



Deliverable 7.2

Project ID	654241
Project Title	A comprehensive and standardised e-infrastructure for analysing medical metabolic phenotype data.
Project Acronym	PhenoMeNal
Start Date of the Project	1 st September 2015
Duration of the Project	36 Months
Work Package Number	7
Work Package Title	Privacy and Ethics
Deliverable Title	D7.2 Report on policies and procedures for sensitive human data management
Delivery Date	M6
Work Package leader	ICL
Contributing Partners	ICL

Authors: Tim Ebbels, Robert Glen, Namrata Kale, Kenneth Haug



Contents

1. Executive Summary	3
2. Project Objectives	3
3. Detailed report on the deliverable	3
3.1. Background	3
3.2. PhenoMeNal ELSI workshops	4
3.3. Associated documents	4
4. Delivery and Schedule	5
5. Background information	5
6. Annexes	7
6.1. PhenoMeNal Guidance on Handling Sensitive Data	7
6.2 PhenoMeNal Terms of Use	9
6.3 ELSI Technical Guidelines	10



1. Executive Summary

Patient and research subject data is very sensitive, and it is paramount to establish a robust governance framework for overall information management including sensitive data. The PhenoMeNal e-infrastructure will be able to cope with data generated from comprehensive clinical, genotypic, 'omics and analytic sources including medical records, electronic health records, clinical measurements, genotypic data, phenotypic data from tissue and biofluid analysis, image and pathology data. All data collected and held within the project will comply with all local laws, regulations and ethics. A guidance document on handling sensitive human data has been produced ([Annex 1](#)). The key points of the guidance are as follows.

- By default, processing and analysis of sensitive data will take place on local infrastructure at the host institution.
- Responsibility for data will be with the host institution/data provider.
- All sensitive data must be linked or unlinked anonymised.
- PhenoMeNal systems should support the principles of informed consent, including recording the country to which consent applies, ability to withdraw consent, potential for recontact of participants and time limits on consent.
- The project should consider recruiting an independent ethical advisor to join the governance committee.

2. Project Objectives

No.	Objective	Yes	No
1	Develop appropriate policies, procedures and management accountability and structures to provide a robust governance framework for information management.	x	

3. Detailed report on the deliverable

3.1. Background

The PhenoMeNal infrastructure is to be designed to accommodate sensitive human data. In particular, it will deal with sources including medical records, electronic health records, clinical measurements, genotypic data, and phenotypic data from tissue and biofluid analysis, image and pathology data. Key elements of the technical design will be influenced by the need to uphold ethical objectives, which comply with



national, EU and international law and data protection principles. These ethical principles will apply both to the storage but also to processing of data on the PhenoMeNal infrastructure. This deliverable summarises the process by which the project has come to an ethical framework which will ensure the infrastructure supports proper use of sensitive human data.

3.2. PhenoMeNal ELSI workshops

Two workshops on Ethical, Legal and Social Issues (ELSI) were held at Imperial College London (ICL):

- I. Workshop I, 20th November 2015, 22 participants, 5 invited speakers (ELSI experts, external to the project). Objective: to discuss best practice in handling sensitive human data.
- II. Workshop II, 27th January 2016, 11 participants (project partners, plus BiomedBridges representatives). Objective: to draft project guidance on best practice in handling sensitive human data.

See deliverable D7.1 for a report on the first workshop. The current deliverable (D7.2) and D7.3 (data provider form) form the outputs from the above workshops.

3.3. Associated documents

The workshops and accompanying discussions have produced three working documents addressing ethical issues within the PhenoMeNal project. These documents are intended to form a framework, which PhenoMeNal developers can follow to ensure that the infrastructure they build complies with and supports ethical standards. The documentation is understood to be 'live' in that it will be continuously under review and updated as and when necessary throughout the lifetime of the project.

- I. **PhenoMeNal Guidance on Handling Sensitive Data:** This document is intended to be a brief overview of the main principles and best practices to be followed when handling sensitive data within Phenomenal. It will influence the broad design of the infrastructure, including workflow logic, access management and data storage frameworks. See Annex 6.1 [PhenoMeNal Guidance on Handling Sensitive Data](#).
- II. **PhenoMeNal Terms of Use:** This document is a short set of terms to which PhenoMeNal users must sign up in order to use the infrastructure. It defines the responsibilities of data providers and data consumers/users, including



ethical obligations with respect to sensitive human data. See Annex 6.2 [PhenoMeNal Terms of Use](#)

- III. **PhenoMeNal ELSI Technical Guidelines:** The broad ethical principles outlined in the ethical guidance document will have implications for the technical design of the PhenoMeNal infrastructure. These are outlined in more detail in the technical guideline document. This is a working document, which will be updated throughout the project as different software, and technical deliverables are produced. See Annex 6.3 [ELSI Technical Guidelines](#).

4. Delivery and Schedule

The delivery is delayed: No

5. Background information

Patient and research subject data is very sensitive, and it is paramount to establish a robust governance framework for overall information management including sensitive data. The PhenoMeNal e-infrastructure will be able to cope with data generated from comprehensive clinical, genotypic, 'omics and analytic sources including medical records, electronic health records, clinical measurements, genotypic data, phenotypic data from tissue and biofluid analysis, image and pathology data. Primarily, all data collected and held within the project will comply with all local laws, regulations and ethics. All personal information will be processed in accordance with accepted Data Protection Principles outlined above. Responsibility for data will be with the host institution/data provider.			
Work package number	WP7	Start date or starting event:	M1
Work package title	Privacy & Ethics		
Participants	EMBL-EBI, ICL, IPB, UB, UL, UOXF, SIB, UU, BBMRI		
Objectives			
Objective 7.1 Develop appropriate policies, procedures and management accountability and structures to provide a robust governance framework for information management.			
Objective 7.2 Raise awareness of information governance within the consortium and assure on going compliance			



Objective 7.3 provide a forum for information exchange on best practice in clinical data sharing and disclosure.

Task 7.1 Organise workshops involving Scientific and Technological Advisory board (STAB) and the Management Committee (MC) and other necessary players to devise best practices in handling sensitive human data, taking into account National and Institutional legal policies. (ICL, EMBL-EBI)



6. Annexes

6.1. PhenoMeNal Guidance on Handling Sensitive Data

PhenoMeNal will develop advanced software pipelines for efficient and secure handling of metabolomics data. A major consideration is that the software architecture and use of the applications and data takes into account ELSI considerations. This document provides guidelines on how data and software for PhenoMeNal can safely facilitate the storage and analysis of sensitive data. These guidelines will evolve during the course of the project as the legal and ethical frameworks evolve, and the software is developed. PhenoMeNal is intended to handle a wide range of data analysis scenarios (from open data to clinically sensitive data). The purpose of this document is to inform the developers (of PhenoMeNal) and data providers of the ELSI considerations, which must be taken into account, and conditions under which PhenoMeNal resources are provided.

PhenoMeNal is developing workflows composed of open-source components developed by the worldwide metabolomics community. One of the major achievements of PhenoMeNal will be to scale those components to work with big data and to ensure that they interoperate. The components themselves will remain under the control of their original authors unless they decide otherwise. Due to this working principle, the PhenoMeNal consortium cannot guarantee that the components are principally safe to use in unprotected compute environments or that no backdoors exist that allow unauthorised access to the data submitted to the PhenoMeNal pipelines. For this reason, our general guidance to users with sensitive data will be to use the PhenoMeNal workflows on their own, local, secure cloud.

This report should be read in conjunction with the PhenoMeNal Terms of Use, which covers 1) the scope of the PhenoMeNal infrastructure, 2) The measures PhenoMeNal takes in order to make computing secure and 3) the precautions that users need to take in order to use the PhenoMeNal infrastructures securely.

Due to the use of a large variety of externally developed software components, which can be run on publicly accessible clouds, it is impossible to ensure full security unless certain precautions are taken. PhenoMeNal will take suitable measures to educate users of infrastructure about potential security threats and countermeasures. As already pointed out in the grant proposal, the **default solution to ensure security will be bring the compute to the sensitive data**, i.e. users will run the portable PhenoMeNal infrastructure in a secure local cloud, close to their sensitive data, so that the data does not have to leave the protected area.



To help users understand the complex landscape of ELSI in the context of human clinical data, PhenoMeNal is working with Biobanking and BioMolecular resources Research Infrastructure (<http://bbmri-eric.eu>, BBMRI-ERIC) and Coordinated Research Infrastructures Building Enduring Life-sciences (<http://www.corbel-project.eu/home.html>, CORBEL) WP7, who are developing ELSI services on the BBMRI infrastructure (coined “ELSI common services”). As a starting point, we are referring our users to the Ethical Governance framework developed by BioMedBridges (<https://zenodo.org/record/11952?ln=en#.VqjYn4Tbxg0>)

Based on the BioMedBridges Ethical Governance framework, a number of points are clear which need to be taken into account when designing and implementing the PhenoMeNal infrastructure.

- I. There should be a clear differentiation of roles between a) data providers and b) data users. Each role will need to be handled differently in respect of access, use etc. In some cases a given entity may take on both roles.
- II. Responsibility for all data stored, accessed or otherwise made available via PhenoMeNal rests with the data provider. This must be clearly stated on the data provider form signed by those contributing data to the project.
- III. All sensitive data must always be linked (this is anonymous to the people who receive and hold it, but contains information or codes that can allow identification) or unlinked (this contains no information that could reasonably be used for any identification) anonymised. This ensures that confidentiality can be maintained. In the case of linked anonymisation, the codes/keys needed to identify a participant must not be stored on the PhenoMeNal architecture or accessible through it.
- IV. Details of the informed consent used in the collection of data must be made available. The information must include whether consent covers storage of data in a country other than that in which the data were obtained. It is recommended that data be labelled according to whether it is ‘broad’ or narrow, and also by to what extent communication with participants is required in cases of clinically relevant findings, or obligations to third parties.
- V. The responsibility for providing feedback to participants resides with the data provider, along with fulfilling any obligations to third parties. These obligations must be stated clearly on the data provider form.
- VI. Many consent forms allow the participant to withdraw their consent at any time, which implies removing their data from the system. This is not possible for unlinked anonymised data. For linked anonymised data, the infrastructure should support withdrawal of consent. However, the data provider form must



make clear that it may not be possible to withdraw data from researchers who have already accessed it.

- VII. Any time limitations on consent or assurances to third parties must be made clear on data provider form and clearly labelled in the data.
- VIII. The project should consider appointing an independent ethical advisor to join the governance committee. This advisor would monitor compliance of the partners with the ethical framework and evaluate infrastructure for ethical compliance. Thus the infrastructure must support such testing and evaluation.

6.2 PhenoMeNal Terms of Use

The following is the first version of the PhenoMeNal Terms of Use which will be made available on the PhenoMeNal website and associated infrastructure pages. Users will be required to accept these terms before accessing the PhenoMeNal infrastructure.

Phenomenal Terms of Use version 1.0

PhenoMeNal is an integrated, secure, on-demand service-driven, privacy-compliant and sustainable European e-infrastructure for processing, analysis and information mining of metabolomics data. The project has been designed to enable maximum benefit from research by making data as accessible as possible to the research community, while protecting the interests of participants from whom the data originate with regard to Ethical, Legal and Social Implications (ELSI) and within the scope of their consent. These Terms of Use reflects PhenoMeNal's commitment to provide this service and impose no additional constraints on the use and transfer of the contributed data than those provided by the data owner.

- I. All users have an obligation of confidentiality and must conform to data protection principles to ensure that data is processed in compliance with the legal and ethical requirements.
- II. The data owners must ensure that they have sought and obtained, where necessary, all appropriate approvals, ethical and legal, for the data collected.
- III. For animal data, the data owner must ensure that national guidelines for their welfare and care during the collection of data have been followed.
- IV. PhenoMeNal does not guarantee the accuracy of any provided data.
- V. PhenoMeNal has implemented appropriate technical and organisational measures to ensure a level of security which we deem appropriate, taking into account the sensitivity of data we handle. However, the data provider holds sole responsibility for the usage and distribution of data.



- VI. Computing of personal and sensitive data on PhenoMeNal infrastructure should be run internally by the users on their secure cloud infrastructures under appropriate firewalls. PhenoMeNal will not hold any liability for any loss or damage to data.
- VII. While we will retain our commitment to privacy of sensitive data, we reserve the right to update these Terms of Use at any time. When alterations are inevitable, we will attempt to give reasonable notice of any changes by placing a notice on our website, but you may wish to check each time you use the website. The date of the most recent revision will appear on this, the 'PhenoMeNal's Terms of Use' page. If you do not agree to these changes, please do not continue to use our services. We will also make available an archived copy of the previous Terms of Use for comparison.
- VIII. Any questions or comments concerning these Terms of Use can be addressed to: phenomenal-help@ebi.ac.uk

6.3 ELSI Technical Guidelines

The principles outlined in the main guidance document have implications for the technical design of PhenoMeNal infrastructure. These implications are outlined at a high level in the PhenoMeNal ELSI Technical Guidelines document.

(<https://docs.google.com/document/d/1yc4UrFBiDAckdoZ5m-dLZ8xB13jFzNze--xTY4xvPGA/edit>)

This is a working document, which will be updated throughout the project to provide developers with up to date advice on issues they should consider when developing the infrastructure.