HumanE AI Net:

The HumanE AI Network

Grant Agreement Number: 952026
Project Acronym: HumanE AI Net

Project Dates: 2020-09-01 to 2023-08-31
Project Duration: 36 months

D11.3 M Ethics requirements

Author(s): George Kampis, DFKI
Contributing partners: CEU
Date: May 4, 2022
Approved by: Paul Lukowicz
Type: Report (R)
Status: final
Contact: George.Kampis@dfki.de

Dissemination Level		
PU	Public	X

DISCLAIMER

DOCUMENT INFO

0.1 Authors

Authors	Institution	e-mail
George Kampis (GK)	DFKI	George.Kampis@dfki.de
János Kertész (JK)	CEU	KerteszJ@ceu.edu

0.2 Document History

Revision		
Date	**Lead Author(s)**	**Comments**
30.09.2020	GK	Empty template
12.02.2022	GK	Initial draft
13.03.2022	GK	Edits and additions
14.04.2022	GK	Continued additions
22.04.2022	GK	Continued additions
03.05.2022	JK	Edits, suggestions
03.05.2022	GK	Final formatting

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This is the “M Ethics Requirement” Deliverable. It is part of WP11, introduced by the EC. This Deliverable is a report containing a risk assessment and details on the measures to prevent the misuse of specific research findings. As such, it discusses the necessary procedures that lead to ethical data discovery and use.

Note that considerable part of the research in the project Humane AI NET is carried out in Micro-Projects (MPs). An MP is worked out following a successful proposal as a response to a semi-open call within the project. As submission and approval of MP-s has been continual, and will remain so, we cannot take responsibly for the data discovery and use practices therein. Therefore, the formulation of these ethics requirements should be understood as directives.

1. INTRODUCTION

This Deliverable is a report containing a risk assessment and details on measures to prevent the misuse of data and research findings including accidental ones. It clarifies what measures are being taken to ensure that outputs from HumanE-AI-Net will not be misused by malicious actors.

Note that considerable part of the research in the project Humane AI NET is carried out in Micro-Projects (MPs). An MP is worked out following a successful proposal as a response to a semi-open call within the project (detailed elsewhere, e.g., in D10.5). As submission and approval of MP-s has been continual, and will remain so, we cannot take responsibly for the data discovery and use practices therein. Therefore, the formulation of these ethics requirements should be understood as directives.

2. RISK ASSESMENT RELATED TO ETHICS

We provide a description here on the basis of the proposal.

Section 5.1.3 of the Proposal states that “Due to the nature of the proposed research, the achievement of HumanE-AI-Net tasks may require the processing of sensitive personal data, including ... user-mobility data, online social network and other web and internet data, [etc.]... These [viz. online data] are the only forms in which humans and human data enter the HumanE-AI-Net project.” (proposal B2, p 165.)

This suggests that non-project personnel should NOT be recruited as research participants or volunteers in any case.

On the other hand, section 5.1.6.3.a ("Informed consent") states: "A copy of the informed consent form and information sheet must be provided to the EC and participants must be informed about the accidental findings policy” (proposal B2, p 168). The wording reflects the potential existence of working with humans (eg. patients in some MPs) and the existence of an informed consent form.

Indeed, outlines of several MPs look as they may involve external human participants - possibly including hospital patients - in their particular activities. In terms of WPs and Tasks,

task T2.3 proposes “the development of tools to observe and assist people with different levels of expertise engaged in solving problems in domains such as chess, math, or circuit design” (proposal B1 p 49). A microproject in WP3 envisages “a scenario in which an AI system is part of the daily life of a chronically ill patient (e.g., diabetes).”, with measurements of “how the user perceives the system.” (proposal B1, p 52.) Another WP3 MP envisages “a setting where [a] meeting is being recorded by a spatial microphone grid and possibly also videorecorded from several angles.” (ibid.) An MP in WP5 - Explanatory tool for Clinical Analysis of Patients - proposes “targeting a variety of stakeholders (practitioners and patients), (proposal B1, p 57.)

Upon the collection of sensitive personal data, the proposal states (proposal B2, p163.): "the achievement of HumanE-AI-Net tasks may require the processing of sensitive personal data, including but not limited to the analysis of user-mobility data, online social network and other web and internet data, the handling of personal data in web pages, among others." And on p165: "HumanE-AI-Net tasks may require the processing of sensitive personal data (..)."

However, in section 5.1.4.3 (Protection of personal data and algorithms), it is stated: "The data collected and processed within the work packages and activities cannot be considered as sensitive data." (proposal B2, p167.)

On observation and tracking of participants, the discussion in proposal B1, pp 46ff suggests that some observation/tracking of individuals may be involved.

Ethical compliance processes concerning the processing of personal data are described in detail in the ethics self-assessment below. It is now confirmed that these measures will apply to personal data in relation to dissemination activities (workshops, summer schools, etc.) as well as to personal data that are involved in the microproject AI systems developed in work packages within the Project.

The slight inconsistency due to the open-ended MPs is handled in the HumanE-AI-Net project and in this document.

A related suggestion is to request an approval of the proposals from the local Ethical Committee of the research site (university or research institute) for MPs involving human subjects in any form. This should be attached upon submission and could be simple (“no sensitive data are collected”) or in detail if necessary.

3. ETHICS SELF-ASSESMENT (AND GDPR COMPLIANCE)

The text below reflects Section 5.1.1 of the proposal.

The HumanE-AI-Net consortium is fully aware of the ethical implications of the proposed research and respects the ethical rules and standards of RIA, and those reflected in the Charter of Fundamental Rights of the European Union. Generally speaking, the ethical, social and data protection considerations are crucial to this project and will be given all due attention.

The project will gather innovative and proactive responses to the structural problems we are seeing with social, and cultural uses of AI and data analytics. Examples include bias in automated mechanisms for hiring new workforce, discrimination in automated bank credit

decisions, unexplained treatments in medical AI, among several others. This will ensure that our research not only develops best practices and resources for practitioners of social and cultural AI and data analytics, but also that it facilitates a better informed, wider engaged, and more equitable participation and impact for such practitioners. In this respect, HumanE-AI-Net will become an essential hub for European and international researchers working across numerous fields, reporting annually on the best-practices, emerging trends, and innovative approaches. These legal and ethical requirements are also designed to ensure the proper professional behavior in compliance with accepted codes of conduct, current law, and upholding high moral standard of good scientific practice. As such, HumanE-AI-Net will pursue the EU views on Responsible Research and Innovation, the privilege of scientific research, and especially the values and norms of EU Data Protection law, inspired to strengthen the protection of personal data as a fundamental right, combined with boosting the free flow of personal data as a common good.

To drive the project along this path, HumanE-AI-Net will maintain and improve two ethical and legal boards: *The project advisory board* and the *the board for operational ethics and legality* (BOEL) with experts in IT law, IT ethics, and data protection, as well as experts in various disciplines such as Computer Science, Digital Humanities, Philosophy, Political Science and Sociologists along with external specialists and stakeholders.

All partners in HumanE-AI-Net will adhere to the Charter of Fundamental Rights of the European Union and to data protection legislation, as well as overarching ethical guidance such as the European Code of Conduct for Research Integrity. We also expect project researchers to adhere to the ethical commitments contained in their professional and institutional codes of conduct.

The project will undertake coordination and investigations that involve collection and sharing of personal data. The overriding principles will be a) active informed consent and b) appropriate ownership of data. The project will manage such data in accordance with the GDPR, local laws and institutional requirements. In addition, HumanE-AI-Net's partner organisations will be responsible for ensuring all ethical principles relating to their country and institutional context are adhered to. In some cases, this may require additional ethical permission. Thus, where relevant and/or necessary, academic researchers in the Consortium will submit their particular research to their institution's research ethics committee for ethical approval. In case of need as judged by the applicants, ethical approval in advance via the Board for Operational Ethics and Legality will be sought (BOEL). This consortium involves partners who are experienced academics and experts with a history of conducting research in the kinds of settings described in this research project.

While the General Data Protection Regulation (GDPR, https://gdpr-info.eu) intends to modernize the legal framework for the processing of personal data, HumanE-AI-Net aims at answering new questions on the scope, interpretation, and application of the GDPR along with the expanding role of AI, machine learning and data mining. Indeed, data processing via AI, big data mining, data analytics in the context of scientific research pose renewed concerns on legal issues, in particular, on the lawfulness, fairness, and transparency of algorithms and data (Art. 5(1)(a) and Art. 22 of GDPR).

Although HumanE-AI-Net does not intend to apply automated decision making or profiling, the consortium intends to elaborate on the social impact of algorithmic biases, explanatory AI, transparency in algorithms, responsibility and other pressing ethical and legal

issues falling within the scope of HumanE-AI-Net. The research in WP5 deals with various ethical issues such as transparency, whether biases are pre- programmed, are unintendedly introduced by the algorithm, or are the result of disproportionate data.

Ensuring adequate information to the data subjects (Art. 13, 14 of GDPR), the exercise of data subjects' rights (Chapter III GDPR) and lawfulness of processing personal data in the context of AI are focal points for the legal research taken by the consortium. The necessity of new regulations is also under debate (e.g., the requirement of algorithm inspecting authorities). Although there is no specific law (at least not in the EU) governing AI and big data, soft law and proclaimed principles are available and will be taken into account. Unfortunately, the particular assumptions, impact, and consequences of such rules and guidelines are unclear. In this sense, HumanE-AI-Net aims at taking part in the international discussion and knowledge transfer around these issues that pertain the relevant scientific communities. In that context (legally) non-binding state of the art policies, e. g. the Ethics Guidelines for Trustworthy AI recently published by the High-Level Expert Group on Artificial Intelligence (https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai) will be considered, as well as the development of a custom- tailored ethical and legal framework for HumanE-AI-Net.

4. PREVENTING MALEVOLENT USE OF RESEARCH DATA AND RESULTS

Below we clarify what measures are being taken to ensure that information from HumanE-AI-Net will not be misused by malicious (criminal or terrorist) actors.

We operate a relatively high number (4-5) of web platforms and other URL-s including a Webdav server at DFKI (detailed in D10.5). Also, we keep several documents (ca. 30, editable by all who have the proper link) at Google Doc. Besides we run some (a total of 10+) mailing lists (for them again see D10.5) to reach all partners, WP leaders, coordinators etc. selectively.

The corresponding URLs are monitored regularly. Some of them (like the Google Doc or the Webdav server) are also monitored by their providers. Human content monitoring could reveal active criminal or terrorist abuses soon. Passive abuses (like data theft) are revealed by the login history and log files checked regularly. The only form of handling sensitive data (such as personal data) by us in the project is the mailing lists run by DKFI. They are provided by the ISG (infrastructure group) of DFKI and guaranteed to be run behind firewalls and using up-to-date crime prevention technology.

Correspondingly, we have marked "No" in the ethics questionnaire under "Misuse: Does your research have the potential for misuse of research results?".

5. ACCIDENTAL FINDINGS

We want to regulate the use of accidental research findings as well.

Since the majority of our work is exerted in independent research carried out within the Micro Projects, and especially since their submission is open towards the future (as mentioned more than once above), and since they may contain external partners, in end effect we have no direct control of their activity. However, a criterion of acceptance of an MP is a statement given (when applicable) which states that the applicant should comply with the EU ethics rules

that implies a careful handling of research findings, including accidental ones. For more details see below.

5.1 EU FUNDING AND THE ACCIDENTAL FINDINGS POLICY

Accidental findings policy as an ethics issue is addressed in the Commission's guidance entitled "*How to complete your ethics self-assessment*". The Commission has published these guidelines in order to help all applicants to get their proposal "ethics-ready" for EU funding. Accidental findings policy is included in the ethics issues checklist published by the European Commission and in particular in its Section 2 (humans). It is therefore concluded that this ethical obligation applies to research that involves human participants. If this is the case, namely a human subject research, the procedures that will be implemented in the event of unexpected accidental findings should be clearly stated. In other words, researchers have an obligation to address the possibility of discovering accidental findings and describing in advance the procedure that shall be followed in such case acting both proactively (for instance, acquiring consent forms by the participants), as well as following such findings (confidentiality, communication to research participants etc.).

If one considers the ethical implications such findings may raise for researchers and at the same time what implications their disclosure to participants may present, it becomes apparent that accidental findings present a range of ethical, legal, and practical challenges, for both their recipients, as well as the researchers who encounter them. In order to better understand these implications and their possible consequences in the context of research conduct, a definition should be included in this report.

5.2 DEFINITION OF ACCIDENTAL FINDINGS

The notion of accidental findings originated in medical and genetic research. In this context, all existing definitions of accidental findings have a medical focus/orientation. A medical definition defines accidental findings as:

"test results that are outside the original purpose for which the test or procedure was conducted or accidental findings are observations of potential clinical significance unexpectedly discovered in research participants and unrelated to the purpose or variables of the study or the medical problems discovered in the course of a research/clinical trial which were not related to the topic of research or a finding concerning an individual research participant that has potential health or reproductive importance and is discovered in the course of conducting research but is beyond the aims of the study"

We can apply the same definition here, removing the words "clinical", "medical", "health" and "reproductive".

5.3 TYPES OF ACCIDENTAL FINDINGS

To mention a significant development in biology, the Presidential Commission for the Study of Bioethical Issues (Bioethics Commission) has issued a report for researchers under the title accidental and secondary findings. According to this report, accidental findings can be either "anticipatable" or "unanticipatable." We can apply this definition here.

An **anticipatable** accidental finding is one that is known to be associated with a test or procedure. Anticipatable accidental findings need not be common or even likely to occur—their defining characteristic is that the possibility of finding them is known.

Unanticipatable accidental findings include findings that could not have been anticipated given the current state of scientific knowledge. Researchers cannot plan for these types of findings specifically. However, they can consider in advance what they might do if a particular kind of unexpected finding arises, for example, one that could be actionable or lifesaving.

5.4 ETHICAL CONCERNS RAISED BY ACCIDENTAL FINDINGS

Questions the researchers should answer in advance before commencing their research include:

- *How should a finding of potential significance be handled in the research setting?*
- *Who shall be responsible to evaluate potential risks and benefits of such disclosure and ultimately take the final decision of whether to communicate such findings or not?*
- *What duties belong to basic research scientists?*
- *Whose responsibility is it to communicate the finding, to follow up, and to treat if needed?*

The answers to these questions bring to the surface further ethical concerns when accidental findings in research may raise. In more detail, possible implications these findings may have on participants should be taken into serious consideration: for example, will research subjects be ready to face the shock and anxiety that may be triggered by unwelcome and potentially bad news, can they afford the costs associated with follow-up, but most importantly do they want such findings to be communicated to them in the first place?

As far as researchers are concerned the main ethical concern that needs to be addressed is: are researchers ethically obligated to share such information with study participants and, if yes, are they qualified to do so? This obligation derives from the broader researcher duty of beneficence to secure participants' well-being by maximizing benefits and minimizing harms. In other words, researchers have an ethical duty to plan for accidental findings to the best possible extent.

5.5 THE ROLE OF ACCIDENTAL FINDINGS RISK ASSESSMENT

Given the constantly increasing number of people working in research today, the advanced medical technologies used and the number of people participating in research, the possibility of accidental findings is becoming increasingly common.

The best way to minimize the risks of accidental findings in research which may potentially entail for both researchers and participants, is the performance of a sound risk assessment.

5.6 ACCIDENTAL FINDINGS POLICY

The first step that needs to be undertaken when conducting a risk assessment is identifying accidental findings. The researchers (including MP participants) shall evaluate in advance the possibility of discovering accidental findings in their research, as well as the extent of such

possibility. Once they identify the possibility of discovering accidental findings, they should recognize and list such findings first by stating if they are anticipated or anticipatable. The possibility of findings that could not be predicted at all, if existing, should also be taken into consideration. Management of accidental findings should be an important part of the risk assessment. Researchers at this stage should be able to categorize the findings and evaluate their magnitude and the significance and potential implications they may have to research participants, seeking expert consultation, if needed. This will affect the next step of the risk assessment, that is communication policy. Specifically, what are the accidental findings should be communicated. How shall these findings be communicated? Who is the person responsible to communicate them to research subjects? The final step of the risk assessment is designing a clear policy outlining what follow- up assistance will be provided. To this effect, the implementation of an appropriate plan for the incorporation of outside expertise to evaluate accidental findings and/or apply the best return and communication policy of such findings to their recipients, could prove very useful.

The main steps of the accidental findings risk assessment could be summarized as the following:

1. Identify accidental findings;
2. Recognize and list accidental findings;
3. Manage accidental findings by categorizing and evaluating these;
4. Communicate accidental findings to the research participants;
5. Design a follow-up policy.

6. INFORMED CONSENT

Informed consent is absolutely necessary in studies involving humans in order to conduct research in an ethical and lawful manner. The European Commission in the same document mentioned (*"How to complete your ethics self-assessment*") provides some useful guidelines to participants in research on how to acquire a right and adequate informed consent from the research participants.

Participants must be given an informed consent form and detailed information sheets that:

- are written in a language and in terms they can fully understand;
- describe the aims, methods and implications of the research, the nature of the participation and any benefits, risks or discomfort that might ensue;
- explicitly state that participation is voluntary and that anyone has the right to refuse to participate and to withdraw their participation, samples or data at any time — without any consequences;
- state how biological samples and data will be collected, protected during the project and either destroyed or reused subsequently;
- state what procedures will be implemented in the event of unexpected or accidental findings (in particular, whether the participants have the right to know, or not to know, about any such findings).

Accidental findings therefore need to be taken into consideration when researchers design their consent forms. More specifically, researchers should inform potential research participants in the informed consent process and forms that:

a. accidental findings may be found;
b. describe the anticipated accidental findings that may arise;
c. inform participants of the process by which accidental findings will be evaluated;
d. inform participants of the circumstances under which these will be communicated to them, as well as of the disclosing process;
e. indicate how participants might opt out of receiving certain findings;
Most importantly, researchers should acquire the human's written and clear consent that they wish such findings, if any, to be notified to them.

7. LINKS PROVIDED

We have here assembled a set of links to key documents, mostly those of the EC. These links are understood to establish a background to the current Deliverable.

European Textbook on Ethics in research, European Commission, Directorate-General for Research

https://ec.europa.eu/research/science-society/document_library/pdf_06/textbook-on-ethics-report_en.pdf

Horizon 2020 Programme, Guidance How to complete your ethics self-assessment, European Commission, Directorate-General for Research & Innovation, version 6.1., February 2019

http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-self- assess_en.pdf

European Commission, Guidance for applicants, informed consent

http://ec.europa.eu/research/participants/data/ref/fp7/89807/informed-consent_en.pdf

European Commission, How to complete your ethics self-assessment

http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-self- assess_en.pdf

The law of Incidental Findings in Human Subjects Research – Establishing Researchers' Duties

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2581517/