midas

Meaningful Integration of Data, Analytics and Services

Grant Agreement No. 727721

Contract Duration: 40 months (1st November 2016 – 29th February 2020)

This project is funded by
The European Union

H2020-SC1-2016-CNECT
SC1-PM-18-2016 - Big Data Supporting Public Health Policies

Deliverable 1.7

Data Management Plan V1.0

Circulation:	PU
Nature:	R - ORDP: Open Research Data Pilot
Version #:	1.0
Issue Date:	25/04/2017
Responsible Partner(s):	University of Ulster
Author(s):	University of Ulster
Status:	Final
Reviewed on:	dd/mm/yyyy
Reviewed by:	
Contractual Date of Delivery:	30/04/2017 (M6)

Grant Agreement No: 727721

Executive Board Document Sign Off

Role	Partner	Signature	Date
Project Coordinator	Ulster	Michaela Black	24 April 2017
WP1 Lead	Ulster	Michaela Black	24 April 2017
WP2 Lead	SET	Paul Carlin	24 April 2017
WP3 Lead	VICOM	Gorka Epelde	06 April 2017
WP4 Lead	KU Leuven	Marc Claesen	12 April 2017
WP5 Lead	VTT	Juha Pajula	25 April 2017
WP6 Lead	DCU	Regina Connolly	12 April 2017
WP7 Lead	Ulster	Jonathan Wallace	24 April 2017
WP8 Lead	Ulster	Michaela Black	24 April 2017
Scientific-Technical Manager	Analytics Eng	Austin Tanney	24 April 2017

Grant Agreement No: 727721

Abstract

This deliverable is the first version of the Data Management Plan (DMP) for the MIDAS project describing datasets that will be used within the project, whether their access is controlled or open, how these will be managed, and where appropriate made available in research repositories.

Copyright

The MIDAS project is funded under the EC Horizon 2020 SC1- PMF-18 Big Data Supporting Public Health Policies

Grant Agreement No: 727721

Document History

Version	Issue Date	Stage	Content and Changes
0.1	24/03/2017	Draft	Initial draft version
0.2	04/03/2017	Draft	Initial draft reviewed by WP1 Lead
0.3	12/04/2017	Draft	Updates to datasets and content based on reviewer comments
1.0	25/04/2017	Final	Updated with further reviewer comments and final version produced

Statement of Originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Grant Agreement No: 727721

Executive Summary

Work Package:	WP1
Work Package leader:	University of Ulster
Task:	T 1.7 – Data Management Plan
Task leader:	University of Ulster

Grant Agreement No: 727721

Glossary of Terms

BIOEF	Public Foundation of the Health Department of the Basque Government
CSV	Comma Separated Variables
DMP	Data Management Plan
DOI	Document Object Identifier
EC	European Commission
EPAG	Ethical and Privacy Advisory Group
FAIR	Findable, Accessible, Interoperable and Reusable
HSCNI	Health and Social Care Northern Ireland
NHS	National Health Service
NI	Northern Ireland
PHE	Public Health England
RP	Periodic Report
RV	Review
THL	National Institute for Health and Welfare
TUDA	Trinity, Ulster, Department of Agriculture Ageing Cohort Study Data
UK	United Kingdom

Grant Agreement No: 727721

Table of Contents

Grant Agreement No: 727721

1 Introduction

This deliverable involves the development of an open data related Data Management Plan (DMP) in accordance with the **Guidelines on FAIR data management in Horizon 2020**[1],[2]. These guidelines assist beneficiaries in making their research data findable, accessible, interoperable and reusable (FAIR) to ensure it is soundly managed.

It is intended that this project will participate in the Commission's Open Research Data Pilot (ORD pilot) which aims to improve and maximise access to and re-use of research data generated by Horizon 2020 projects and takes into account the need to balance openness and protection of scientific information, commercialisation and Intellectual Property Rights (IPR), privacy concerns, security as well as data management and preservation questions

While open access to research data thereby becomes applicable by default in Horizon 2020, the Commission also recognises that **there are good reasons to keep some or even all research data generated in a project closed**. The Commission therefore provides robust opt-out possibilities at any stage.

To this end, the MIDAS project will deposit the relevant data in a research data repository, and to the maximum extent possible implement provisions for third parties to access, mine, exploit, reproduce and disseminate this data. Accompanying these measures, it is intended that the project will provide the information necessary for validating the project's results.

This Data Management Plan describes the data management life cycle for the data to be collected, processed and/or generated by a Horizon 2020 project. As part of making research data findable, accessible, interoperable and reusable (FAIR), a DMP should include information on:

- the handling of research data during & after the end of the project
- what data will be collected, processed and/or generated
- which methodology & standards will be applied
- whether data will be shared/made open access and

[1] http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[2] http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

- how data will be curated and preserved (including after the end of the project).

1.1 Opting Out – Partially or Entirely

Not all data can be open. The MIDAS project can therefore opt out at any stage (either before or after signing the grant) and so free the project retroactively from the obligations associated with the conditions – if:

- participation is incompatible with the obligation to protect results that can reasonably be expected to be commercially or industrially exploited
- participation is incompatible with the need for confidentiality in connection with security issues
- participation is incompatible with rules on protecting personal data
- participation would mean that the project's main aim might not be achieved
- the project will not generate / collect any research data or
- there are other legitimate reasons.

The Commission's approach can therefore be described as ***"as open as possible, as closed as necessary"***. During the lifetime of the project, a total opt-out is possible for any of the reasons highlighted above. Alternatively, projects can also choose to keep selected datasets or even all data closed for any of the reasons above, via the Data Management Plan.

Types of data covered by the Open Research Data Pilot:

1. the **'underlying data'** (the data needed to validate the results presented in scientific publications), including the **associated metadata** (i.e. metadata describing the research data deposited), **as soon as possible.**
2. **any other data** (for instance curated data not directly attributable to a publication, or raw data), including the **associated metadata, as specified and within the deadlines laid down in the DMP** – that is, according to the individual judgement by each project/grantee.

1.2 Data Management Plan Updates

This document is the first version of the DMP due for submission by month 6 of the project. The document uses the Horizon 2020 FAIR Data Management Plan template[3]. As per the template guidelines:

[3] http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

Grant Agreement No: 727721

"It is not required to provide detailed answers to all the questions in the first version of the DMP that needs to be submitted by month 6 of the project. Rather, the DMP is intended to be a ***living document*** *in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur. Therefore, DMPs should have a clear version number and include a timetable for updates. As a minimum, the DMP should be updated in the context of the periodic evaluation/assessment of the project. If there are no other periodic reviews envisaged within the grant agreement, an update needs to be made in time for the final review at the latest."*

The timetable for updates to the DMP are outlined in table 1.2a. Should additional updates be required throughout the life of the MIDAS project, this timetable can be updated at the point of periodic reporting.

Table 1.2a Timetable for DMP Updates

Version	Date Due	Notes
1.0	M6 - 30/04/17	First version of DMP
2.0	M18 - 30/04/18	Update 2.0 in line with D1.3 - Periodic Report 1 (RP1) (aligned to first review RV1)
3.0	M33 - 31/07/19	Update 3.0 in line with D1.4 - Periodic Report 2 (RP2) (aligned to second review RV2)
4.0	M40 - 29/02/20	Update 4.0 in line with D1.5 - Periodic Report 3 (RP3) (aligned to third review RV3)

2 Data Summary

Provide a summary of the data addressing the following issues:

- **What is the purpose of the data collection/generation and its relation to the objectives of the project?**
- **What types and formats of data will the project generate/collect?**
- **Will you reuse any existing data and how?**
- **What is the origin of the data?**
- **What is the expected size of the data?**
- **To whom might it be useful ('data utility')?**

Managing big data for 'health in all' is a monumental challenge for policy makers. MIDAS is addressing this challenge by developing and delivering an integrated

Grant Agreement No: 727721

solution which will liberate knowledge from data silos and unify heterogeneous big data sources to provide evidence-based actionable information and transform the way care is provided.

Health care authorities currently collect and store vast quantities of health care data. This data is stored in heterogeneous data silos using different standards. Currently the data is not being used to its full potential. The MIDAS project seeks to overcome this by mapping, acquiring, managing, modelling, processing and exploiting existing heterogeneous healthcare data and other governmental data. The project also seeks to overlay this with existing open and social data from various sources to provide enhanced insights into the data.

As such, the MIDAS project will use existing data from controlled and open sources to meet the objectives of the project but currently no plans exist to collect new data.

The project will however, generate new synthetic data. Synthetic data, also known as ‘artificial data’ is data that is simulated from real data using statistical models in order to represent the population yet avoid any divulgence of actual patient records. This will involve creating synthetic data from the real population datasets that are made available in this project. The synthetic datasets will be used by the MIDAS consortium in the development of the MIDAS platform as a representative of real data that avoids various governance and confidentiality issues since real patient or citizen records are not provided or disclosed. The synthetic data will also be provided to an Open Portal and thus contribute to the European Commission (EC) Open Data Initiative according to Horizon 2020 recommended practice.

“Knowledge Results” means data sets or aggregated data sets, provided by policy makers and health care providers, including the data oriented results derived from these data sets and the know-how derived from data set analysis. Due to the controlled nature of the data sets used to produce results and knowledge, such Knowledge Results will not be Open Access.

The project will generate:

- Synthetic data - created by modelling and simulating real data to produce a dataset not related to patients or citizens but that represents real data. The format is unknown at this point and will be provided in a DMP update. This data will be made available via Open Access.
- Knowledge Results - existing, controlled health care data sets and results derived from these data sets. The format is unknown at this point and will be

provided in a DMP update. Due to the controlled nature of the data sets used produce results and knowledge, such Knowledge Results will not be Open Access.

The MIDAS project will re-use existing data from controlled and open sources to meet the objectives of the project.

A list of datasets, whether they are open or controlled, country of origin, and details of the dataset Table 2a. When synthetic datasets are generated, details will appended to this table in further versions of the DMP.

Table 2a. Potential Datasets Used in MIDAS

Data Source	Open / Controlled	Origin	Description
NHS Datasets	Controlled	UK (All regions)	Public health provider data
MyData	Controlled	Finland	Individual health and wellness data
Private/ Commercial Entities Data	Controlled	Basque Country and others	Grocery selling per municipality, localisation data captured by telecom operators
European Open Data Portal[4]	Open	Europe	Various public data on a range of topics
Twitter[5] [6]	Social Open	Global	Social - Messages and metadata about messages from the public
Liveuamap[7]	Open	Global	Media data
EventRegistry[8]	Open	Global	Media data
OSAKIDETZA - National Health Datasets	Controlled	Basque Country	Public health provider data
Open Data Euskadi[9]	Open	Basque Country	Various public data from the Basque Government (health, tourist, environmental, employment, education, economy...)
National Health Datasets	Controlled	Finland	Data related to prevention of mental health in young persons, prevention and reasons for drinking and drug usage, unemployment and promotion of working options/possibilities
National Health Datasets	Controlled	UK (NI)	Data related to children at risk
National Health Datasets	Controlled	Republic of Ireland	Details to be updated in DMP V2.0
OpenDataNI	Open	UK (NI)	Various
Open Data by THL[10]	Open	Finland	Social and health care
Sotkanet[11]	Open	Finland	Statistical information on welfare and health in Finland

[4] https://www.europeandataportal.eu/
[5] https://dev.twitter.com/rest/public
[6] https://dev.twitter.com/streaming/overview
[7] http://world.liveuamap.com/
[8] http://eventregistry.org/
[9] http://opendata.euskadi.eus/
[10] https://www.thl.fi/en/web/thlfi-en/statistics/statistical-databases/open-data
[11] https://www.sotkanet.fi/sotkanet/en/index?

Grant Agreement No: 727721

Statistics Finland[12]	Open	Finland	Statistics Finland (official national statistics center)
THL health and welfare data/statistical data[13]	Controlled	Finland	Health and welfare data/statistical data
THL regional level health and welfare research data[14]	Controlled	Finland	Health and welfare research data
National Patient Data Repository, Finland[15]	Controlled	Finland	Patient healthcare data
Northern Finland Birth Cohort data[16]	Controlled	Finland	Birth cohort data
National My Kanta pages, Finland[17]	Controlled	Finland	Health records and electronic prescriptions
MyData in the City of Oulu[18]	Controlled	Finland	Individual health and wellness data for Oulu
OukaDW	Controlled	Finland	Healthcare data from City of Oulu
Fingertips[19]	Open	UK (England)	Wide range of aggregate public health data for England and English sub populations
NHS England Data Catalogue[20]	Open	UK (England)	Wide range of health and care data for England
PHE Data Gateway[21]	Open	UK (England)	Portal to range of tools, datasets on health and care in England
CancerData[22]	Controlled	UK (England)	Cancer incidence, survival and other statistics; access to raw data sets possible
Data.gov.uk[23]	Open	UK (All regions)	UK open government data site
LG Inform[24]	Part open	UK (England, Scotland, Wales)	Local government information
UK Data Archive[25]	Part open	UK (All regions)	Survey repository
Open prescribing[26]	Open	UK (England)	Reuses data on monthly prescriptions and costs by general practices in England
Global burden of disease[27]	Open	UK (England)	Comparative data on the Global Burden of disease for English regions split by deciles of deprivation
Open Data Ireland[28]	Open	Republic of Ireland	Ireland's Open data Portal
TUDA	Controlled	UK (NI) & Republic of Ireland	Trinity, Ulster, Department of Agriculture (TUDA) Ageing Cohort Study Data
Diabetes dataset - HSCNI	Controlled	UK (NI)	Northern Ireland citizens and diabetes dataset

[12] https://www.stat.fi/index_en.html
[13] http://www.terveytemme.fi/
[14] http://www.terveytemme.fi/ath/
[15] http://www.kanta.fi/en/earkisto-esittely
[16] http://www.oulu.fi/nfbc/
[17] http://www.kanta.fi/en/omakanta
[18] https://www.oulunomahoito.fi/
[19] https://fingertips.phe.org.uk
[20] https://data.england.nhs.uk
[21] https://www.gov.uk/guidance/phe-data-and-analysis-tools
[22] https://www.cancerdata.nhs.uk/
[23] https://data.gov.uk
[24] http://lginform.local.gov.uk
[25] http://data-archive.ac.uk
[26] https://openprescribing.net
[27] https://vizhub.healthdata.org/gbd-compare/england/
[28] https://data.gov.ie/data

Intego	Controlled	Belgium	Primary care records maintained by Academic Centre for General Practice at KU Leuven

The datasets in Table 2a will be used for the development of a test platform for installation within the Policy Board data gatekeeper locations (health authorities in partner countries). The test platform will then be installed within the gatekeeper premises' and applied to the controlled datasets held by the gatekeepers. This data will never leave the gatekeepers premises. The open and social datasets will be overlaid where possible with the controlled data in order to develop the MIDAS platform throughout the duration of the project.

The datasets listed in this version of the DMP may expand in later versions and additionally it may be decided that certain datasets will not be used.

The size of the data is unknown at this point but will be included in a later DMP update. University of Ulster storage will cover the data storage for the test datasets (TUDA and Diabetes/HSCNI) and synthetic datasets. Exchange of data with project members at different sites will be secure and redundant through the use of University of Ulster cloud storage.

The data described will be useful to the MIDAS consortium, and other research groups working on similar research questions to MIDAS.

3 FAIR Data

3.1 Making data findable, including provisions for metadata

- **Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?**
- **What naming conventions do you follow?**
- **Will search keywords be provided that optimize possibilities for re-use?**
- **Do you provide clear version numbers?**
- **What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.**

The MIDAS consortium intends to share any synthetic datasets produced in publicly accessible research data repositories. Initially the data will be made available via the

University of Ulster Research Data Repository. The data may also be deposited in the Registry of Research Data Repositories[29], Zenodo[30] and additional repositories as decided upon by the consortium.

These repositories assign DOIs for clear identification and citability of datasets. Additional metadata of the dataset will be offered within a separate file in a standardised way. Files will be uniquely identifiable and versioned by using a name convention consisting of project name, dataset name, method used, ID, place and date. Suitable keywords will also be added.

3.2 Making data openly accessible

- **Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.**
  Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.
- **How will the data be made accessible (e.g. by deposition in a repository)?**
- **What methods or software tools are needed to access the data?**
- **Is documentation about the software needed to access the data included?**
- **Is it possible to include the relevant software (e.g. in open source code)?**
- **Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.**
- **Have you explored appropriate arrangements with the identified repository?**
- **If there are restrictions on use, how will access be provided?**
- **Is there a need for a data access committee?**
- **Are there well described conditions for access (i.e. a machine readable license)?**
- **How will the identity of the person accessing the data be ascertained?**

[29] http://www.re3data.org/
[30] https://www.zenodo.org/

The datasets listed as “controlled” within Table 2a cannot be shared as these do not belong to the MIDAS project and the MIDAS consortium are using these as a third party. Datasets listed as “Open” are already openly available. Any synthetic datasets produced within the project will be made openly available.

Synthetic datasets will be made accessible by depositing these within the University of Ulster Research Data Repository. These datasets may also be deposited in other open research data repositories to be agreed amongst the consortium, such as the Registry of Research Data Repositories[31] and Zenodo[32].

Datasets will be deposited as CSV files that can be accessed using a spreadsheet processing, e.g. Microsoft Excel or alternative open source software packages. No additional documentation will be needed to access the software. The data, metadata and documentation will be deposited in the repositories as described.

The use of synthetic datasets will not be restricted. The MIDAS project has an Ethical and Privacy Advisory Group (EPAG) in place. Data access considerations will be monitored by the EPAG periodically.

As the data that will be provided and accessed is synthetic (and therefore anonymous), the identity of persons accessing it is not required. This policy may change in a future version of the DMP, for example, a user may need to subscribe to a repository providing basic identifying details such as name, email address, location and organisation, in order to gain access to the data.

3.3 Making data interoperable

- **Are the data produced in the project interoperable, that is allowing data exchange and reuse between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?**
- **What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?**

[31] http://www.re3data.org/
[32] https://www.zenodo.org/

- **Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?**
- **In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?**

The data, metadata and documentation will be compliant to disciplinary standards, open file formats and will use controlled vocabularies and the standard metadata schema for easy interoperability and reuse. Research Data Alliance guidance will be followed.[33]

For applicable synthetic data, further information will be provided in a later DMP update when the synthetic data has been produced. This will be used to provide metadata at various levels as needed.

3.4 Increase data reuse (through clarifying licences)

- **How will the data be licensed to permit the widest re-use possible?**
- **When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.**
- **Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.**
- **How long is it intended that the data remains re-usable?**
- **Are data quality assurance processes described?**

The data will be licensed under Creative Commons CC BY 4.0 International[34]. Applicable data will become available at the end of the project. Parts of the data may become available prior to this as a result of journal publications. There will be no embargo period.

The data can be reused by other scientists and interested parties.

The data quality is ensured by different measures. These include validation of the sample, replication, comparison with results of similar studies and control of

[33] http://rd-alliance.github.io/metadata-directory/
[34] https://creativecommons.org/licenses/by/4.0/

systematic distortion.

As open formats are used for data archiving, the data will remain re-usable until the repository withdraws the data or goes out of business.

4 Allocation of Resources

- **What are the costs for making data FAIR in your project? How will these be covered? Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).**
- **Who will be responsible for data management in your project?**
- **Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?**

The costs for data preparation to be FAIR are unknown at this stage but will be estimated in a future DMP update. Expenses may consist of additional publication and documentation costs of the repositories where applicable. Data preparation and management costs during the project will be covered by the project.

The University of Ulster, as the Project Coordinator and Data controller for MIDAS, will be responsible for data management plan updates, backup and storage of datasets held within the University of Ulster infrastructure (i.e TUDA, Diabetes/HSCNI and synthetic datasets), and data archiving and publication within repositories.

Long term preservation should not result in additional cost.

5 Data Security

- **What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?**
- **Is the data safely stored in certified repositories for long term preservation and curation?**

For the controlled datasets that the MIDAS project has access to (TUDA and Diabetes/HSCNI), i.e. data that is not held by the Policy Board gatekeepers, and any

synthetic datasets generated, the data is stored on a University of Ulster institutional server with regular backups taking place. Backups are also checked regularly. The server and backups are managed and monitored by the University of Ulster IT team. Following consultations with this team, no additional costs are expected for storage and backup at this point.

Access to the controlled datasets is provided only to technical consortium members requiring access for development purposes. Access will be provided via OpenVPN as well as HTTPS for secure data transfer. Only project members with clearance through data access agreements will access the data. Access will be granted at different access levels depending on the level of access required. Secure passwords will be used for access.

Long term preservation will be provided by the University of Ulster Research Data Repository and any additional repositories selected.

There are plans for controlled datasets originating in Finland to be provided to the MIDAS consortium via THL, City of Oulu, University of Oulu and MyData. This data needs to be cleaned and secured initially. Some of the data cannot go outside Finnish borders and therefore is planned to be located in the secure CSC virtual environment (ePouta). Planning and approval for this data is in the early stages and therefore more information will be provided in a future DMP update.

6 Ethical Aspects

- **Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).**
- **Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?**

Ethical aspects of data usage are handled within the deliverables of Work Packages 2 and 8. Ethical approval of controlled data used within the project has been granted for use by the consortium and the data will not be shared outside of the project. Open data is already accessible. There are no ethical issues surrounding the synthetic data that will be generated within the project as this data will not be linked to any persons.

Grant Agreement No: 727721

Any activity being undertaken within the MIDAS project using data whether controlled or open is reviewed and monitored by the EPAG.

7 Other Issues

- **Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?**

The University of Ulster procedures for data management and any procedures as required by the owners of controlled datasets (i.e. Policy Board).

8 APPENDIX A

Excerpt from Article 29, Section 29.3 of the Grant Agreement

ARTICLE 29 — DISSEMINATION OF RESULTS — OPEN ACCESS — VISIBILITY OF EU FUNDING

29.3 Open access to research data

Regarding the digital research data generated in the action ('data'), the beneficiaries must:

(a) deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user — the following:

(i) the data, including associated metadata, needed to validate the results presented in scientific publications as soon as possible;

(ii) other data, including associated metadata, as specified and within the deadlines laid down in the '**data management plan**' (see Annex 1 of the Grant Agreement);

(b) provide information — via the repository — about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (and — where possible — provide the tools and instruments themselves).

This does not change the obligation to protect results in Article 27, the confidentiality obligations in Article 36, the security obligations in Article 37 or the obligations to protect personal data in Article 39, all of which still apply.

As an exception, the beneficiaries do not have to ensure open access to specific parts of their research data if the achievement of the action's main objective, as described in Annex 1 (of the Grant Agreement), would be jeopardised by making those specific parts of the research data openly accessible. In this case, the data management plan must contain the reasons for not giving access.