The ADLIFE Project has received funding from the European Union under the Horizon 2020 Programme, grant reference number 875209.

Horizon 2020
European Union Funding
for Research & Innovation

D1.1 Data Management Plan

Deliverable No.	D1.1 update 2023	**Due Date**	30/06/2023
Description	This document is the updated version of the Data Management Plan of the ADLIFE project. It presents the implemented data processing and flows, and specifies the intentions of the project towards open research asset access after the project.		
Type	Report	**Dissemination Level**	CO
Work Package No.	WP1	**Work Package Title**	Co-ordination and management
Version	1.04	**Status**	Draft

Authors

Name and surname	Partner name	e-mail
Dipak Kalra	i~HD	dipak.kalra@i-hd.eu

History

Date	Version	Change
14/06/2023	1.01	Initial content for all sections
20/06/2023	1.02	First updating of all sections except not the yellow highlight portions of section 2.1 which need technical partner review
26/06/2023	1.03	Incorporating updates from technical partners and KG
26/06/2023	1.04	Final edits

Key data

Keywords	Data management, data protection, open research data
Lead Editor	Dipak Kalra
Internal Reviewer(s)	

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This deliverable contains original unpublished work except where clearly indicated otherwise.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.

Abstract

This document is the second version of the Data Management Plan of the ADLIFE project. It updates the first version, Deliverable 1.1 that was submitted in August 2020. The first version indicated the project intent, and this version now reflects the implemented reality of how the project will acquire and collect health data via its pilot sites through use of the ADLIFE implemented solution, and collect evidence of its benefit to patients and health systems through a large-scale pilot and clinical trial.

The first part of this deliverable presents the data processing and flows that will take place within each of the pilot sites, to identify patients who match the eligibility criteria for the ADLIFE study, and then how these data will be processed including pseudonymisation and anonymisation steps, before being made available for wider consortium use. This wider use has already included the design and development of technical components, the training and validation of artificial intelligence algorithms. During the pilot studies data will be used for the generation of evidence of health outcomes and health economic impact from use of the ADLIFE solutions. That chapter is like a data-oriented summary of the clinical research protocol, which was submitted as deliverable 11.1 and has since been further refined.

There are other project deliverables that specify the information governance, data protection and information security policies and measures to be adopted. Information security measures and guidelines on pseudonymisation and anonymisation were published in D11.2. A Data Protection Impact Assessment template and guidance for completion by each pilot site was published in D11.3. After those, a Data Processing Agreement template was developed that incorporated relevant European Commission Standard Contractual Clauses, and an information security code of practice checklist was developed for consortium wide use. Since these two items are not targeted for inclusion in any other deliverable, they are included as annexes to this deliverable.

The second half of this document focuses on the formal Data Management Plan template published by Horizon 2020. This mostly covers the potential of the project to make available open research data at the end of the project, and how it intends to comply with the FAIR principles. Some of the responses to the template questions are provisional, because decisions about exactly what data is permitted to share, how the data will be made available and how it will be documented with suitable metadata will be determined later in the project.

There is a closing short section on other types of knowledge asset that will be developed in the project, some of which we hope to make available as open source or open access.

This Data Management Plan will be maintained as a living document throughout the project. If it is considered appropriate, a final version of this Data Management Plan will be published as an updated deliverable, in the final year of the project, to provide definitive answers to the template questions.

Table of contents

..... 1

TABLE OF CONTENTS 4

1 THE ADLIFE PROJECT FROM A DATA MANAGEMENT PERSPECTIVE 5

2 ADLIFE DATA SUMMARY 6

2.1 THE ENVISAGED DATA FLOWS AND PROCESSING WITHIN THE ADLIFE PILOT SITES 9

2.2 ADLIFE INFORMATION GOVERNANCE INSTRUMENTS 12

3 ADLIFE OPEN RESEARCH DATA AND OPEN ACCESS 15

3.1 ANONYMISED POPULATION HEALTH DATASETS 15

3.2 KNOWLEDGE ASSETS AND PUBLICATIONS 16

4 ADLIFE DATA MANAGEMENT PLAN TEMPLATE 17

4.1 DATA SUMMARY 17

4.2 FAIR DATA 18

4.2.1 Making data findable, including provisions for metadata 18

4.2.2 Making data openly accessible 19

4.2.3 Making data interoperable 20

4.2.4 Increase data re-use (through clarifying licences) 20

4.2.5 Allocation of resources 21

4.2.6 Data security 21

4.2.7 Ethical aspects 22

5 OPEN ACCESS STRATEGY FOR KNOWLEDGE ASSETS AND PUBLICATIONS .. 23

5.1.1 Aggregated data sets 23

5.1.2 Clinical guidelines 23

5.1.3 AI algorithms 23

5.1.4 Dissemination resources 24

ANNEX 1: ADLIFE DATA PROCESSING AGREEMENT TEMPLATE 25

ANNEX 2: ADLIFE CHECKLIST OF ORGANISATIONAL DATA PROTECTION AND INFORMATION SECURITY REQUIREMENTS 33

1 The ADLIFE project from a data management perspective

ADLIFE is a Horizon 2020 funded project developing innovative digital health solutions to support the healthcare planning and care delivery for patients with advanced (severe) long term conditions or multiple conditions (multimorbidity, with chronic obstructive pulmonary disease and/or heart failure). ADLIFE's solutions will include integrated health information for the multi-professional care team, personalised care planning with patient inclusion, a patient self-management and empowerment platform, and an AI-driven Early Warning System monitoring for acute health deterioration. ADLIFE aims to demonstrate positive patient and clinician experience of using the solutions, improved health outcomes including better quality of life and fewer hospitalisations, and a reduced burden of disease on the family and the health system.

The **ambition** of ADLIFE is to:

- demonstrate that the ADLIFE personalised care model can be deployed and replicated on a large scale in different environments and be trusted with regard to data access, protection and sharing;
- achieve quantified gains in health status, preventing unnecessary suffering (by qualitative analyses), slowing down clinical and functional deterioration (through functional assessment) and improving Patient Reported Outcomes (PROMs);
- obtain improvements in efficiency by making a better use of resources and increasing the coordination among care stakeholders;
- protect functionality and enhance autonomy, empowering patients to participate in decisions making on their own health and adapting to their changing conditions and context.

ADLIFE's technology innovations are getting set to be deployed, used and evaluated in healthcare environments in Spain, UK (England, Scotland), Denmark, Israel. The aim of these large-scale pilots is to demonstrate the effectiveness of the ADLIFE intervention when deployed in clinical real conditions. The project will generate evidence of benefits for the patients, the families and carers, the professionals and the health system.

These sites will each enrol patients into a multi-site quasi-experimental design (non-randomized, non-concurrent, controlled trial) and mixed method analysis, to enable the health outcome and economic impact of ADLIFE to be assessed. The sites will therefore be the primary project sources of patient level data for the trial. Their data have already been used for the technical developments such as the implementation and validation of the digital tools, including the training and validation of the AI components.

A robust approach to data protection and data management has been adopted prior to any contact with patients or their health data. This approach is summarised in this report, with details already provided in three WP11 deliverables and two annexes to this deliverable.

If it is agreed and approved that some individual level or aggregated health and care data can be made available as open data after the project, policies regarding its data management, storage, protection, access and ownership will be developed and included as an updated Data Management Plan before the end of the project.

2 ADLIFE Data Summary

This chapter summarises the categories of patient level data that will be processed through the project and the ways in which it will be processed.

Figure 1 presents a high-level diagram of the overall study design, the details of which are documented in the research protocol (D11.1), and subsequent updates.

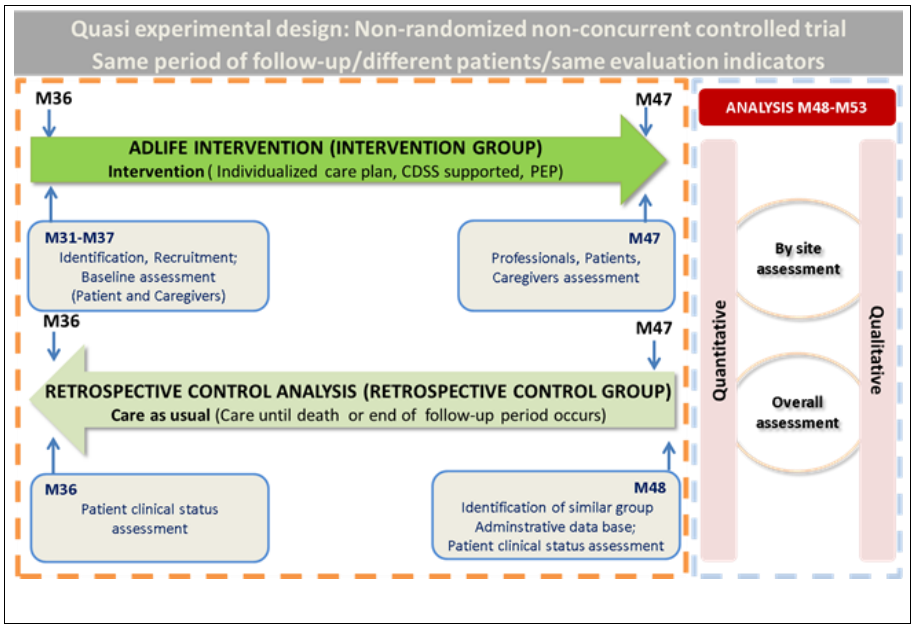


Figure 1: ADLIFE study design

The empirical methodology of the pilot study involves the allocation of eligible patients to intervention and control groups at the sites. Informed consent will be obtained from the intervention group for the use of their health data, as they will also receive care that is supported by ADLIFE innovations. No consent will be obtained from the control group patients, and so their data may only be used anonymously, with ethics committee approval and subject to approved safeguards. Data from potentially eligible patients (prior to the allocation of patients to study arms) will be used for some system development activities, as training data. Mock (synthetic) data that is unrelated to real patients will also be used for some developments. The details of this empirical methodology are given in the ADLIFE research protocol (Deliverable 11.1).

The information architecture of the ADLIFE study design involves the following categories of patient level data collection, processing and communication:

- Mock/synthetic healthcare data for software development and testing.
- Training healthcare data obtained without consent, if permitted by the site, for defining and refining the prediction models behind the clinical decision support systems (artificial intelligence algorithms). However, the developed algorithms will only be used in the care of patients who have given consent.
- Control healthcare data without consent and intervention healthcare and patient/health care professional reported data with consent, routinely collected for software usage and project evaluation purposes.

Figure 2 presents the planned data flows, focusing on the pilot sites which will do most of the patient-level data processing, and all of the personal data processing.

For readability, arrows are not shown on the diagram, but the data flows are vertical from top to bottom on the diagram. The vertical colour coded zones represent the data collections by category. The numbered steps on the edges of the diagram are elaborated in the text below, by number. The data tenants are represented by colour coded boxes to reflect the sovereignty of data in every step. The pilot sites are:

- Basque Country (Osakidetza), Spain
- NHS Lanarkshire, United Kingdom
- University Hospitals Coventry and Warwickshire NHS Trust, United Kingdom
- Odense University Hospital, South Denmark
- Assuta Ashdod Hospital and Maccabi Healthcare Services Southern Reg, Israel

These are all consortium partners or contractual third parties. Each of them will act as the primary legal entity and data controller for the processing of personal data originating from their site on behalf of ADLIFE. In that data controller role, they will be responsible for all of the data objects and data flows shown in blue on Figure 2.

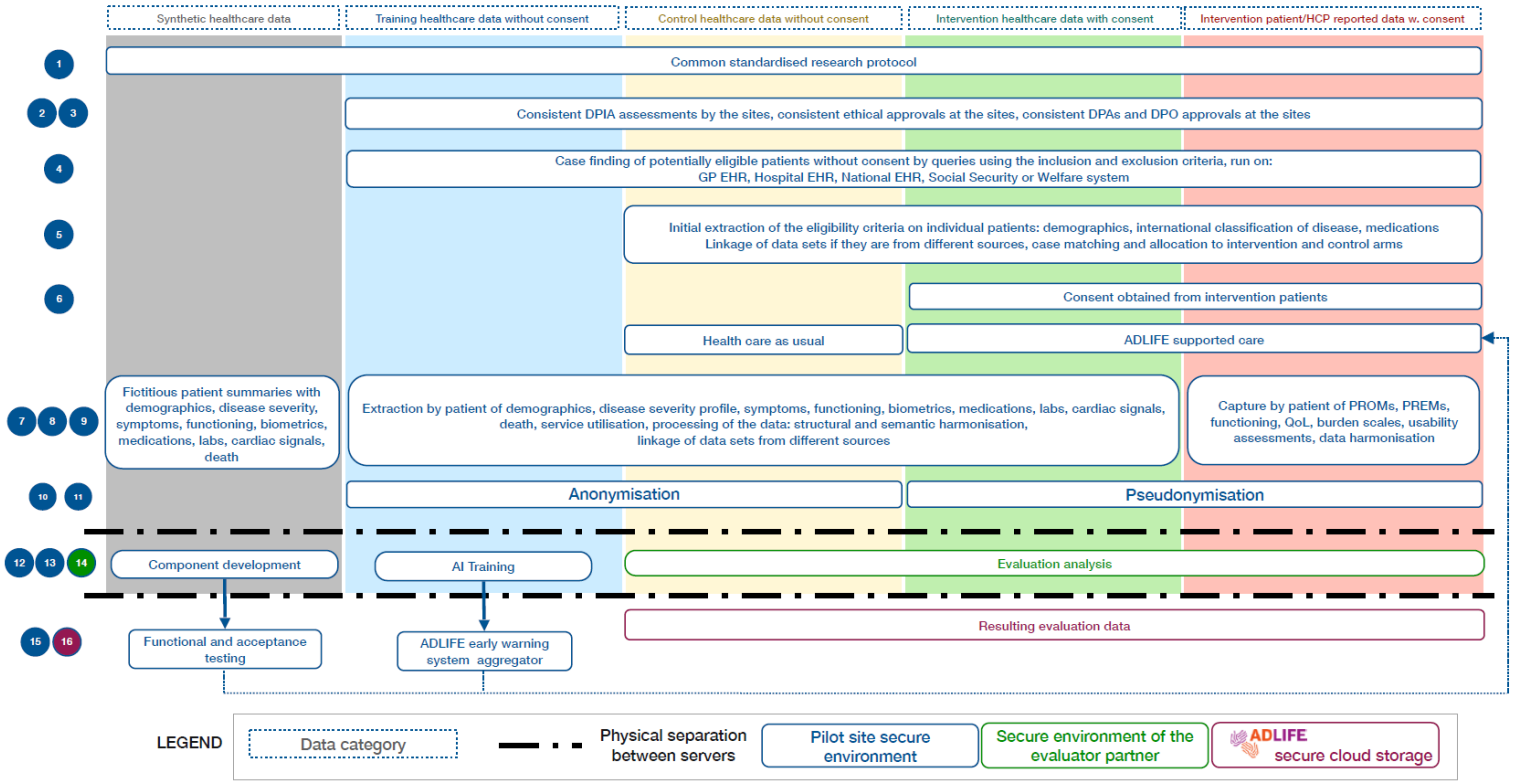


Figure 2: The envisaged data flows and processing within the ADLIFE pilot sites

Abbreviations used: DPIA = Data Protection Impact Assessment, DPO = Data Protection Officer, GP = General Practitioner, EHR = Electronic Health Record, PROM = Patient Reported Outcome Measure, PREM = Patient Reported Experience Measure, AI = Artificial Intelligence

2.1 The envisaged data flows and processing within the ADLIFE pilot sites

Please note that this is a data-oriented summary. The full details of the project's scientific methodology are given in the research protocol (Deliverable 11.1). The numbered points below relate to the numbers shown on the left and right edges of Figure 2.

1. The pilot site partners, alongside other partners in the consortium, have developed the research protocol that specifies the details of how the trial will be conducted. This includes the eligibility criteria, the recruitment methodologies, the data that will be needed for the study, how it will be processed and communicated - and in which forms - with other partners, and how it will be processed at the site and by other partners. Specifically, sharing of data with other partners is documented using 'Data Sharing Storyboards' with the common theme that no identifiable personal data leaves the healthcare organisation/data controller. The protocol includes a brief summary of the data protection measures to be adopted, which are elaborated further in this data management plan and are presented in detail in D11.1 (and its subsequent updates).

2, 3. Each of the sites has conducted a Data Protection Impact Assessment (DPIA) as required by the GDPR. They have been assisted in the conduct of this through a generic DPIA guide (D11.3) and support from partner Kronikgune, the co-ordinator and WP11 leader. Each site has to conduct a DPIA because it is the legal entity overseeing the processing of its data. The project itself is not a legal entity and cannot be a data controller. The partners are also collaborating on their ethics committee applications, so that these can also be consistent with each other.

Osakidetza and AMCA obtained early permission to generate anonymised data sets on eligible patients to be used to develop the AI predictive models, which were further processed to ensure anonymisation by the technical team, corresponding to Figure 2 point **2.** Later, all pilot sites have obtained ethics committee approval for the actual piloting, for the intervention and for use of the control and intervention data corresponding to Figure 2 point **3.** Sites were guided through generic application templates and support from Kronikgune.

Each site will consult and adhere to its data protection and information security policies and guidance, aligned with the checklist included here in Annex 2. They usually also required obtaining the approval of the local data protection officer to the research protocol, unless the DPO was part of the ethics committee, to review the intended data flows across the consortium and the data protection measures to be adopted.

4. The research protocol defines a set of inclusion and exclusion criteria for participant recruitment. All of the pilot sites run the same search on their electronic health record system in order to identify equivalent candidate recruitment pools of patients. Where necessary, the extraction query may be extended to other healthcare repositories such as local GPs, a national EHR system or other social security or welfare databases. The query will be run by personnel authorised to access the electronic health record system, and the results of this extraction will be restricted to authorised personnel.

The first wave of datasets from Osakidetza and AMCA, which were robustly anonymised, were used for guiding the development of the AI prediction models. Other component development and functional testing used synthetic data to limit the need for patient data (anonymised or not) to leave each site.

5. The initial extraction of patient level data from the eligible patients will need to be run prior to obtaining any informed consent. The case-finding data cannot be anonymised initially because there will be a need to identify patients who will be allocated to the intervention arm

(Step 5). It will therefore need to be run by personnel authorised to access the electronic health record system, and the results of this extraction will need to be restricted to authorised personnel who will identify patients to be allocated to the intervention arm and will be directly contacted for recruitment (Steps **5** and **6**). The methodology for case matching of patients to these study arms is defined in the research protocol.

6. Intervention arm patients will be contacted by telephone, email, post or face-to-face encounters, and briefed according to ethically approved Patient Information Sheets. Written informed consent will be sought from all of the patients allocated to the intervention arm. The patient information sheet and consent form template, and GDPR Transparency Notice are part of the ethics committee application which will have been approved prior to use. They will all be expressed in a locally relevant language. Informal caregivers may also be nominated by intervention arm patients to be included (not shown on Figure 2). The remaining eligible patients, plus any who decline to participate actively, may be allocated to the control arm whose (retrospective) electronic health record data will also be extracted.

7. The synthetic (mock) data needs to reflect the structure and semantics of the real data but is directly based on real patient data. Synthetic data that draws on the profile of real patient populations but does not actually include patient data has been used for some ADLIFE component development and testing (e.g. for the semantic interoperability suite).

8. A detailed dataset has been defined in the research protocol, but additional data items may be identified through interacting with the technical partners in the project and with the evaluation partners. It is expected that at least some data items would be extracted for the following categories of health and health care information, as indicated in Figure 2:

- demographics (especially age, gender)
- disease and its severity profile
- current symptoms
- functioning such as mobility
- biometrics such as blood pressure, pulse rate and rhythm, body weight
- medications
- laboratory test results, radiology investigation results, cardiac signals such as ECG
- information about if and when a patient has died and the cause of death
- healthcare service utilisation such as hospital admissions, clinic visits.

Some of this information will also be required retrospectively on the control arm patients and for extracting a training dataset from the source systems (but retained at each site). The same data would be extracted retrospectively and on a regular prospective basis for patients in the intervention arm. These data have been determined as being relevant and necessary for the conduct of the ADLIFE project and study, to comply with the data minimisation principle of the GDPR. The ADLIFE study, in this context, means the development, testing, deployment and patient care use of components such as clinical decision support and MDT care planning (for intervention arm patients).

9. On patients in the intervention arm, there will be additional patient reported data collected through a combination of the patient empowerment platform and occasional questionnaires. This will include PROMs, PREMs, functioning, Quality of Life scores, measures of the burden of care. Both patients and healthcare professionals will complete standardised instruments: technology acceptance and adoption assessments and satisfaction questionnaires relating to the use of the ADLIFE system and its perceived benefits. Caregivers will also complete the burden of care, wellbeing and satisfaction questionnaires and interviews.

10. In order to make healthcare data accessible for research within the consortium, it needs to be protected. Intervention arm patients have given consent and have a GDPR legal basis to be used as personal data. Anonymization is necessary for control arm patients and for the AI training data since these will not have a GDPR legal basis.

The anonymization rules build on D11.2. f_get_arx_dataset() is the main function where the whole anonymization process takes place. The parameters needed are the anonymised IDs, the file with the demographic data, the file that is going to be anonymised and its corresponding date fields.

This function performs the following steps to complete the process with each of the files:

- An outerjoin with the data of both files is made.
- The birth_date variable of the merge data is taken, and their values are ordered from lowest to highest and the range between max and min in days is calculated.
- From that range of days previously calculated is chosen randomly a value that is used to subtract from the minimum birth date and a random date reference is obtained.

This new date previously calculated is always smaller than the minimum birth_date value and is used to calculate the deltas. Deltas are achieved by subtracting this date from the date fields to be anonymised and finally the result of the difference is stored as time deltas in days.

11. In order to link the retrospective healthcare data for the purposes of study evaluation, the prospectively extracted healthcare data and the patient reported data for each patient, the intervention arm datasets need to be pseudonymised by staff within the partner's healthcare organisation (data managers) who have authority to access personal health data. (For healthcare delivery to control and intervention arm patients, staff will continue to use identifiable data. Only the demographic traits required for the research will be extracted from the data sources, no other identifiers. The pseudonymisation via assignment of study ids to patients and health professionals is done by the authorized staff in each pilot site within the Administrator menu of the PCPMP, which automatically assigns a study id to each individual when registered for the first time. These study ids can be updated by the pilot site admins if necessary. The mechanisms for linkage, such as linkage tables, will be kept securely, and will be known to the corresponding pilot site admin(s) only At the end of the pilot study, pilot site admin will generate the data export from the PCPMP Administrator menu that contains the clinical parameters that have been defined in the Data Collection Guide for all intervention arm patients, in a pseudonymised way. These data exports can be safely handed over to the ADLIFE Evaluation Team.

The research protocol, supplemented by the wording and agreement of each consented patient, will determine what happens to already-collected data in situations where a patient withdraws from the study or dies.

12, 13, 14. The end result of all of this data processing, under the data controllership of each of the healthcare partners, will be 6 categories of anonymised/pseudonymised data:

- Synthetic (mock) healthcare data
- Training healthcare data
- Control arm healthcare data
- Intervention arm healthcare data
- Intervention arm patient reported data
- Intervention arm healthcare professional reported data

Data for the conduct of the trial will only be available to the site from which the data originated. For study evaluation, the relevant individual-level data only needs to be made available to the evaluation partner, with no other consortium members needing access to even the anonymised data unless they have identified a legal basis and need to do so. (Aggregated analysis results will be shared and used by the consortium for scientific publications and dissemination.)

The preferred storage and usage for research within the consortium of the above datasets will be in physical servers that are different from the ones where all the above processes took

place but within the secure environment of each pilot site, under the terms of the ethical approval. These datasets will be transferred to these servers for partners to use for:

- the design, implementation and validation of the software components within the ADLIFE solution (Step 12)
- the training and validation of the AI used within the early warning system (Step 13)
- the population of ADLIFE components for the provision of healthcare, by MDT members, patients and caregivers (Step 14)
- the clinical evaluations of the impact of the ADLIFE solution on health outcomes and burden of care (Step 16)
- the health economic implications of the impact of the ADLIFE solution on cost of care (Step 16)

Comentado [ÁD1]: should be step 16, not 17; already chaged

The data for project evaluation purposes will only be available to the site from which the data is originated, it will be kept within the healthcare provider infrastructure. For the evaluation process, the data managers from the sites will upload data on specific folders of SharePoint with restricted access (only for data managers and evaluator partners) - this will be the ADLIFE secure cloud storage. For the statistical analyses, the evaluation data will be stored by evaluator partners in physical servers within the secure environment of the evaluator partner, under the terms of the ethical approval.

15. The iterative communication process between site systems (Step 12) and ADLIFE developer teams at the development phase will allow the developer teams to access synthetic data for developmental and testing purposes in a preproduction environment. The testing with real, pseudonymised data will be conducted at the site, by authorised personnel on a staging environment prior to going live. The ADLIFE consortium will provide training to staff in how to perform this securely.

16. The evaluation team will proceed with the credible and feasible analysis on control and intervention data within the evaluator secure environment. The results of such analysis, which will not comprise patient level data, including the clinical evaluations of the impact of the ADLIFE solution on health outcomes and burden of care and the health economic implications of the impact of the ADLIFE solution on cost of care will be uploaded to the ADLIFE secure cloud storage (hosted on MS SharePoint). They will then be used by consortium partners scientifically. Some aggregate data sets may also be made available as open research data.

Comentado [ÁD2]: evaluator, not pilot; already changed

Kronikgune will host the ADLIFE cloud storage of the evaluation data for the necessary duration of the research and publications. The access and processing of data from this ADLIFE centralised cloud platform will be audited.

The above numbered steps have described the way in which data will be processed in order to inform the development of the ADLIFE toolkit components, and to conduct the trial of their use. For patients in the control arm, healthcare delivery will continue as normal throughout the study period, and routinely collected health data will continue to be accumulated, which will be part of the eventually extracted control study data. Care for the intervention arm patients will be supported by the ADLIFE components. Healthcare staff will access identifiable patient data, as usual.

2.2 ADLIFE information governance instruments

Almost all of the governance instruments foreseen in the original DMP have in practice proved necessary, and they have been developed. The list of actual instruments developed and used is summarised here.

1. The research protocol (initially submitted as D11.1 but with later updates that proved necessary as the actual deployment, integration and patient recruitment opportunities were finalised at the sites). This includes, in the relevant language, patient and caregiver information sheets, informed consent forms for patients and for caregivers and GDPR Transparency Notices. These were included within the ethics committee application made at each site.

2. Information security measures and guidelines on pseudonymisation and anonymisation (D11.2). This document is being used as the project basis for processing, including transforming, sensitive personal data items.

3. Data Protection Impact Assessment template and guidance for completion by each pilot site (D11.3). As anticipated, the sites each have a DPIA template that they are required to use, and the deliverable served as anticipated as a source of content from which each site could select and personalise for the relevant headings in their template. These have now been completed and the revised DMP will report that, plus any important risks that were identified and needed special safeguards. Fortunately, relatively limited numbers of risks were identified through the introduction of new processing and the European nature of the consortium. Appropriate safeguards are in place at each site, with the support and agreement in each case of the Data Protection Officer and ethics committee.

4. It was originally considered possible that special data sharing agreements might be required in order for pseudonymised and anonymised data sets to be transferred from pilot sites to the partners (in the other partner European countries) conducting the evaluation analyses. In practice the site DPOs have confirmed that the consortium agreement terms cover these anticipated data flows and processing, and that a supplementary data sharing agreement is not required.

5. However, in contrast, a relatively complex process was required to develop, agree and sign Data Processing Agreements. ADLIFE initially started using a previously used DPA from the C3-Cloud project, which had several partners in common with ADLIFE. Through consultation with the partners and especially with their DPOs an updated and universally accepted version of this agreement template was created. As mentioned earlier in this report, the European Commission at that point published Standard Contractual Clauses (SCCs) that provide a standardised and well-respected legal text for the data processing requirements that our template was intending to address. Considerable work was then undertaken with valued help from some of the DPOs to modify our template such that it did not duplicate or clash with the SCCs, and became a lightweight project specific covering document to the relevant SCCs. It was important for the partners to fully understand which of two different SCCs they would need to utilise, with which other partners, in order to provide comprehensive data processing agreement coverage. This is important because each DPA is signed one-to-one between a pilot site and each of the technical partners providing some of the components to be deployed at its site It also proved necessary to modify the plans for signing the agreement for relationships between the pilot sites and UK partners (with the technical partner, now university of Birmingham, and with the pilot site in Strathclyde) due to Brexit and the enactment of a UK GDPR. The ADLIFE DPA is included as Annex 1 of this deliverable, as it is not otherwise part of any other forthcoming deliverable.

6. Finally, an ADLIFE information security policy and code of practice was developed utilising the Five Safes model as the over-arching framework. Because each site has its own internal policies that have historically been developed and approved within its organisation, it was agreed that it is not appropriate ADLIFE to try to impose a new policy on each organisation just for the purposes of our project. We therefore decided that we should produce a checklist of measures and commitments that we would expect all partners to be able to agree to, for them to verify against their internal policies and codes of practice and available security measures. This has enabled a coherence of trustworthy practices across the consortium

without disrupting each partner. This checklist is included in Annex 2, as it is not otherwise part of any other forthcoming deliverable.

The project coordinator, Kronikgune, will store a copy of each of these instruments, for the record and for possible inspection by the European Commission if required. Kronikgune will additionally store copies of all ethics committee submissions and approvals, and any constraints or requirements imposed by the committees.

3 ADLIFE open research data and open access

The data sets and knowledge assets that will be created through ADLIFE and have potential for wider reuse or open access are:

1. One or more anonymised population health datasets of health and care information on patients with advanced long-term conditions and/or multimorbidity (chronic obstructive pulmonary disease and/or heart failure)
2. Aggregated health outcomes and health economic data that has been used within the study and its published results, and might be further reused by others
3. Clinical guidelines and decision support services that have been specifically tailored for advanced conditions and for combined use in cases of multimorbidity, expressed in a human readable form and as clinical decision support services
4. AI algorithms suitable for inclusion within an Early Warning System for patients with advanced long-term conditions
5. Academic publications, conference presentations and posters, and other dissemination materials that showcase the methodology, results and solutions of the project

3.1 Anonymised population health datasets

The project cannot commit at this stage to make anonymised population health data sets available as open access data. This is desirable but will be subject to ethical approval and organisational approval at the pilot sites, possibly also additional verification of the robustness of the anonymisation. The challenge we will face is that rich clinical data sets are very difficult to anonymise robustly.

Limited opportunities for open access data were foreseen early in the project because of data protection concerns. Interactions between WP 1, WPs 8 and 11 have tracked the developments at each pilot site when conducting their DPIA, which has reinforced the caution that was first anticipated. We have now experienced first-hand at each pilot site the concerns of the DPOs and ethics committees about the processing of personal data, and the limitations on data flows and data use they impose. We will keep the topic open and see whether any aggregated anonymised data might be publishable by the end of the project. However, at this stage we foresee it being more likely that aggregated analytics data will be shareable, usually as the reproducibility dataset underpinning each of the academic papers that we produce that contain pilot evaluation results.

If the whole of the dataset cannot be made available, selected subsets of the data, which are not so rich, might be possible to release as open research data. This will be explored towards the end of the project when the dataset can be comprehensively assessed.

The responses given in the next section are based on the assumption that some data may be available for open data sharing. The project will therefore take care to ensure that relevant contextual metadata, in accordance with the FAIR principles, are captured and included within the data sets. Descriptive metadata to allow for discovery will be added later in the project. Data access arrangements, also in accordance with the FAIR principles, will also be specified. Data will be made open access via the EU Open Research Data Pilot (ORDP). The Data Management Plan template specified by the European Commission is provided in the next chapter.

3.2 Knowledge assets and publications

The numbered items 2 to 5 in the list above, covering aggregated data sets, clinical guidelines and decision support rules, algorithms and dissemination materials are discussed later in this report in a chapter on open access.

4 ADLIFE Data Management Plan template

This section is based on the Data Management Plan template provided by the European Commission for Horizon 2020 projects[1].

4.1 Data summary

Please see the previous section for a detailed description of the data processing activities to be undertaken, primarily at the pilot sites. The responses given in this section refer to the expected generation of anonymised population health data sets. These will have underpinned the main research results that will be published at the end of the project, and also have the potential to be reused by other researchers.

Aspect	Response/explanation
Purpose of the data collection/ generation and its relation to the objectives of the project	Health and care data collected from patients with advanced long-term conditions from the healthcare sites across Europe and Israel, plus patient reported data on quality of life and health outcomes. The purpose of the data collection and processing is to conduct an evaluation of effects of a digital solution enabling cross sector care on health and healthcare utilisation for patients with multimorbid chronic conditions. The dataset will include patients who have utilized the ADLIFE project digital solutions, who will enable the health and economic evaluation of the solutions.
Types and formats of data generated or collected by the project	Database tables, the format of which will comply with the HL7 FHIR standard and are documented in D3.1.
Any re-use existing data and how this will be done	The existing data will be electronic health record information on these patients, that had been collected as part of routine care in the years prior to the commencement of the project and its trial. Data will be extracted using the electronic solutions, either the ADLIFE solution for data included in the study, or the anonymisation tools outlined above. Data will be reused either for evaluation of the ADLIFE solution or to ensure the ADLIFE solution contains the full relevant clinical history of participants during the study.
The origin of such data	Hospital and general practice electronic health records, national electronic health records and Social Security or welfare system

[1] Full template available here for reference: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/reporting/h2020-tpl-oa-data-mgt-plan_en.docx

	records. The exact combination of which data sources will vary between the pilot sites.
Expected size of the data	To be confirmed by the end of 2023
Likely users of the data	Health and care professionals providing care to participating patients and/or involved in evaluating the ADLIFE digital solution. Public and population health researchers investigating the impact of long-term conditions and multi morbidity, health economists and health informatics researchers.

4.2 FAIR data

4.2.1 Making data findable, including provisions for metadata

Aspect	Response/explanation
Are the data produced and/or used in the project discoverable, identifiable and locatable by means of a standard identification mechanism	If we define an anonymised dataset that can be made available as open research data we will liaise with the EC's OpenAire project (https://www.openaire.eu) and with the Fair4Health project (https://www.fair4health.eu) to make the data available.
What standard identification mechanism used (e.g. persistent and unique identifiers such as Digital Object Identifiers)	We will use DOI.
Is meta-data available through catalogue?	It will be, once the data that we will be sharing is finalised and the repository/portal that will be used has been decided.
Can meta-data be used for search?	Yes, it will be. Please see https://www.thieme-connect.com/products/ejournals/html/10.1055/s-0040-1713684
Naming conventions used	We will align with the Horizon 2020 FAIR4Health on this point (one of our partners is in this project). We will adopt the naming conventions of HL7 FHIR. (https://wiki.hl7.org/FHIR_Guide_to_Designing_Resources#Naming_Rules_.26_Guidelines)
Clear versioning supported?	It will be.
Additional keyword search?	It will be.
What metadata will be created using which standards?	We will use HL7 FHIR Profiles as the metadata of the data to be shared. To be elaborated later in the project, on advice from FAIR4Health, and any other advice from the European Commission.

4.2.2 Making data openly accessible

Aspect	Response/explanation
Will data be made openly available as the default?	If we can define a dataset that we can robustly anonymise and have permission to make available as open data, we will make this available by default.
Which datasets will NOT be openly available and why?	We will not make data available if we believe, or we are advised, that it is not possible to robustly anonymise the data, because of distinctive patterns in the data due to some of the population profiles that are included.
How will the data & meta-data be made accessible (e.g. by deposition in stated repository)?	By deposition in the chosen repository. We expect to use the (partner) SRDC onFHIR repository as a means of sharing anonymised data and its metadata.
If known repository, what arrangements explored?	To be determined later in the project.
If project-specific access, then:	To be determined later in the project.
– Data Access Committee	A committee comprising some or all of the partners involved in the project will determine the policies for which data will be made open access.
– Any conditions for access (i.e. a machine-readable license)	The conditions that the committee will determine will include confirming which data sets (patient level and aggregate level, all robustly anonymised) are suitable for open data access and ensuring that the necessary approvals have been obtained from the originating sites. The committee will determine the time interval after the project when the data will be made available, how it will be discovered and accessed, the open access licence terms that will apply and how data access will be requested and granted. The committee will also make long-term sustainability decisions, including a sustainable business model for the data sets.
What methods or software tools will be needed to access the data?	To be determined later in the project. We expect to use the (partner) SRDC onFHIR repository as a means of sharing anonymised data, and guideline modelling tools as appropriate once the language for CIGs has been finalised.
– Documentation for software	This will be provided.
– Availability of software	The SRDC onFHIR repository, which is shared as Open Source on GitHub
Institution and researcher vetting process/procedures - describe	To be determined later in the project.

4.2.3 Making data interoperable

Aspect	Response/explanation
Are the data produced in the project interoperable	Interoperability standards will be adopted by design, including the use of HL7 FHIR, in order to harmonise the data coming from the different pilot sites. This is necessary for the project itself, for the data analytics work that will be undertaken, but also serves the benefit that any research data we make openly available later will also be standardized.
If not, explain why not	N/A
Data and metadata vocabularies, standards or methodologies used	To be determined later in the project. Within the project, for coded data, we will be adopting ICD10 and ATC for diagnoses and medications, and will be mapping site specific terminologies to these. It is expected that any patient level dataset shared would include these terminologies.
Standard vocabularies used	To be determined later in the project.
Mappings from uncommon or project-specific ontologies or vocabularies to more commonly used ontologies	We do not anticipate the need to adopt uncommon ontologies or vocabularies.

4.2.4 Increase data re-use (through clarifying licences)

Aspect	Response/explanation
Will data be available for onward data-sharing/re-use?	Yes
Approach to data licensing for onward use	To be determined later in the project, but it will aim to facilitate onward sharing.
Likely date for data availability for onward use	During 2024
Explain any restriction on date of availability	None is anticipated.
Possible restrictions on onward data-sharing	To be determined later in the project, but it will aim to facilitate onward sharing.
Data retention policy (including availability for data-sharing)	We will follow data retention policies as determined at sites. Evaluation/open access data will be retained according to EC guidance.
Description of data quality assurance processes	To be determined later in the project.

4.2.5 Allocation of resources

Aspect	Response/explanation
Estimated project costs for making data FAIR	This will be determined later. However, we do not anticipate substantial costs because we will establish the anonymized data repository according to standards and with suitable metadata as an intrinsic part of the project. We will use the Open Source OnFHIR Repository, which is supported by SRDC.
Data management responsibility across the project	The co-ordinator, Kronikgune, will take lead responsibility for this, but other technical partners will support.
Resources required for long term preservation (costs and potential value, who decides and how what data will be kept and for how long)	To be determined later in the project, under the responsibility of the co-ordinator, Kronikgune. This would include a sustainability business model for data sets that will be made available open access, covering costs of long-term storage, the maintenance of the data sets if necessary, and any human resources required to manage data sharing.

4.2.6 Data security

Aspect	Response/explanation
Data security measures used (including data recovery as well as secure storage and transfer of sensitive data)	As indicated earlier in this report, a number of information governance and data protection and information security instruments are being developed and will be included within forthcoming project deliverables. Additionally, a core component of the ADLIFE solution will be focussed on ensuring security and privacy of the solution (SPS). The onFHIR repository, integrated with SPS components, will enable authentication, authorization, anonymization and audit logging. As a minimum, no personal, identifiable data will leave study sites.
Where data will safely be stored (in certified repositories for long-term preservation and curation). Provide detail	To be determined later in the project, under the responsibility of the co-ordinator, Kronikgune.

4.2.7 Ethical aspects

Aspect	Response/explanation
Any ethical or legal issues that can have an impact on data sharing	There are no moral ethical issues. Processes for handling the withdrawal of patient from the study, for their death during the study period, including what should happen to their study data, are documented within the research protocol, D 11.1. There are potentially data protection issues which we will examine carefully before determining which data items and on which population profiles can be made available as open research data.
References to ethics deliverables and ethics chapter in the Description of the Action (DoA) – if relevant	Work package 11 deliverables D11.1, D11.2, D11.3
Questionnaires dealing with personal data	Some of the anonymised data will have been derived from questionnaires completed by patients, covering topics such as health outcomes, quality-of-life and burden of care.
How is informed consent for data sharing and long term preservation sought in such questionnaires?	It will not be a GDPR requirement to obtain informed consent for the scientific use of anonymised data. However, this data reuse will be described in the Transparency Notice accompanying the informed consent forms, and to be approved by the ethics committees representing the pilot sites. These statements will cover all of the intended uses of the data post project, by project partners for future research or teaching, as well as external researchers.

5 Open access strategy for knowledge assets and publications

The section summarises the intentions of the consortium towards other assets that will be developed through the research in addition to research data sets. Regarding the potential for open access assets, the project has just completed its engagement with the European Commission‘s Horizon Booster programme, to explore the potential sustainability and commercialisation opportunities. Commercial exploitation might prove in conflict with open access knowledge resources, so at this stage we do not anticipate being able to commit to open access knowledge assets such as AI algorithms and computerised guideline representations.

5.1.1 Aggregated data sets

As part of conducting the research evaluations of the health outcomes impact and health economic impact of ADLIFE, the consortium will generate a number of aggregated health outcomes and health economic data tables. Some of the content of these will be derived from our research data, and other comparative data may have been derived from public sources such as the academic literature. Many of these data tables will be used to develop and publish our results, and may therefore be also held by publishers as ESCROW research data. However, we will seek permission from journal publishers to publish these data tables alongside the deliverables to which they relate, on relevant data repositories, with links to these on the project website, so that they are open and available without any data sharing or licensing restrictions. A specifically formulated data access committee, comprising some or all of the project partners, will determine the policies for this data access. These policies will include which data sets (patient level and aggregate level, all robustly anonymised) are suitable for open data access and ensuring that the necessary approvals have been obtained from the originating sites. The committee will determine the time interval after the project when the data will be made available, how it will be discovered and accessed, the open access licence terms that will apply and how data access will be requested and granted. The committee will also make long-term sustainability decisions, including a sustainable business model for the data sets. A sustainability business model will be developed for covering the costs of long-term discovery and storage, the maintenance of the data sets if necessary, and any human resources required to manage the data sharing arrangements.

5.1.2 Clinical guidelines

ADLIFE will use clinical guidelines that have been developed by professional societies or Health Technology Assessment organisations. We will need to adapt these in order that they focus on the necessary care pathways for patients with advanced conditions and for patients who have multi morbidity, for which multiple guidelines need to be used in parallel.

We would ideally like to make the human readable version of these adapted guidelines available publicly, to some extent in academic literature. We will liaise with the publishers of these guidelines to determine if we can additionally publish these on our web site.

5.1.3 AI algorithms

ADLIFE will develop novel AI algorithms for inclusion within an Early Warning System for patients with advanced long-term conditions. We have not yet determined the extent to which these algorithms will be commercially exploited in the future, and what levels of detail could

be made available as open source components. We will make this clear later in the project, building on the outcomes of our recent engagement with the Horizon Results Booster Programme.

5.1.4 Dissemination resources

The partners are committed to investing significant effort in creating academic publications, conference presentations and posters, and other dissemination materials that showcase the methodology, results and solutions of the project. We would ideally like all of this material to be open access, and we will prioritise open access journals. Many conferences request to host presentation slides, and sometimes video recordings of presentations, on their website, to which we will always agree. In situations where conferences do not make these materials available, we will seek permission to host them on our website ourselves, for public access. Publications made by ADLIFE partners will additionally need to comply with the project's publication policy.

The ADLIFE website will be used to disseminate project findings, by sharing links to publications, conferences, abstracts etc. (Access to the papers from our website will be dependent upon the copyright and sharing policies of the journals.) Academic partners are required to deposit author accepted copies of publications in their institutional repositories as per relevant open access policies. Our social media channels will also be used as dissemination channels, as appropriate, throughout the project.

All of this dissemination support is being handled by WP2 and is being periodically reported in their deliverables.

Annex 1: ADLIFE Data Processing Agreement template

THIS AGREEMENT dated [DATE] is made **BETWEEN:**

(1)

[Entity name] (Partner short name if an ADLIFE partner) whose address is [Address including country] (the **"Site"**); and

(2)

[Entity name] (Partner short name if an ADLIFE partner) whose address is [Address including country] (the **"Technical Partner"**).

(Additional entities may be added, if required, for example if the Technical Partner has an affiliate or is an affiliate to another entity)

Each a "Party" and together the "Parties"

WHEREAS:-

(A) The Parties are collaborating on the research project entitled "ADLIFE" funded by the European Commission under grant agreement No 875209 (the **"Project"**) and on 1 January 2020 entered into a consortium agreement in respect of the Project (the **"Consortium Agreement"**);

(B) When the processing does not imply an international data transfer, this Agreement will be supplemented by the standard contractual clauses established in the Decision (EU) 2021/915 of 4 June 2021 on standard contractual clauses between controllers and processors under Article 28(7) of Regulation (EU) 2016/679 of the European Parliament and of the Council and Article 29(7) of Regulation (EU) 2018/1725 of the European Parliament and of the Council, to ensure compliance with Article 28(3) and (4) of Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (the **"Decision (EU) 2021/915"**);

(C) When the processing implies an international data transfer, this Agreement will be supplemented by the standard contractual clauses established in the Decision (EU) 2021/914 of 4 June 2021 on standard contractual clauses for the transfer of personal data to third countries pursuant to Regulation (EU) 2016/679 of the European Parliament and of the Council, to ensure appropriate safeguards within the meaning of Article 46(1) and (2)(c) of Regulation (EU) 2016/679 for the transfer by a controller or processor of personal data processed subject to that Regulation (data exporter) to a processor or (sub-) processor whose processing of the data is not subject to that Regulation (data importer) (the **"Decision (EU) 2021/914"**).

(D) This Agreement is supplemental to the Consortium Agreement and applies only in respect of the processing of the personal data; in the event of contradiction between the Consortium Agreement and this Agreement the latter shall prevail;

(E) The Technical Partner, as part of its allocated responsibilities for the Project in accordance with the Consortium Agreement, shall carry out certain data processing activities in respect

of the data on behalf of the Site as described in Annex II of Decision (EU) 2021/915 or Annex I of Decision (EU) 2021/914 ("Description of the processing").

NOW IT IS AGREED as follows:

1. **DEFINITIONS**

The definitions established in REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (the "General Data Protection Regulation" or "GDPR") shall apply. The following definitions shall additionally apply:

"**Data Protection Law**"	means any law, statute, declaration, decree, directive, legislative enactment, order, ordinance, regulation, rule or other binding restriction (as amended, consolidated or re-enacted from time to time) which relates to the protection of individuals with regards to the Processing of Personal Data to which a Party is subject, including the GDPR and all national legislation in force in respect of the protection of personal data as well as the Privacy and Electronic Communications (EC Directive) Regulations 2003;
"Good Industry Practice"	means, at any time, the exercise of the level of skill, care and diligence that would reasonably be expected at that time from a leading and experienced organisation carrying out activities similar to those carried out under this Agreement, and which seeks to perform its contractual obligations in full and in compliance with all applicable legal standards, including Data Protection Law;
"Jurisdiction"	means the country, territory or jurisdiction of the Site, for the purposes of determining any applicable national law referred to in this Agreement;
"Damages"	means all damages, penalties, sanctions, penalties, losses, damages, costs, charges, claims, amounts paid in compensation, as well as expenses (including legal fees (on a solicitor/client basis), disbursements, costs of investigation (including forensic investigation), costs of litigation activity, out-of-court dispute resolution (including ex gratia payments), judgment payments, interest and fines), other professional fees and charges, reimbursements, costs of breach notification, including notifications to the data subject, cost of claims handling (including credit reference checks, setting up contact centres (e.g. call centres) and making ex gratia payments), all arising out of contractual or non-contractual liability, civil wrongs (including negligence), breach of statutory duty or otherwise;
"**Permitted Purpose**"	means the purpose of the Processing as specified in Annex II of Decision (EU) 2021/915 or Annex I of Decision (EU) 2021/914 ("Description of the processing");
"**Personnel**"	means all persons engaged or employed from time to time by the Technical Partner in connection with this Agreement, including employees, consultants, contractors and permitted agents;
"**Security Requirements**"	means the requirements regarding the security of Personal Data, as set out in the Data Protection Law (including, in particular, the seventh data protection principle of the DPA and/ or the measures set out in Article 32(1) of the GDPR (taking due account of the matters described in Article 32(2) of the GDPR)) as applicable;
"Site"	means an entity that is the data Controller of personal data from patients, caregivers, health professionals or other parties that is processed as part of the ADLIFE project

"Technical Partner"	means an entity that is a signatory of the ADLIFE Consortium Agreement and which may require access to data under the control of the Site for the Permitted Purpose relating to the processing undertaken as part of the ADLIFE project;
"Third Party Request"	means a written request from any third party for disclosure of Personal Data where compliance with such a request is required or purported to be required by law or regulation.

2. **DATA PROTECTION**

2.1 **Arrangement Between The Parties**

2.1.1 The Parties shall each Process the Personal Data. The Parties acknowledge that the factual arrangements between them dictate the classification of each Party in respect of the Data Protection Law. Notwithstanding the foregoing, the Parties anticipate that, in respect of the Personal Data, as between the Site and the Technical Partner for the purposes of this Agreement, the Site shall act as the Controller and the Technical Partner shall act as the Processor, as follows:

(a) The Site shall be the Controller where it is Processing Personal Data in relation to the evaluation of the ADLIFE system and

(b) The Technical Partner shall be the Data Processor where it is Processing Personal Data in relation to the Permitted Purpose in connection with the performance of its obligations under this Agreement.

2.1.2 Where necessary, each Party shall duly notify the data protection authority of its respective country.

2.1.3 Each Party shall operate in accordance with European Commission standard contractual clauses between controllers and processors as specified in Decision (EU) 2021/915 or Decision (EU) 2021/914.

2.1.4 Nothing within this Agreement relieves the Technical Partner of its own direct responsibilities and liabilities under Data Protection Law.

2.1.5 The Technical Partner undertakes to the Site to take all necessary steps to ensure that it operates at all times in accordance with the requirements of Data Protection Law and the Technical Partner shall assist the Site in complying with its obligations under Data Protection Law arising as a consequence of this Agreement, which are further detailed in this Clause 2 (Data Protection). The Technical Partner shall not cause, either by act or omission, the Site to knowingly fail to comply with any of its obligations under the provisions of the Data Protection Law.

2.2 **Data Processor Obligations**

2.2.1 To the extent that the Technical Partner carries out the Processing of any Personal Data in the capacity of Processor on behalf of and in the name of the Site (in the capacity of Controller) it shall:

(a) only Process the Personal Data for and on behalf of the Site for the purposes of performing its obligations under this Agreement, and only in accordance with the terms of this Agreement and any instructions from the Site;

(b) keep a record of any Processing of the Personal Data it carries out on behalf of the Site;

(c) unless prohibited by law, notify the Site immediately of becoming aware of it if it considers, in its (reasonable) opinion, that the applicable EU law in the jurisdiction requires it to act differently from the instructions on the Site, in particular where it considers that any of the instructions on the Site given under Paragraph 2.2.1(a) infringes any rule of Data Protection Law;

(d) take, implement and maintain appropriate technical and organisational security measures which are sufficient to comply with at least the obligations imposed on by Annex III of the Decision (EU) 2021/915 or Annex II of the Decision (EU) 2021/914 and where requested provide to the Site evidence of its compliance with such requirements promptly;

(e) maintain the Personal Data in a form that allows it to be distinguished from other data or information processed by the Technical Partner;

(f) within thirty (30) calendar days of Site's request, allow its data processing facilities, procedures and documentation to be subject to scrutiny, inspection or audit by Site (and/or persons acting on its behalf including any auditors it may appoint), in order to verify compliance with the terms of this Clause 2 (Data Protection), and to provide reasonable information, assistance and cooperation to Site, including access to relevant Personnel. Without prejudice to the foregoing, the cost of the audits shall be borne by Site. Furthermore, the audits shall not affect the information or data of third parties or unreasonably disrupt the normal operation of the Project Technical Partner's business. The results and/or information obtained from the audits shall be confidential and shall be returned to the Technical Project Partner upon termination of this Agreement; and shall not be used beyond the scope of this Agreement;

(g) refrain from disclosing Personal Data to third parties (including subcontractors) in all circumstances without the prior written consent of the Site, except in relation to Third Party Requests where a law or regulation prohibits the Technical Partner from notifying the Site;

(h) promptly comply with any request from the Site to amend, transfer or delete any Personal Data;

(i) notify the Site without delay after receipt by it of any request made by a data subject or data protection authority and shall:

  (i) not disclose any Personal Data in response to any Request made by a data subject or data protection authority Correspondence without first consulting with and obtaining the prior written consent of the Site; and

  (ii) provide the Site with all reasonable co-operation and assistance requested by the Site in connection with any Request made by a data subject or data protection authority;

(j) notify the Site without delay if becoming aware of any actual, suspected, announced or threatened Personal Data Breach in relation to the Personal Data (and subsequently confirm such information in writing) and in addition shall::

  (i) conduct, or assist the Site in conducting, such investigations and analyses as the Site may reasonably request in relation to such Personal Data Breach, in accordance with Personal Data Breach notification, as set out in Article 33(3) GDPR;

  (ii) implement the necessary corrective actions or measures to restore the security of the Personal Data concerned; and

(iii) assist the Site in making appropriate notifications to the National Data Protection Authority of the Site and to the Data Subject(s) concerned;

(k) comply with the obligations imposed upon the Processor under Data Protection Law;

(l) act with reasonable care, in accordance with Good Industry Practice, to assist the Site in complying with the obligations imposed on the Site by Data Protection Law, including, but not limited to:

(i) compliance with the Security Requirements;

(ii) obligations relating to notifications that Data Protection Law requires to be sent to the National Data Protection Authority and/or the relevant Data Subjects;

(iii) conducting Data Protection Impact Assessments (and, where required by Data Protection Law, consulting with the Data Protection Agency and any other relevant Regulatory Authority on matters relating to such Data Protection Impact Assessments); and

(iv) notifying the Site of Personal Data Breaches without delay after becoming aware of them, unless the Personal Data Breach is unlikely to endanger the rights and freedoms of natural persons.

(m) As soon as one of the following events occurs:

(i) termination or expiry of the Consortium Agreement, which, for clarification purposes, shall also entail the automatic and immediate termination of this Agreement;

(ii) termination or expiry of this Agreement (as applicable); or

(iii) the date on which the Personal Data ceases to be relevant or necessary for the Permitted Purpose

the Technical Partner shall cease Processing of all Personal Data (if any) and shall permanently and securely return and/or destroy so that it is no longer retrievable (in accordance with the Site's written instructions) all Personal Data and all copies in its possession or under its control and, where requested by the Site, will certify that such destruction has taken place (without delay, and in any event within fifteen (15) days of such request) except to the extent necessary to comply with any applicable EU law requiring the retention of Personal Data;

(n) not carry out (nor instruct nor permit a third party to carry out) a transfer of any Personal Data to a Third Country, except with the prior written consent of the Site and in accordance with European Commission's standard contractual clauses for the transfer of personal data to third countries as specified in Decision (EU) 2021/914 and any terms and conditions the Site may impose on such transfer, to the extent deemed necessary to satisfy the requirements of ensuring that transfers of Personal Data outside of the European Economic Area (EEA) are afforded adequate protections as set out in Data Protection Law;

2.2.2 Except as otherwise provided, this Agreement does not transfer ownership of, or create any licence (implied or otherwise), to any intellectual property rights in any of the Personal Data.

2.3 **Staff of the Technical Partner**

2.3.1 The Technical Partner shall disclose Personal Data only to those Personal Data which the Technical Partner requests to co-operate in the performance of its obligations under this Agreement (the "Project Personnel") and shall ensure that no other Personnel have access to such Personal Data.

2.3.2 Technical Partner shall only disclose Personal Data to its Personnel when the following conditions have been met in relation to such Project Personnel:

(a) That the Technical Partner has taken (and maintains in place) all reasonable measures to ensure the reliability and integrity of each member of its Personnel;

(b) That each member of its Personnel has undergone adequately clear pre-employment checks;

(c) that each member of its Personnel has received, and continues to receive on an annual basis, reasonable levels of training in data protection law and in the handling and management of Personal Data; and

(d) each member of its Personnel shall have entered into appropriate contractually-binding confidentiality undertakings.

2.4 **Sub-contractors**

2.4.1 The Technical Partner shall not sub-contract the performance of any of its obligations under this Agreement without the prior written consent of the Site. All approved subcontractors are listed in Annex IV of Decision (EU) 2021/915 or Annex III of Decision (EU) 2021/914 (“List of sub-processors”).

2.4.2 When it is necessary to subcontract all or part of the Personal Data Processing, the Technical Partner shall communicate this fact previously in writing to the Site. Subcontracting may be carried out if, after thirty (30) days, the Site does not expressly object.

2.4.3 In the event the Site refuses the subcontractor, the Technical Partner will propose a new subcontractor and/or provide the necessary information to demonstrate the subcontractor's compliance with the GDPR.

2.4.4 Notwithstanding any consent or approval given by the Site, the Technical Partner shall remain primarily responsible to the Site for the adequacy of data protection with respect to the GDPR. Notwithstanding any consent or approval given by the Site, in the event of a breach by the Sub-processor of any of its obligations under the GDPR or Applicable Law, the Technical Partner shall remain primarily responsible to the Site for the adequacy of data protection with respect to the GDPR.

2.4.5 Notwithstanding any consent or approval given by the Site, the Technical Partner shall at all times be primarily responsible to the Site for the actions, errors and omissions of any subcontractor to whom it discloses Personal Data and shall be liable to the Site for the actions, errors and omissions of such subcontractor in the same manner as if they were the Technical Partner's own actions, errors and omissions, to the extent that the Technical Partner would have been liable to the Site under this Agreement for such actions, errors and omissions.

2.5 Notwithstanding anything to the contrary contained in this Agreement, this Agreement shall remain in full force and effect for so long as the Technical Partner continues to Process any Personal Data.

3. **RECOVERABLE LOSS**

3.1 Without prejudice to Section 4 of the Consortium Agreement, the Parties agree that nothing in the Consortium Agreement shall preclude the Site from any proven direct Loss incurred as a result of the acts or omissions of the sole negligence of the Project Technical Partner in connection with this Agreement.

4. **INDEMNITY**

4.1 A Party (the "Defaulting Party") shall indemnify upon request and keep indemnified the other Party (the "Indemnified Party") from and against:

4.1.1 any and all penalties or fines imposed on the Indemnified Party by the corresponding data protection authority; the costs of any investigative, corrective or compensatory actions imposed by the Data Protection Agency or any other regulatory authority, or of opposing any potential or actual enforcement action taken by any data protection authority and/or any other regulatory authority;

4.1.2 any Loss suffered or incurred by the Indemnified Party, awarded or agreed to be paid by the Indemnified Party pursuant to a claim, action or challenge made by a third party against the Indemnified Party (including by a Data Subject); and

4.1.3 except to the extent that Paragraphs 4.1.1 and/or 4.1.2 apply, any Losses suffered or incurred, awarded against, or agreed to be paid by the Indemnified Party, in each case, to the extent arising as a result of a breach by the Defaulting Party (or its subcontractors) of this Agreement and/or their respective obligations under Data Protection Law,

in each case, to the extent that it arises as a result of a breach by the offending Party (or its subcontractors) of this Agreement and/ or of their respective obligations under Data Protection Law.

4.2 Subject to Clause 4.3, the maximum liability of any Party under or in connection with this Agreement or its subject matter, and however arising, whether for breach or negligence and whether in contract, tort, breach of statutory duty or otherwise, shall not exceed FIVE HUNDRED THOUSAND EUROS (500,000€).

4.3 For the avoidance of doubt, nothing in this Agreement limits or excludes the liability of any Party for any other liability which cannot by law be limited or excluded.

4.4 Neither Party will be liable for any indirect loss and damage, lucrum cessans, loss of income or profit or goodwill, loss of data and/or use. Any claim must be brought within one year since the act or omission took place.

5. **TERMINATION**

5.1 Notwithstanding anything to the contrary contained in this Agreement, in the event of a breach of this Agreement by the Technical Partner, the Site may terminate this Agreement and/or request that the General Assembly constituted in accordance with the Consortium Agreement terminates the Technical Partner's participation in the Consortium Agreement immediately upon written notice to that effect being given to the Project Partner.

6. **MISCELLANEOUS**

6.1 The parties shall comply with all laws, rules and regulations including, without limitation, all domestic and foreign export control and anti-corruption laws and regulations, applicable in the jurisdiction of the Site. The parties agree that access to the Data and results of the Project to a Party under this agreement is granted with the specific understanding and requirement that responsibility for ensuring compliance with all applicable export control laws and regulations are being undertaken by the parties. Both parties further understand and acknowledge their obligations to make a prompt report to each other and appropriate authorities regarding any access to or use of the Data and results of the Project hereunder that maybe in violation of applicable export control laws and regulations. In addition, each party hereby

agrees that no results of the Project, Data, know-how or other information or assistance furnished by a Party pursuant to this Agreement, or any product or revision thereof, shall be re-exported or otherwise used by other Party or its authorized transferees outside of that Party's principal domiciliary country in contravention of any applicable export control laws. These obligations shall survive any satisfaction, expiration, termination, or discharge of this Agreement or any other obligations.

AS WITNESS the hands of authorised signatories for the parties on the date first mentioned above.

SIGNED for and on behalf of **[SITE]:**	**SIGNED** for and on behalf of **[TECHNICAL PARTNER]:**
Name:	Name:
Position:	Position:
Signature:	Signature:
Date: Date:	

AS WITNESS the hands of authorised signatories for the parties on the date first mentioned above.

SIGNED for and on behalf of **[SITE]:**	**SIGNED** for and on behalf of **[TECHNICAL PARTNER]:**
Name:	Name:
Position:	Position:
Signature:	Signature:
Date: Date:	

COMMISSION IMPLEMENTING DECISION on standard contractual clauses between controllers and processors under Article 28 (7) of Regulation (EU) 2016/679 and Article 29 (7) of Regulation (EU) 2018/1725

Content to be inserted without change from

https://ec.europa.eu/info/law/law-topic/data-protection/publications/standard-contractual-clauses-controllers-and-processors

COMMISSION IMPLEMENTING DECISION on standard contractual clauses for the transfer of personal data to third countries pursuant to Regulation (EU) 2016/679

Content to be inserted without change from

https://eur-lex.europa.eu/eli/dec_impl/2021/914/oj

Annex 2: ADLIFE Checklist of organisational data protection and information security requirements

Introduction

This document is intended to facilitate all partners in the ADLIFE consortium to have consistent organisational, personnel and information security measures relating to the handling (processing) and communication of personal health data obtained as part of the project's research. Specific interactions between partners, such as the support provided by technical partners to pilot sites that are using and assessing technology components, are covered by separate agreements such as Data Processing Agreements.

All partners that may have contact with personal research data are expected to have undertaken a Data Processing Impact Assessment (DPIA), which is a kind of risk assessment, and from this to have determined if they need to develop any specific measures, safeguards and operational practices, that are additional to their normal organisational practices for handling personal data.

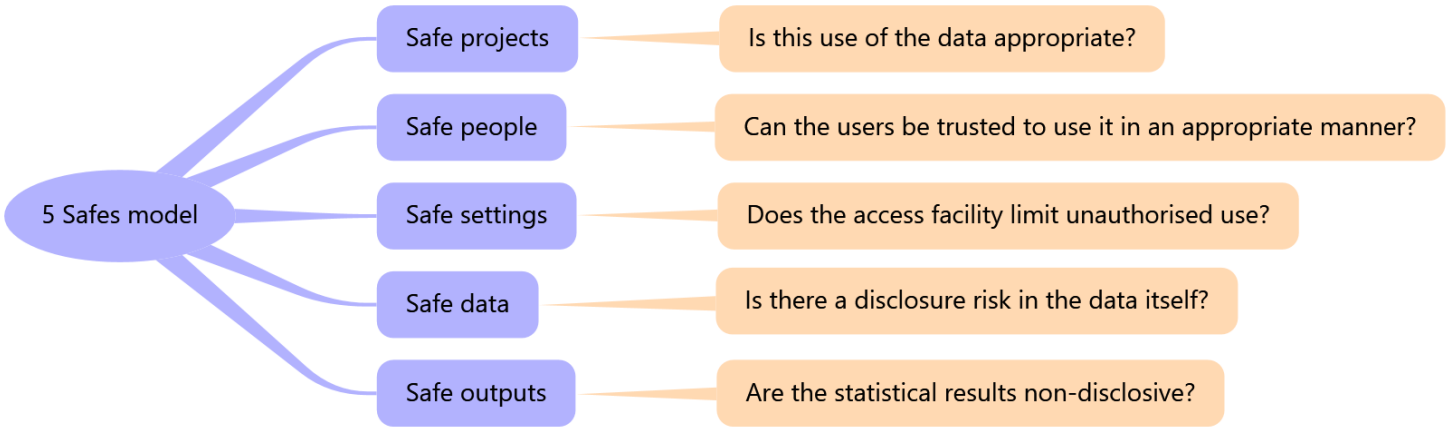
This document lists the practices and safeguards that would normally be expected in an organisation that handles personal data and is involved in health research projects that process personal data, such as health data collected from our recruited patients and evaluation feedback we obtain from patients, caregivers and health professionals through questionnaires and interviews. This document is not itself a code of practice or an information security policy, deliberately, because it is assumed that all partner organisations already have such documents or equivalent documents. This document is a kind of checklist against which existing documents should be validated, and consideration should be given to adding anything specified here that is missing from those existing documents.

The objective is therefore not to impose new rules or policies on organisations, but to enable them to verify that they are able to meet a project-wide universal set of data protection and information security requirements. This will allow all partners to exchange personal data with each other in confidence that similar standards are adopted throughout the consortium amongst the partners who may need to handle personal health or research data.

NOTE

Where this document mentions 'Article X' then this is a reference to an Article in the EU General Data Protection Regulation (GDPR).

The 5 Safes Model



The '5 Safes' model is a useful way of initially considering security issues. Not all issues will necessarily fit under a single one of its five headings, so may appear under just one heading or as slightly differently-emphasised requirements under two or more headings. For example 'auditing of user activity' could appear under 'Safe People' or 'Safe Settings' as:

- **Safe People**: Users must be made aware that their activity when using organization systems will be monitored and possible misuse investigated, potentially leading to dismissal or other sanctions
- **Safe Settings**: Administrators must run and review regular reports to check on potential misuse, following established procedures for further investigation of any possible incidents indicating misuse

This 5 Safes model has been used as the high level structure of this checklist.

Suggested Approach

Safe projects			
GDPR compliance			
	• Dataflow/Controllership	Establish which parties are controllers or processors for each processing step Each controller needs to ensure that the relevant processing is recorded on their Article 30 'Records of processing activities'	Project leads for each organization will need to inform their organisation's Data Processing Officer of the intended processing.
	• Transparency Notices	Article 13 or Article 14 requirements will apply to all controllers, but further there is a general obligation for 'transparency' so project websites should also explain what is happening to personal data and where to contact the various controllers	
	• Processor Agreements	Ensure that there is a Data Processing Agreement for each processor as per Article 29	
	• DPIAs	Each controller will need to make their own risk assessment (which may or may not be a formal DPIA, though this is likely to be necessary in the realm of healthcare data use at scale)	Developing a project wide common DPIA (or equivalent document) ensures a common understanding of the overall position as well as consistency of presentation to the public

Risk Management			

	• Risk registers	It is good practice to maintain a project risk register (probably part of the Description of Action) as well as an IT Security Risk Register for each data operation (at controller or processor)	This is likely to include:
			• Risk rating grid
			  • Risk appetite criteria
			  • Risk category criteria

	• Incident Management	Each organisation needs to ensure that it has effective incident management processes and procedures covering: • Incident detection • Incident containment • Incident investigation • Lessons Learned • Board-level reporting	There also needs to be effective coordination between partners whenever an incident occurs. Maintaining a clear list of DPOs and Information Security Officers for each organisation is advisable, possibly best held by the project co-ordinator.
	• Business Continuity Planning (BCP)	It is good practice for each organisation to ensure that it is resilient against adverse events, including information security incidents, with appropriate backup and recovery mechanisms as well as effective and timely reporting to supervisory authorities where necessary.	

Safe people			
Acceptable Use Policy (AUP)		All institutions should have a formal AUP which all staff are required to read and accept as part of their employment	This is a minimum requirement to ensure that staff know what may or may not be done when using organizational systems, equipment or when processing data that is, or might potentially be, personal data.
	• Confidentiality	The need to protect corporate information in the interests of the organization The need to protect personal information in the interests of the individual as well as to preserve trust in the organization and healthcare research generally.	
	• Use of Corporate systems	Making clear what is appropriate use and what is not (e.g. not trying to locate records of friends or family – including one's own). Should be clear what limits there are on using corporate systems for personal use, e.g. email or web-browsing	This may need to differ between ordinary users and those with special privileges over the management of systems
	• Unacceptable behaviour	Being explicit about behaviours which would lead to disciplinary sanctions, including bringing the organization into disrepute, or to criminal charges.	This should be aimed at educating staff about the law and the conditions of their employment as well as making clear what levels of oversight may be applied to detect unacceptable behaviours
	• Reporting incidents or suspicious behaviour	Covering types of data or information system related incident and how to report any suspected incidents or misuse	This should link to broader incidents, such as physical security or health and safety matters

Engagement & Training		Organisations need to provide effective training about IT security, engaging all staff in safe practices, rather than relying only on specialist staff to have that knowledge	

	• Awareness training	Staff need to be given initial training about security and confidentiality issues and what organizational policies and procedures apply	Training may be through reading, on-the-job instruction, online or classroom sessions, but it is important to ensure that all staff with access to healthcare data have this basic training and that it is kept up-to-date. Topics are likely to include:
			• Legal requirements
			• Organisational & Employment requirements
			• Threats, such as
			  • Counter-measures, e.g. Screen-locking
			  • Social Engineering, e.g. Phishing
			  • Physical Security

	• Training records	It is important that the organization can demonstrate (evidence) that all staff have been appropriately trained and that refresher training and topic understanding can be arranged as needed	
	• Established processes and procedures	Processes and procedures to support all of these requirements need to be documented and properly implemented – as part of this staff need appropriate training on what these procedures are and when to use them	Simply having documented procedures is not enough; staff need to know how to react to a potential incident and be able to handle it in a timely and efficient manner according to those procedures.

	• Knowledge testing & Process tests	Possibly as part of formal training or as a separate exercise, it is important that staff know their particular roles and can work cohesively to manage any incident	Scenario testing is a useful desk exercise, both to explore possible implications of adverse events as well as to test staff understanding of how they are supposed to respond
	• Key staff skills & training	Key staff (e.g. system administrators and auditors) will require specific additional formal training when they take up a new role. On-the-job training is unlikely to be sufficient, especially for unusual events or circumstances.	
Identity & Access Management			
	• Joiners, Movers, Leavers (JML) procedures	It is important that staff taking up a new role have the relevant access rights to the information they will need to fulfil their role; equally, when their role may change that their access rights are altered to reflect the new role, and when staff leave that their access rights are removed. Ideally, system access privilege assignment/revocation should be integrated with human resources (HR) processes.	It is important that access rights are only granted on an 'as needed' basis, rather than broad rights given regardless of role, though there will be some systems to which all staff may need full access. Those personnel changing role may need to retain access rights during a transition period in order to provide continuity, but these retained rights must be routinely reviewed and not retained indefinitely. When staff leave, system accounts may need to be retained for audit trail purposes, but actual access should be prevented.

	• Multi-factor authentication	A typical logon process involves an ID (email or username) and a password – the password is a single-factor authentication of the identity of the person (as given by the ID). The logon/password combination gives a certain level of assurance that the person is authorized to access the system. Having two distinct authentication factors greatly reduces the chance of an attacker gaining access.	Authentication factors are typically: • Something you know (e.g. password, answer to security question) • Something you have (e.g. key-fob, mobile phone, or SmartCard) • Something you are (e.g. fingerprint, iris-scan, face recognition)
	• Password policy	Organisations need to determine what is a minimum level of password strength and complexity and ensure staff know how to produce a safe but memorable password.	Requiring too complex a password is likely to lead to increased login failures and technical support calls; similarly, requiring users to change passwords frequently is likely to them using the weakest possible password and also needing technical support calls without significantly improving overall security
	• Time-outs/caching	Systems will usually have various settings for screen time-outs, inactivity log-outs, etc.	These settings need to be risk-assessed to ensure that both security and usability are not compromised – there is usually some trade-off between these two requirements.
	• Privileged access	Administrator or similar accounts with special privileges, e.g. granting access to others, direct access to data, or ability to amend or delete core files. These need careful consideration and may sometimes need to be activated only on a temporary basis if needed for an exceptional event.	There can be a need to balance security with responsiveness to emergencies. Certainly, close monitoring of use of such accounts is critical as well as additional training to ensure that staff know they will be monitored closely.

	• Geographic restrictions	Remote access is increasingly common, but it is possible to limit or block access from other countries, particularly outside the EU.	This may need to be tailored both by systems and role of user; many users should not need access outside their 'home' country, but others may need to travel abroad on business, but may not then need access to sensitive data or systems.
Logging & Monitoring			
	• Activity reporting	It is a clear security requirement that not only are access controls and suitable IT security measures put in place, but there are also checks that these are effective. Anti-malware software should report possible events encountered, and systems should maintain user activity logs ('audit trails') which should be pro-actively interrogated rather than just reviewed in the light of a possible incident.	Typical reports might include: • Failed logons • Out-of-hours activity • Inactive or redundant user accounts • Activity on supervisor accounts These reports may highlight events or accounts that need more detailed investigation.
Mobile & bring your own device (BYOD) security			This topic probably falls outside the remit of most research projects, though may be relevant for individual organisations, so included here for completeness only
	• Remote working	Process and procedures for staff accessing systems remotely	Relevant perhaps for instances where access is being granted to a statistical query system for remote access by licensees.
	• Portable drives (e.g. USB sticks)	Process and procedures for staff using portable storage devices	Easiest to simply forbid. Users should not be allowed to make copies of any data, except through formally approved mechanisms (e.g. institutional back-up systems).

Key roles assigned			
	Proper oversight and handling of incidents require that key roles are clearly assigned to individuals, who are also trained to carry out these roles, either with specialist training or specific training in relation to the role.	Typically, there will be at least these roles: • Senior Information Risk Owner (SIRO) role • Information Asset Owners • Data Protection Officer (DPO) role • Information Security Officer role	These roles may be assigned in addition to other roles, e.g. SIRO will typically be a senior board member with other responsibilities, but it is important that individuals have sufficient time, resources, and authority to carry out these additional roles. Staff should have access to information about who these role-holder are, in the event that any of them need to be contacted (in the event of a query, incident, data breach etc.)

Safe settings			
Architecture & Configuration			
	• Information Security Management System (ISMS)	Need to document all information security policies and procedures as part of an overall Information Security Management System (ISMS)	The ISMS needs to be appropriate to the scale of the organisation, but needs to cover the major areas of concern and be designed as an overall system of information security to avoid exploitable weaknesses
Asset Management			
	• Asset Registers	GDPR Article 30 requires a 'Register of Processing Activities' (RoPA) which may be part of a wider Asset Register for gauging information risks (see earlier entries)	The finance department may maintain an 'asset register' of what has been purchased, but this may be limited and not include 'information assets' which may be crucial to operations and need to be adequately protected
	• Legacy system management	It is important to identify any assets that are no longer updated or maintained as these can pose security risks	This is particularly true for commercial software or hardware that is no longer actively supported by the supplier because a maintenance licence or agreement has ended
Vulnerability Management			
	• Anti-malware	Anti-virus software and equivalent	
	• Network Perimeter defences	• Firewalls • Email gateway security • Web gateway security	
	• Patch Management	• Supplier assessment • Customer assessment	It is important that software 'patches' are applied in a timely and effective manner

	• Supply Chain Security	Part of overall Business Continuity Planning	
Help Desk			
	• Incident Reporting	As noted above under Incident Management, procedures need to be effective in identifying and reacting to potential data security incidents which may need to be notified to Supervisory Authorities (Article 33), especially given the 72-hour deadline (regardless of actual working hours)	
	• Rapid Reaction Taskforce	Part of Incident Management is the ability to set up a Rapid Reaction Taskforce (or equivalent team) to address and resolve possible incidents. This requires pre-established procedures for selecting and informing participants of an incident so that a coordinated and effective response can be made in a timely fashion.	

Safe data			
Data Security			
	• Encryption	• At rest • In transit	Encryption of personal data assets should now be a standard requirement for both data storage and data transmissions.
	• Data Retention & disposal schedule	Need to determine business requirements for data retention and then legal basis for holding; should also consider form in which data is to be held (pseudo/anonymized)	Should also consider situation of data subject makes request to have data erased and whether request can be upheld or not – helps test rationale for retention. Need to consider user and contact data as well as mainstream data holdings
	• Back-up & recovery policies	Need to be able to recover in case of system failure or attack (e.g. ransomware) Need to balance costs of storage with recovery time	Also need to test systems, so can data be recovered in various scenarios; what about transition to new systems?
	• Data deletion & destruction process	There need to be established processes for deletion of data (records or datasets) when no longer required. This may require purging of records from datasets and/or purging of datasets from back-up systems	This has been covered above under Contractual Terms, but here is about having procedures to effect this. This is a danger of retaining electronic data 'just in case' rather than having a formal process for removing data which is no longer realistically required.

Safe outputs			
Data Release (or data sharing) Approvals process			
	• Application process defined and published	There needs to be a clear application process, usually involving an Application Form (paper or electronic) whereby an applicant justifies their request to have access to the data	This process should be published, ideally with likely timescales for consideration of requests. An appeals process might also be considered, though most data-holders allow revised applications to be made, so the need for a formal appeals process may be minimal. NOTE: In the case of a European project the Consortium Agreement may specify the permission and terms where by research (foreground) data is to be shared within the consortium, and with external parties such as sub-contractors and third parties.
	• Acceptance criteria	There need to be clear criteria of what is or is not acceptable in terms of an application to be both fair and to save unnecessary applications.	There may be an element of judgement in applying these criteria, so there needs to be a clear process of determination with more problematic applications being referred to senior managers or an independent board for determination.

Contractual restrictions			
	Access Licence or Data-sharing Agreement	There should be a formal contractual document (which may be predetermined or variable) which determines how the accessor organisation or data recipient shall protect and use the data received or facility made available. It is likely to include at least the following elements (as well as more usual contractual terms for data delivery, restrictions on use, and termination of the agreement/licence):	
	• No further release	It is a usual requirement that the data provided will not be passed onto other parties where licence conditions may not apply or particularly where EU data protection law (or equivalent) may not apply.	Special allowances may be made for submissions to regulators, including the relaxation of small cell-counts (see below) though with the further restriction on actual publication of the unsuppressed data. Where sub-licensing is permitted, then sub-licences must be subject to the same limitations as the main agreement or licence (excepting permission to further sub-license).
	• Destruction at end of project/fixed term	Both to meet Article 5(1)e 'Storage limitation' as well as limiting the effective terms of the licence or agreement, there should be a requirement that dataset received or access for analysis should be limited by a fixed period, after which any data received shall be destroyed.	This only applies to individual-level data and would not normally apply to derived data of an aggregate or statistical nature, though intellectual property (IP) restrictions might still apply.

	• Small cell-count restrictions on publication	Tabular aggregate data should be restricted from containing small cell-counts which might permit reconstruction of some aspect of the individual-level data which might reveal some personal information.	Applicants are likely to want to share or publish analytic results or statistics derived from the shared or accessed data, but open publication of such data may permit 'reconstruction' attacks
	• No attempt to re-identify	Anomalous records should not be investigated by recipient – to avoid accidental re-identification	Any suspected data errors, especially poor quality de-identification, should be referred back for investigation rather than refined by recipient
	• Staff training in confidentiality & contract restrictions	Need to ensure restrictions are applied in practice	Can be danger that only contracts departments reads the contract restrictions
De-identification			
	• Data minimisation	Article 5(1)c requires that only the minimum personal data required to achieve the purpose is used. It is also simply good IT practice as well.	This requires that either any data accessed or shared is anonymised (to prevent re-identification) to fall outside GDPR requirements, or that data is more generally limited in scope (range of records, range of attributes, value grouping/blurring, re-coding (including use of pseudonyms)) before it is made available to the licensee
	• Pseudonym management	Different licensees should only receive data under different pseudonym-coding schemes to avoid the possibility of the different datasets being linked and possibly identifying individuals.	This requires that records are kept of what coding schemes (e.g. seed values) were used for each data release; this may also allow longitudinal incremental releases of data to the same licensee over time without requiring a further release of the entire dataset (and would also require licensee to delete previous version).

	• Risk Assessment	As per DPIA requirements (Article 35), risk assessments will be needed either for individual releases or the generic process for de-identification of releases.	Without such a risk assessment, it may be hard to meet the Recital 26 ‘reasonably likely’ test for showing that data is outside the scope of GDPR. A project-based DPIA may well include the necessary risk assessment for data releases.