This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 964248

Project Acronym: *DeDNAed*

Deliverable 8.2
Data Management Plan

Due date of Deliverable: 31.08.2021

Actual submission date: 31.08.21

Lead beneficiary for this Deliverable: TUC

Level of Dissemination: public

Document History

Version	Date	Edited by	Comment
1	15.08.2021	TUC, Julia Hann	1st draft
1.1	18.08.2021	TUC, Andreas Morschhauser	1st internal review
1.2	19.08.2021	TUC, Julia Hann	1st revised version
1.3	26.08.2021	BNN, Christina Pfeifer & Susanne Resch	Final review
2	30.08.2021	TUC, Julia Hann	Final version

Table of Contents

Acknowledgement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 964248

Abbreviations

AGA	Annotated Model Grant Agreement
CA	Consortium Agreement
CV	Curriculum vita
D	Deliverable
DM	Dissemination Manager
DMP	Data Management Plan
DOI	Digital Object Identifier
EC	European Commission
FAIR	Findable, Accessible, Interoperable, Re-usable
GA	Grant Agreement
GAs	General Assembly
GDPR	General Data Protection Regulation
IEM	Innovation and Exploitation Manager
IPR	Intellectual Property Right
NP	Nanoparticle
ORDP	Open Research Data Pilot
PC	Project Coordinator
PM	Project Manager
PMT	Project Management Team
PQP	Project Quality Plan
PDF	Portable Document Format
V	Version
WP	Work Package
WPL	Work Package Leader

1. Introduction

The Data Management Plan (DMP) as described in this document defines the implementation of the general data handling mechanisms of the project based on the definitions and regulations in the Grant Agreement (GA) and the Consortium Agreement (CA). The purpose of this document is to provide guidelines and principles that ensure a high scientific and organisational quality of the *DeDNAed* project throughout its lifetime.
This DMP is drawn up in English which is the language governing all documents, notices, arbitral proceedings and journal publication thereto.

As the DMP reflects dynamic processes in relation with the project environment, it will be partly updated and refined during the course of the action at least in two times during the project with a final version in M36. Any new version will be added to the project platform and website with an email notice to all partners.

2. Document Management

The Project Coordinator (PC) is responsible for the document management in the whole project. Deliverable documents to the Commission as listed in the GA and all other reports, minutes or presentations shall be based on the document templates applicable for all documents to be created within the scope of this work. In addition, it shall be noted that all documents (internal and external) will be written in English.

2.1. Project Quality Plan

The Project Quality Plan (PQP) (deliverable 1.2) details the templates to be provided for document production, the nomenclature to be followed and the intervals at which reports must be delivered to the Project Manager (PM) or the PC in order to maintain a high scientific standard. In addition, a specific procedure for evaluation/review within the consortium of deliverables for submission to the European (EU) Commission is set out. The procedure for convening project meetings or plenary meetings and their documentation as well as the veto rights of the individual partners with the corresponding deadlines to be met are also defined. All information regarding these points can be found in the publicly accessible Deliverable “1.2 - Project Quality Plan”.

2.2. Data Exchange and Communication Platform

The TUCcloud of the Chemnitz University of Technology was selected as the exchange platform for data sharing in between the consortium of DeDNAed project in order to enable the best possible security and a long-term internal data storage of project related data. TUCcloud gathers all sorts of documents generated during the project lifetime. Within the TUCcloud, a folder structure has been implemented, which provides well-arranged access to all documents. All internal and external reports, regardless of which issue or document type as well as source code, shall always be distributed by using the TUCcloud. The direct distribution of documents or source code via email is allowed.

3. Data Management Plan

To meet all rules within the given framework, DeDNAed will take the approach to inweave processes of the building up of the data management plan - as a real living document - into the tasks of the work package (WP) 8 as three deliverables (D8.2/8.4/ 8.6). There are two types of digital research data relevant for the Open Research Data Pilot (ORDP).

- On the one hand, these are the data needed to validate the results of DeDNAed in scientific publications, as well as the associated metadata (i.e. data describing the deposited research data).
- On the other hand, data and related metadata are generated or collected that are not cleared for publication as indicated by the beneficiaries themselves in their data management plan (e.g. curated data not directly attributable to a publication, or raw data), or contain business or exploitation related or proprietary information.

Therefore, the DMP and its regular updates will interact not only with all scientific tasks itself but also with the DM and the work done within WP 8 to oversee planned publications.

3.1. DMP content

The content and specific questions and guidelines is largely based on the template provided by the EU Commission [2]:

3.1.1. Data Summary

Purpose of the data collection/generation and its relation to the objectives of the project?

During the project we will collect, store and process data related to the engineering, realization and performance optimization of the DeDNAed optical biosensor and obtained from the characterization and assessment of the sensor in lab. Moreover, data relevant for the specific targeting of the sensor operability in relation to the end-users' will be provided.
Collected data are needed to:

- design, characterize, develop and optimize the DNA origami based biosensor platform and its integration in the whole biosensor assay;
- design parameters of the nanograting, (bio)functionalization procedure of the grating surfaces, its fabrication method, structural and optical characterization and immobilisation performance evaluation;
- to compare operability of the sensor with respect to state-of-the-art detection technologies;
- collect requirements and specifications for the bio-recognition element and DNA origami template in the view of the real-setting applications;
- demonstrate the use of the sensor in relevant environment and provide key parameters and knowledge allowing the uptake of the sensor into end-users' processes;
- fill in a database useful for the interpretation of analytical data generated by the sensor;
- assess the risk for health of nanoparticles integrated in the DeDNAed biosensor platform.

Relation to the objectives of the project

The data collected in the project as a whole aims to develop a biosensor platform that is:

- highly sensitive,
- highly selective,
- fast,
- easy to customise,
- simple in its applicability,

combining biological recognition elements like DNA aptamer and antibodies with DNA origami and an optical detection method.
On page 5 of annex 1 - Part B of the Grant Agreement, the necessary objectives to achieve the above are listed.

(1) Establishment of DNA origami as "nano breadboard" for recognition elements
(2) Proof of signal enhancement through spatial alignment of recognition elements
(3) Demonstration of detection of food containments and bio markers on novel sensor platform
(4) Transfer of the sensor platform to a flexible substrate

Types and formats of data generated/collected

In table 1 an overview of the generated and/or collected research data is shown. A detailed description for each partner you can find in the annex.

Reasearch data that are generated and/or collected during the project	**Accessibility**	**How data will be shared/exploited**	**Formats**
Scientific documents	Public	Public DeDNAed reports, research publications in open access journals, patent applications	Written documents (.pdf, .doc, .docx)
Experimental Data like Microscopy / Spectroscopy / Measurement Data	Public/ Confidential, shared with partners	Corresponding experimental data of public deliverables/reports or research publications will be published at least at the repository in open access; IPR relevant experimental data will be closed but shared in the consortium	Measurement data in raw format: (.csv, .xlsx, .mi, .nid, .wip, .ngs), measurement report (.docx, .pdf, .txt) , microscopy images (.tiff, .png, .bmp , .jpg, .eps, .mov , .avi)
Production of technical documents for working procedures and fabrication data	Confidential, shared with partners	Technical documents will be shared with project partners to allow the results of the research to be applied on similar problems.	Input data (GDSII, OASIS, DXF, CIF, …), report data (.txt, .xls, .doc)
Definition of specific technical data sheets for assessment of possible markets in the food safety and quality monitoring	Public	Data sheet will be shared with the commercial partners and/or stakeholders for marketing of experimented products	.pdf, .doc, .docx
Reports on progress of exploitation including analysis of drivers and barriers influencing the process	Confidential, shared with partners	Analysis and reporting on exploitation will be performed yearly by partners coordinated by the IEM/IPR Manager in order to maximize the use and effect of available results. In particular, the Exploitation Manager will manage IPR	Digital notebooks (.one), presentations (.pptx, .pdf)
Data related to Interleukin-6, Aflatoxin and special DNA probes monitoring	Public	Report on the performance of DeDNAed sensor as public deliverable	Presentations (.pptx, .pdf), written

			documents (.doc, .docx, .pdf)

Table 1: Overview of collected/generated data

Re-use of existing data

Some of the partners could reuse existing technology or data to compare the new sensing system against existing solutions or provide some know-how solutions.

Origin of data

The research data will be collected during the experimental activity in the phase of the engineering and optimization of the sensor platform and its integration on a solid or flexible substrate, reused from other partners, or collected from legislative and scientific sources, etc. The data will be either generated via Personal computers and corresponding software for design, measurement and office applications or directly by used measurement equipment (microscopes, spectrometers, etc.) and equipment during the wafer processing.

Expected data size

The size of generated/collected data differs between the partners and depends on the used methods and number of measurements. Biggest size will be expected by design and CAD-files. For most of the partners the size of data will not exceed 50 GB. However, there may be some single exceptions that could overshoot these standards – even up to 100-1000GB – but this should only concern raw data.

Data utility: To whom it will be useful

- Generated data are beneficial to the DeDNAed project, end-users, and the scientific community in general.
- Data are essential for partners of the DeDNAed consortium to develop the several components of the sensor and the final sensor platform.
- Data are useful for scientific production (e.g. publications, patents, analysis, impact, assessment, validations, etc.), that are relevant to end-users, stakeholders and the scientific community.
- CAD data and layout designs of the sensor platform and its components are of interest in terms of interface definition between the partners and a coherent overall system design.
- Experimental data are needed for evaluation of the expected performance of the DNA origami hybrid and the sensor platform and for counter-checking the methods and results among partners.
- Raw data (.mov and .tiff) could be shared, if other partners have certain software or know-how of data analysing, which the original partner has not.
- The final data about the evaluation of DeDNAed use will be beneficial for policy makers, risk managers and risk assessors.

3.1.2. FAIR

Making data findable, including provisions for metadata

All the generated and collected data during the project will be provided with keywords, a DOI (Digital Object Identifier) and with relevant metadata. There are 2 main scenarios.

- The first is data collected and processed from the sensor evaluation. These datasets will be annotated with information describing data source (in case of evaluation: time, location, use-case partner etc.), actual content of the data (size, format, type of information etc.) and optionally keywords providing additional context to the dataset.
- The second one is data produced for dissemination (i.e., diagrams, pictures, graphs, reports, etc.) resulting from the engineering of the sensor and from research activities. These will be annotated with relevant keywords characterizing nature of the data.

Naming convention for the shared and published research documents will follow the general naming convention, that is already defined in D1.2 the Project Quality Plan for deliverables and reports:

1. first include the project acronym
2. be descriptive of the contents of the document/file
3. indicate the date (dd_mm_yyyy) of issue or of reference of the document/file, if necessary
4. include the short name of the partner that has created/last modified the document/file
5. include, at the end, the version number –version numbers should start at 1.0 and be incremented just by the author of the document by 0.1 for minor revisions, and by 1.0 for major revisions/new releases.

In the first version of the DMP of the DeDNAed project, a precise standard for the enclosed meta data has not been established yet, but it will include the following terms:

- Data set nature
- Scale of data set
- Information on existence of similar data sets
- Possibilities for integration
- Search keywords

Any dissemination of results (in any form, including electronic) must:

a) display the EU emblem and
b) include the following text:

"*This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 964248*".

c) must indicate "that it reflects only the author's view and that the [Commission][Agency] is not responsible for any use that may be made of the information it contains.

Making data openly accessible

Reference document: GA, page 47-49, H2020 AGA — Annotated Model Grant Agreement (V5.2 – 26.06.2019), pages 245-254.

Article 29.2 sets out the rules for the open access on digital research data. Beneficiaries of actions participating in the Open Research Data Pilot (ORDP) must give open, free-of-charge access to the end-user to digital research data generated during the action. This includes data statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images. The ORDP applies to 2 types of digital research data:

- Primary application: the data needed to validate the results presented in scientific publications and associated metadata (i.e. data describing the deposited research data) and
- Secondary application (on voluntary basis): Other data and associated metadata, as specified by the beneficiaries themselves in their data management plan (e.g. curated data not directly attributable to a publication or raw data).

Open access to research data means taking measures to make it possible for third parties to access, print, exploit, reproduce and disseminate data — via a research data repository. A 'research data repository' means an online archive for research data; this can be subject-based/thematic, institutional or centralized.
Scientific publications and public deliverables of DeDNAed project will be made openly available by the repository platform Zenodo.org. Associated metadata and documentation will be deposited together with the files on Zenodo.org There is also the possibility to store research publications (publicly searchable) in the OpenAire approved institutional repository of each partner. Some data will contain sensitive information (closed reports, device layouts, microscopy data) or will contain business- or exploitation- or IPR-sensitive information. To protect IPR and personal data, no access will be given to restricted/closed data. Such closed data can be shared in the consortium through the project collaborative platform (TUCcloud) hosted by TUC. The TUCcloud is the central cloud storage in the data centre of the University Computing Centre (URZ) of the TUC and is based on the *Nextcloud* software (https://www.tu-chemnitz.de/urz/storage/cloud/index.html). The access to the TUCcloud is password secured and only possible by invitation email.

Zenodo was established in 2013 as a collaboration between CERN and the OpenAIRE project, which was commissioned by the EC to support their Open Data policy by providing a catch-all repository for EC funded research (https://www.openaire.eu/). Since then, it has welcomed research from all over the world and from every discipline.

Zenodo supports the sharing, curation and publication of data and software. It assigns all publicly available uploads a DOI to make the data set easily and uniquely citeable. Zenodo supports making the data findable by allowing harvesting of all content via the OAI-PMH protocol; citation information is also passed to DataCite and onto the scholarly aggregators. Furthermore, OpenAIRE integrates into existing reporting lines to the EU Commission, allowing the funding body to be notified when data is shared. Finally, Zenodo allows the creation of collections, which could be used to increase the cross-visibility of data sets created within the DeDNAed project. All these features make it an ideal choice in order to gain the desired visibility and accessibility for DeDNAed data.

Further, every publication will be published on the project website *DeDNAed*.eu and promoted on social media (Twitter and LinkedIn).

Making data interoperable

The project will follow the standards and formats of the informatics community. Research data will be made available in PDF/A format to ensure the documents' portability across systems in a long-term perspective. Closed data for internal DeDNAed use will be stored and shared in a defined folder hierarchy using TUCcloud instance hosted by TUC. Access will be granted only to DeDNAed members.

Increasing the data re-use (clarifying licences)

All public research data should be published under a Creative Commons license wherever possible for re-use, and a DOI should be assigned. Peer-reviewed publications will be made available based on journal requirements. Data associated with research publications (including layouts and microscopy data, unless subject to intellectual property protection) are manually managed by the creating

researchers before being uploaded to Zenodo as a public repository. This happens after the research publications have been published in institutional repositories. Other published outcomes (e.g. project publications, posters, presentations, public deliverables, etc.) will be uploaded to the project website and/or shared file server and assigned a DOI through Zenodo (https://zenodo.org/), where a DeDNAed project community is established (https://zenodo.org/communities/dednaed/?page=1&size=20). The metadata in the institutional repositories are curated by subject librarians to ensure sufficient quality and discoverability.

Research data that contributed to a PhD thesis need to be stored for at least ten years but will try to make the data of DeDNAed available for as long as it is technically possible.

3.1.3. Allocation of Resources

By using public repositories whenever possible, there should be no further costs for the public dissemination of the data.

Scientific research should also be conducted through green or gold open access whenever possible. The costs for such a publication are covered by the DeDNAed project.

Possible costs for the long-term management of the TUCcloud data platform have to be estimated.

3.1.4. Data Security

- Zenodo manages secure storage and data recovery of published data;
- TU Chemnitz manages TUCcloud secure storage and data recovery/back-up of closed but consortium shared data;
- Additional data storage via self-archiving on partners institute servers with secure storage and data recovery.

3.1.5. Ethical aspects

There are no ethical concerns regarding data being processed within the project.

3.2. Updates of the Data Management Plan

The DMP will be updated over the course of the project whenever significant changes arise (new data, new innovation potential, patents, new consortium members, etc.). As a minimum, the DMP will be updated in accordance with the periodic evaluation/assessment of the project as D 8.4 and D8.6. The GA meetings will allow to monitor continuously whether any update of the DMP is necessary in between those periods.

3.3. Protection of Personal Data

Reference document: GA, page 60-61, 298-299; H2020 AGA — Annotated Model Grant Agreement (V4.1 – 26.10.2017)[1], pages 285-288; CA, page: 23-25.

Article 39, explicitly 39.2, sets out the rules regarding the processing of personal data (e.g. name, address, identification number, e-mail, CV, bank account number, phone number, medical records) by the beneficiaries. The processing of this personal data means any operation (or set of operations) which is performed on personal data that includes its

- collection

- recording
- organisation and storage
- adaptation or alteration
- retrieval and consultation
- use
- disclosure by transmission, dissemination or otherwise making available
- alignment or combination
- blocking, deleting or destruction.

Some potentially sensitive data (meaning e.g. racial or ethnic origin) will be collected in the research process, due to the nature of the technology developed.

This chapter sets out the framework for the exchange of personal data between the parties, which alone determines the purposes and means of the processing of personal data as controllers on the basis of Chapter 12 of the CA on the protection of personal data, pages 23-25. It sets out the principles and procedures to be followed by the parties and the responsibilities that the parties owe to each other without being joint controllers within the meaning of Article 26 of the GDPR.

Accordingly, each partner acts as a data controller and is responsible and liable for dealing with any data subject requests under Articles 12 to 23 of the GDPR. However, each other party shall provide its reasonable assistance in dealing with such requests. Further, the parties ensure that they not transfer personal data to third country without complying with Applicable Data Protection Laws.

When collecting (Data Controller) and processing personal data (Data Processor), the parties agree to implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk pursuant to Section 32 of the GDPR. For the avoidance of doubt, personal data shall always be treated as confidential, and shall be protected with an adequate level of safety and confidentiality, subject to any applicable legal, regulatory or contractual requirements.

If a party comes to the conclusion that processing of data is to be carried out on behalf of a controller (Article 28 GDPR - Processor) or two or more parties jointly determine the purposes and means of processing (Article 26 GDPR - Joint controllers) and therefore they require a further contractual agreement in addition to the above-mentioned section, such party shall notify the parties without delay and all affected parties shall undertake to establish such agreements without undue delay. The agreements must be especially in accordance with Article 26 or Article 28 GDPR and must among other things but not limited to, for example the duty of the processor at the choice of the controller, to delete or return all the personal data to the controller after the end of the provision of services relating to processing, and deletes existing copies unless Union or Member State law requires storage of the personal data.

Otherwise, Consortium Parties agree that the data:

(a) is to be used only for the research purposes as described in the Project;
(b) will not be used for commercial purposes; and
(c) will not be disclosed to any Third Party.

4. References & Bibliography

[1] https://ec.europa.eu/research/participants/data/ref/h2020/other/gm/reporting/h2020-tpl-oa-data-mgt-plan-annotated_en.pdf; access 09.08.21

[2] https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf, access 09.08.21

5. Annex

Partner	Question	Answer
TUC	Data Collection	
	The purpose of the data collection/ generation and its relation to the objectives of the project	Data is collected for the following purposes: - internal management of the project - collection of requirements and specifications for the processing of the thin-film system - Optimization of immobilization quality of DNA origami hybrids - giving data background for reports - sharing data between the partners - - giving comprehensive data background for combined publications
	Types and formats of data that the project will generate/ collect	The nature of the data is largely experimental, apart for the activity correlated to the project management and dissemination (which are of relevant entity in the action of TUC as project coordinator). Data formats are: - Written documents and reports (.doc, .docx, .txt) - Measurement data in raw format (.csv, .xlsx, .mi, .nid) or report data format (.docx, .pdf) - microscopy images and pictures raw and processed data (.tiff, .png, .eps, .jpg) - movie data raw and compressed (.mov., .avi)) - descriptions/ notes(.txt) - presentation material (.pptx, .pdf) - publications (.pdf) - - digital notebooks (.one)
	Reuse of existing data	A part of the work is based on previous know-how and published results in regard to the thin-film system
	Origin of the data	The data will be either generated via PCs and corresponding software for design, measurement and office applications or directly by measurement equipment such as microscopes, measurement and process equipment and so on.
	Expected size of the data	The overall amount of data generated is expected to be less than 50GB. - .mov's and tiff's are of big sizes (each .mov about 500MB and each .tiff of about 20MB) - single CAD-assembly folders or layout-files can have up to 2GB - other files are of irrelevant small size (<5 MB)
	Data utility	CAD/layout-data is of interest in terms of interface definition between the partners and a coherent immobilization system design - experimental data for cross-checking of method and result - raw data (.mov and .tiff) are sometimes reasonable to share, if other partners have certain software or know-how of analysing the data, which the original partner has not. But it will not be of high necessity - publications are relevant for end-users, policy makers and the scientific community
	Ethics and Data Protection	
	How is it planned to obtain informed consent and secure confidentiality	Communication and administrative and research activities will follow the obligations/agreements specified in the Grant Agreement and the Consortium Agreement; the TUC cloud for internal data sharing is password secured and access is provided by email invitation only. On the DeDNAed website, a Privacy Policy is included and users are asked

		to actively consent the use of their data via the contact form.
	What ethical issues should be answered?	n.a.
	Is there a need to anonymize data during the research period or before data can be shared?	n.a.

Partner	Question	Answer
CICB	Data Collection	
	The purpose of the data collection/ generation and its relation to the objectives of the project	Data is collected for the following purposes: - internal management of the project - collection of requirements and specifications for the design and synthesis of DNA aptamers - collection of specifications for the synthesis of metallic nanoparticles and nanoclusters - collection of data obtained during characterization of DNA aptamers, metallic nanoparticles, and nanoclusters - giving data background for reports - sharing data between the partners - giving comprehensive data background for combined publications
	Types and formats of data that the project will generate/ collect	The nature of the data is largely experimental, apart for the activity correlated to the project management and dissemination Data formats are: - Written documents and reports (.doc, .docx, .txt) - Measurement data in raw format (.csv, .xlsx, .mi, .nid) or report data format (.docx, .pdf) - microscopy images and pictures raw and processed data (.tiff, .png, .eps, .jpg) - movie data raw and compressed (.mov., .avi)) - descriptions/ notes(.txt) - presentation material (.pptx, .pdf) - publications (.pdf) - digital notebooks (.one)
	Reuse of existing data	A part of the work is based on previous know-how and published results in regard to the design and synthesis of DNA aptamers, metallic nanoparticles, and nanoclusters
	Origin of the data	The data will be either generated via PCs and corresponding software for design, measurement and office applications or directly by measurement equipment such as microscopes, measurement and process equipment and so on.
	Expected size of the data	The overall amount of data generated is expected to be less than 50GB. - .mov's and tiff's are of big sizes (each .mov about 500MB and each .tiff of about 20MB) - other files are of irrelevant small size (<5 MB)
	Data utility	Data on DNA sequence and shape of metallic nanoparticles and nanoclusters is of interest for CICB and partners for preparation of DNA origami - experimental data for cross-checking of method and result - raw data (.docx, .txt, .pdf,.tiff, .jpg) are reasonable to share with other partners if needed, to analyse the features of resulting DNA aptamers, metallic nanoparticles, and nanoclusters. This is not

		expected to be of high necessity. - publications are relevant for end-users, policy makers and the scientific community
	Ethics and Data Protection	
	How is it planned to obtain informed consent and secure confidentiality	n. a.
	What ethical issues should be answered?	n. a.
	Is there a need to anonymize data during the research period or before data can be shared?	n. a.

Partner	Question	Answer
KSI	Data Collection	
	The purpose of the data collection/ generation and its relation to the objectives of the project	Data is collected/generated for the following purposes: - internal management of the project - design of DNA sequences for optimal DNA hybridization - characterization of DNA origami, gold nanoparticles and hybrids thereof - optimization of synthesis and purification protocols - giving data background for reports - sharing data between the partners - giving comprehensive data background for combined publications
	Types and formats of data that the project will generate/ collect	Data formats are: - Written documents and reports (.doc, .docx, .txt) - Measurement data in raw and processed formats (.csv, .xlsx, .tiff, .txt, .ibw) - Simulation data (.txt, .json, .mb) - presentation material (.pptx, .pdf) - publications (.doc, .docx, .pdf)
	Reuse of existing data	A part of the work is based on previous know-how and published results.
	Origin of the data	The data will be either generated via PCs and corresponding software for design, measurement and office applications or directly by measurement equipment such as microscopes.
	Expected size of the data	- maya binaries are about 50 MB big (about 20 of them will be needed) - tiffs are about 5 MB big and about 200 will be generated - microscopy files .ibw are about 5 MB each and about 100 will be generated - other files are of irrelevant small size (<5 MB)
	Data utility	- maya binaries are used for simulation and graphics production - experimental data (.tiff and .ibw) for characterization of DNA origami, nanoparticles and hybrids - publications are relevant for end-users, policy makers and the scientific community
	Ethics and Data Protection	
	How is it planned to obtain informed consent and secure confidentiality	The generated data will be stored on a file server in the institute and only persons involved in the project have access to this storage.

	What ethical issues should be answered?	no
	Is there a need to anonymize data during the research period or before data can be shared?	no

Partner	Question	Answer
UM	Data Collection	
	The purpose of the data collection/ generation and its relation to the objectives of the project	Data is collected for the following purposes: - collection of the results and analysis of the data for the project - characterisation of the DNA origami systems and of the sensing elements - determination of the sensing performances and optimisation of the system - giving data background for reports - sharing data between the partners - giving comprehensive data background for combined publications
	Types and formats of data that the project will generate/ collect	Data formats are: - written documents and reports (.doc, .docx, .txt) - measurement data in raw format (.wip, .xlsx, .txt, .ngs) or report data format (.docx, .pdf) - microscopy images and pictures raw and processed data (.tiff, .png, .jpg) - descriptions/ notes(.txt) - presentation material (.pptx, .pdf) - publications (.docx, .pdf)
	Reuse of existing data	A part of the work is based on previous know-how and published results in regard to the characterisation of DNA structures using Raman spectroscopy and SERS and to the SERS sensors already development in UM.
	Origin of the data	The data will be either generated via PCs and Mac and corresponding software for design, measurement and office applications or directly by measurement equipment such as Raman spectrometers.
	Expected size of the data	The size of the whole data generated by the UM partner is estimated to be around 100GB divided in around 50GB of raw data (Raman spectroscopy data…) and 50 GB of analysed data (reports, presentations…)
	Data utility	- Raw data for sharing information and cross-analysis with other partners - Analysed data for partner exchanges within the WP and with the project coordinator for work program definition or deliverables writing - Results and analysis exchanges for publications and consortium meeting
	Ethics and Data Protection	
	How is it planned to obtain informed consent and secure confidentiality	N.A.

	What ethical issues should be answered?	N.A.
	Is there a need to anonymize data during the research period or before data can be shared?	N.A.

Partner	Question	Answer
UP		Data Collection
	The purpose of the data collection/ generation and its relation to the objectives of the project	Data is collected for the following purposes: - internal management of the project - collection of requirements and specifications for the processing of the thin-film system - Optimization of immobilization quality of DNA origami hybrids - giving data background for reports - sharing data between the partners - - giving comprehensive data background for combined publications
	Types and formats of data that the project will generate/ collect	The nature of the data is largely experimental, apart for the activity correlated to the project management and dissemination (which are of relevant entity in the action of TUC as project coordinator). Data formats are: - Written documents and reports (.doc, .docx, .txt) - Measurement data in raw format (.csv, .xlsx, .mi, .nid) or report data format (.docx, .pdf) - microscopy images and pictures raw and processed data (.tiff, .png, .eps, .jpg) - movie data raw and compressed (.mov., .avi)) - descriptions/ notes(.txt) - presentation material (.pptx, .pdf) - publications (.pdf) - - digital notebooks (.one)
	Reuse of existing data	A part of the work is based on previous know-how and published results in regard to the thin-film system
	Origin of the data	The data will be either generated via PCs and corresponding software for design, measurement and office applications or directly by measurement equipment such as microscopes, measurement and process equipment and so on.
	Expected size of the data	The overall amount of data generated is expected to be less than 50GB. - .mov's and tiff's are of big sizes (each .mov about 500MB and each .tiff of about 20MB) - single CAD-assembly folders or layout-files can have up to 2GB - other files are of irrelevant small size (<5 MB)
	Data utility	CAD/layout-data is of interest in terms of interface definition between the partners and a coherent immobilization system design - experimental data for cross-checking of method and result - raw data (.mov and .tiff) are sometimes reasonable to share, if other partners have certain software or know-how of analysing the data, which the original partner has not. But it will not be of high necessity - publications are relevant for end-users, policy makers and the

		scientific community
	Ethics and Data Protection	
	How is it planned to obtain informed consent and secure confidentiality	The generated data will be stored on a file server in the institute and only persons involved in the project have access to this storage.
	What ethical issues should be answered?	n.a.
	Is there a need to anonymize data during the research period or before data can be shared?	n.a.

Partner	Question	Answer
TEC	Data Collection	
	The purpose of the data collection/ generation and its relation to the objectives of the project	The primary data will be gathered through experimental methods done during the research project by different formats, mainly excel files with optical measurements.
	Types and formats of data that the project will generate/ collect	Data formats will be generated, maintained, and made available in a standardized way to ensure the data would be easy to understand, reuse and interoperate among different parties who are interested in utilizing them (XLSX, DOC, PDF and PPT.)
	Reuse of existing data	Some of the data from previous studies performed related to biofunctionalization of surfaces and to atomic cluster integrated in antibodies can be used as reference in this project.
	Origin of the data	in-house scientific measurement
	Expected size of the data	The overall amount of data during a comparable EU-funded project with 3 years runtime can be estimated to be around an overall data volume of 15 GB.
	Data utility	Obtained data will be useful to the partners that are involved directly in the detection of the analytes. It will be also useful to choose which material/strategy should be the best for the system.
	Ethics and Data Protection	
	How is it planned to obtain informed consent and secure confidentiality	Data should be stored in at least two different locations to avoid data lost. Data should be encrypted whenever necessary (e.g. confidentiality issues). The use of USB flash drives should be limited. Every worker in TECNALIA has his/her own password-protected user account to access the systems. The password must satisfy complexity requirements and shall be changed every 90 days. The access to networks folders and programs where information is stored/managed depends on user permissions which are decided by factors such as division, role in the company, role in the project, etc. The permissions are managed by administrators only and must be asked by authorized persons through authorized channels. TECNALIA has two-level backup. The first level is the system “previous versions” service that allows a user to recover a copy of the work (5 copies a day, two weeks period) by his/her own. Moreover, every day TECNALIA makes full backup of the working information. There are

		daily, weekly, monthly and yearly copies. The recover from this backup requires a formal procedure. In Spain, the protection of personal data and guarantee of digital rights is based on the Organic Law 3/2018, published in the official gazette no. 294, on the 6 of December 2018. This Law adapts Spanish law to the General Data Protection Regulation (GDPR) and introduces novelties through the development of certain matters contained in the GDPR.
	What ethical issues should be answered?	N/A
	Is there a need to anonymize data during the research period or before data can be shared?	N/A

Partner	Question	Answer
BNN	Data Collection	
	The purpose of the data collection/ generation and its relation to the objectives of the project	Data is collected/generated for the following purposes: - communication and dissemination of the DeDNAed project activities and results, - creation of a SbD concept and recommendations.
	Types and formats of data that the project will generate/ collect	The nature of the data is mostly written reports, lists for project-internal monitoring and coordination, as well as graphics and related visual material (project branding and related templates) used for communication and dissemination purposes. Data formats are: - Written documents and reports (.doc, .docx, .txt, .xlsx, .pdf), - presentation material (.pptx, .pdf), - publications (.pdf), - graphics (.png, .jpg, .pdf, .eps, .svg), - project website (https://dednaed.eu/).
	Reuse of existing data	For communication and dissemination, agreed relevant input will be provided from all project partners.
	Origin of the data	The data will be generated via PCs and corresponding software for design and office applications. User data from the website is tracked via the WordPress plugin “WP Statistics”. Social Media data is tracked from the DeDNAed Twitter and LinkedIn channels.
	Expected size of the data	The overall amount of data generated is expected to be less than 10GB.
	Data utility	Communication and dissemination of Consortium-agreed relevant data will be of interest for all stakeholder groups (i.e., academia, industry, regulators and policy makers, general public). The DeDNAed brand guidelines define, how to use the DeDNAed brand and its components such as graphics and images.
	Ethics and Data Protection	
	How is it planned to obtain informed consent and secure confidentiality	Communication and dissemination activities will follow the obligations/agreements specified in the Grant Agreement and the Consortium Agreement; on the DeDNAed website, a Privacy Policy is included and users are asked to actively consent the use of their data via the contact form.

	What ethical issues should be answered?	n.a.
	Is there a need to anonymize data during the research period or before data can be shared?	n.a.