A
Other issues	N/A

DMP Components in WP5 – Business planning and innovation management

DMP Component	Issues to be addressed
Data Summary	**Market Intelligence Data**

	The purpose of the Market Intelligent Dataset is to include, document and manage all the data that will be available and/or come out of the market analysis exercise during the DIANA project. The data will have the following formats: *.xlsx, *.doc, *pdf files. Market reports and market data related to the pilots' regions and countries will be in *.xlsx, *.doc, *pdf format. In order to make the results of the market analysis more visible, a series of meta-data tables and figures will be produced in similar forms to feed the drafting of Business Modelling, Business Planning and Exploitation tasks. Thus, all deliverables of type report related to WP5 are expected to include data and meta-data of this kind. These documents will be in *.doc and/or *.pdf format. Existing data and reports in partners' repositories may also be used or reused. Generated meta-data will be produced only within the time duration in the context of DIANA project as described in the GA and DoA. The total file of this dataset will be approximately 1 Gb. This will include text and .xlsx files. These data would be useful to anyone who would like to reflect on the market intelligence in the domains relevant to the DIANA project. **Contact details of external stakeholders** The purpose of the Databases is to manage and document all the data that will be produced regarding the external stakeholders that will be contacted and interviewed within the context of WP5 of the DIANA project. More specifically external stakeholders will include: Potential lead users / early adopters from the Advisory Board of the project as well as other similar ones during

	the pilot operation of the DIANA service platform (based on suggestions made by partners). Interviews will aim either to validate and improve DIANA business models or establish direct contacts and relationships that may result in a more pro-active adoption of DIANA value propositions at an early stage but its target groups. The data will have the following formats: *.xlsx, *.doc, *pdf files. Reports related to the external stakeholders' interviews will also be in *.doc format. In order to make the results of the stakeholders' interviews more visible, a series of *.xlsx format tables and figures will be produced mainly with coded recommendations. An *.xls table will also be created to include the list of the contact details of the external stakeholders contacted and interviewed. Thus, all deliverables of type report related to WP5 are expected to include data and meta-data of this kind. These documents will be in *.doc and/or *.pdf format. No existing data will be used or reused. Data will be produced only within the time duration in the context of DIANA project as described in the GA and DoA. The total file of this dataset will be approximately 1 Gb. This will include text and .xlsx files. These data would be useful to anyone who would like to reflect on novel ideas and recommendations for business modeling and value propositions of services in domains relevant to the DIANA project.
Making data findable, including provisions for metadata	A series of market intelligence meta-data tables and figures will be produced to feed the drafting of Business Modelling, Business Planning and Exploitation. Unique and persistent identifiers will not be used for this dataset. Keyword search will be based on geographical and typological terms.

	Versioning will follow the versioning approach defined by the project.
Making data openly accessible	The datasets will not be publicly available. They will only be accessible through the DIANA Dropbox/ wiki and only the members of the consortium will have access to that material. The administration of the DIANA wiki will only be accessible by the Coordinator (AGROAPPS) of DIANA and the databases will be renewed when new data will be available.
Making data interoperable	N/A
Increase data re-use	The dataset cannot be accessed and there is no permission related with this dataset. Final data will be publicly available within the related deliverables deposited in the project website. Any individual or third party interested to access and reuse the data can download the deliverable from the project website in *.pdf format.
Allocation of resources	All costs related to the WP5 data collection and processing are covered by the project budget with dedicated person months under the WP5.
Data security	The WP5 market data will be preserved and shared with the members of the consortium through the DIANA Dropbox and/or Wiki. The data is collected for internal use in the project, and not intended for long-term preservation. The WP leader is keeping a quarterly backup on a separate disk. The preservation of the contact details of the external stakeholders that will be contacted and interviewed within the context of PW5 will concern the entire time of the project to facilitate communication and engagement with them. The data will be preserved and shared with the members of the consortium through the DIANA Dropbox and/ or Wiki. The data is collected for internal use in the project, and not intended for long-

	term preservation. The WP leader is keeping a quarterly backup on a separate disk.
Ethical aspects	To carry out the interviews with external stakeholders, the partners will employ those informed consent procedures that abide to ethical standards and guidelines of the HORIZON 2020 and in compliance with applicable international and national law. The anonymity and confidentiality of the research participants, as well as their right to turn-down or withdraw from the process at any point (i.e. volunteer participation), will be guaranteed and be made clear in written (consent form) as well as in oral form where possible. Any personal information will be handled according to the principles laid out by the Directive 95/46/EC of the European Parliament and of the Council on the "Protection of individuals with regard to the processing of personal data and on the free movement of such data" (24 October 1995) and its revisions, while no sensitive personal data will be collected. More specifically, the interviewees will be provided with 'information sheet', which will:

	• Inform them about for the aim of the project; • Outline the aim, method and implication of the interviews, the nature of the participation and any benefits, risks or discomfort that might be involved; • Explain the reason of their invitation to be interviewed; • Explicitly state that participation is voluntary and the interviewee has the right to refuse to participate and to withdraw his/her participation, at any time — without any consequences; • Explain that the results of the interviews will be used only for the purposes of the project; • Indicate the potential benefits of the project, such as receiving information about the outcomes of the project, the opportunity to influence evolutions in the field as well as references to related research if they are interested; • Reveal the lack of risks involved in this project other than what they would encounter in daily life; • Declare that any personal information (e.g. name, contact details, audio/video recordings, etc.) will be handled in accordance with the Directive 95/46/EC of the European Parliament and of the Council on the "Protection of individuals with regard to the processing of personal data and on the free movement of such data" (24 October 1995) and its revisions, as well as the national law on data protection;

	• Explain the need of permission for the interview to be audio-recorded and that they can revoke this permission at any time; • Clarify that the interviewer will hold all information and data collected securely and in confidence and that all efforts will be made to ensure that they cannot be identified as a participant in the interviews (except as might be required by law). To ensure that the interviewee has fully understood the scope and implications of his/her participation and does not feel pressured or forced to provide any information, he/she will be asked to give his/her consent in writing by signing the 'informed consent form'. If the consent cannot be given in writing, for example because of illiteracy, the non-written consent must be formally documented and independently witnessed. If requested, both the "information sheet" and the "informed consent form" will be provided to the interviewees in their national language, while emphasis will be given to ensure that the terms are clear and fully understandable to them.
Other issues	N/A

DMP Components in WP6 – Enabling environment and awareness raising

DMP Component	Issues to be addressed
Data Summary	Data collection is an essential procedure for the elaboration of the Enabling environment and awareness raising. Lists of communication recipients containing stakeholders and their e-mail addresses will be created and they will be in *.xls format.

	Information regarding direct/indirect competitors and data regarding water management authorities and farmers.
Making data findable, including provisions for metadata	All data will be publicly available through the DIANA website.
Making data openly accessible	Data concerning e-mail addresses will not be openly available because they are personal data.
Making data interoperable	N/A
Increase data re-use	N/A
Allocation of resources	This dataset does not require specific allocation of resources for its maintenance.
Data security	Automated backup of files.
Ethical aspects	N/A
Other issues	N/A

DMP Components in WP7 – Project Management

DMP Component	Issues to be addressed
Data Summary	**Contact details of project partners and advisory board** The purpose of the Databases is to manage and document all the data that will be produced regarding the project partners and Advisory Board members during the DIANA project. Advices and quality control, from the members, ensure the development and application of state-of-the-art methodologies, algorithms and generated services. The data will have the following formats: *.xlsx, *.doc, *pdf files. Reports related to the AB will also be in *.doc format. In order to make the results of the AB more visible, a series of *.xlsx format tables will be produced mainly with coded recommendations and their relation to specific WP, deliverable or responsible partner. An *.xls table will also be created to include the list of the contact details of the members.

	Thus, all deliverables of type report related to the project's management will be produced, such as Data Management Plan, financial reports and management reports. These documents will be in *.doc and/or *.pdf format. No existing data will be used or reused. Data will be produced only within the time duration in the context of DIANA project as described in the GA and DoA. The total file of this dataset will be approximately 1 Gb. This will include text and *.xlsx files. These data would be useful to anyone who would like to reflect on the impact that the AB would have to the evolution of the project itself.
Making data findable, including provisions for metadata	The inclusion of metadata for the current dataset has not been yet decided. Unique and persistent identifiers will not be used for this dataset.
Making data openly accessible	The datasets will not be publicly available. They will only be accessible through the DIANA Dropbox/ Wiki and only the members of the consortium will have access to that material. The administration of the DIANA wiki will only be accessible by the Coordinator (AGROAPPS) of DIANA and the databases will be renewed when new data will be available.
Making data interoperable	N/A
Increase data re-use	The dataset cannot be accessed and there is no permission related with this dataset. Final data will be publicly available within the related deliverables deposited in the project website. Any individual or third party interested to access and reuse the data can download the deliverable from the project website in *.pdf format.
Allocation of resources	All costs related to the AB data collection and processing are covered by the project budget with dedicated person months under the WP7 "Project Management".

Data security	Preserving contact details of the project partners and advisory board members for the entire time of the project will facilitate the internal communication. The data will be preserved and shared with the members of the consortium through the DIANA Dropbox and/ or Wiki. The data is collected for internal use in the project, and not intended for long-term preservation. The WP leader is keeping a quarterly backup on a separate disk.
Ethical aspects	N/A
Other issues	N/A

Conclusion

The DMP reflects the data management strategy and the procedure that DIANA will follow in order to identify issues and missing information related to data management that can be further clarified until the submission of the 2nd DMP. The DMP is not a fixed document but it will be updated twice during project lifespan (M18, M36).

Abbreviations

AB	Advisory Board
API	Application Programming Interface
BBHC	Biologische Bundesastalt, Bundessortenamt und Chemische Industrie
CSV	Comma-seperated Values
DoA	Description of Action
DMP	Data Management Plan
EC	European Commission
EO	Earth Observation
EODC	Earth Observation Data Centre
EU	European Union
FTP	File Transfer Protocol
GA	Grant Agreement
IP	Internet Provider
LAN	Local Area Network
PDF	Portable Document Format
SQL	Structured Query Language
SSL	Secure Sockets Layers
TIFF	Tagged Image File Format
VLAN	Virtual LAN
WP	Work Package
XML	Extensible Markup Language