GLACIATION

Green responsibLe privACy preservIng dAta operaTIONs

Deliverable D3.3 – Ethical and Privacy Impact Assessment and Recommendations (Intermediate)

GRANT AGREEMENT NUMBER: 101070141

This project has received funding from the European Union's HE research and innovation programme under grant agreement No 101070141

GLACIATION

Project acronym:	**GLACIATION**
Project full title:	**Green responsibLe privACy preservIng dAta operations**
Call identifier:	**HORIZON-CL4-2021-DATA-01-01**
Type of action:	**RIA**
Start date:	**01/10/2022**
End date:	**30/09/2025**
Grant agreement no:	**101070141**

D3.3 – Ethical and Privacy Impact Assessment and Recommendations – Intermediate

Executive Summary:	D3.3 outlines the progress and status of ethical and privacy impact assessment and recommendations		
WP:	3		
Author(s):	Ken Brown		
Editor:	DELL		
Leading Partner:	UCC		
Participating Partners:	UCC, LUH, HIRO, DELL, ECOM, ETH, LAKE, SOGEI, UNIMI, ENG		
Version:	1.2	**Status:**	Submitted
Deliverable Type:	(R) Report	**Dissemination Level:**	(PU) Public
Official Submission Date:	31/03/2024	**Actual Submission Date:**	31/03/2024

Disclaimer

The GLACIATION consortium consists of the following partners:

No.	Partner Organisation Name	Partner Organisation Short Name	Country
1	MINISTERO DELL'ECONOMIA E DELLE FINANZE	MEF	IT
2	EMC INFORMATION SYSTEMS INTERNATIONAL UNLIMITED COMPANY	EISI (DELL)	IE
3	HIRO MICRODATACENTERS B.V.	HIRO	NL
4	GOTTFRIED WILHELM LEIBNIZ UNIVERSITAET HANNOVER	LUH	DE
5	THE LISBON COUNCIL FOR ECONOMIC COMPETITIVENESS ASBL	LC	BE
6	UNIVERSITA DEGLI STUDI DI MILANO	UNIMI	IT
7	UNIVERSITA DEGLI STUDI DI BERGAMO	UNIBG	IT
8	GEIE ERCIM	ERCIM	FR
9	EURECOM	EURECOM	FR
10	SAP SE	SAP SE	DE
11	UNIVERSITY COLLEGE CORK - NATIONAL UNIVERSITY OF IRELAND, CORK	UCC	IE
12	SOGEI - SOCIETÀ GENERALE D’INFORMATICA S.P.A.	SOGEI	IT
13	LAKESIDE LABS GMBH	LAKE	AT
14	ENGINEERING - INGEGNERIA INFORMATICA SPA	ENG	IT
15	EIDGENOESSISCHE TECHNISCHE HOCHSCHULE ZUERICH	ETH	CH

Document Revision History

Version	Description	Contributions
0.1	Table of Content	LUH
0.2	1st draft for editing	UCC
1.0	Version 1	UCC
1.1	Addressing review suggestions	LUH
1.2	Final Version	UCC

Author

Author	Partner
Ken Brown	UCC

Reviewers

Name	Organisation
Alessio Chellini	MEF
Javad Chamanara	LUH
Aidan O'Mahony	Dell

Table of Contents

List of Terms and Abbreviations

Abbreviation	Description
AI	Artificial Intelligence
ALTAI	Assessment List for Trustworthy AI
CC0	Creative Commons 0
DKG	Distributed Knowledge Graph
DMP	Data Management Plan
DoA	Description of the Action
DPIA	Data Protection Impact Assessment
DRI	Decentralised Resource Identifier
EAB	External Advisory Board
EGTAI	Ethics Guidelines for Trustworthy AI
FAIR	Findable-Accessible-Interoperable-Reusable
GA	Grant Agreement
GDPR	General Data Protection Regulation
GEP	Gender Equality Plan
GLACIATION	Green responsibLe privACy preservIng dAta operations
H2020	Horizon 2020
HE	Horizon Europe
HLEG-AI	High Level Expert Group on AI
ML	Machine Learning
Mxy	Month xy of the project's duration
PA	Public Administration
R&I	Research and Innovation
RRI	Responsible Research and Innovation
UCx	Use Case x
WP	Work Package

Executive Summary

In all projects and deployments, the use of Artificial Intelligence should be reviewed in order to ensure ethical and trustworthy practice, particularly in cases which involve human end users or human subjects, or which use private personal data. GLACIATION proposes the use of AI systems to manage data movement and workload placement in distributed edge cloud systems.

The objectives for this report were to review the guidelines on ethical and trustworthy AI, review the use of AI in GLACIATION, highlight the aspects of the guidelines that are particularly relevant, and recommend actions for the remainder of the project.

The AI sub-systems proposed by GLACIATION are not expected to interact with human end users, and they do not process personal data. The main requirements for trustworthy AI are to ensure that the systems are reliable, traceable, explainable and auditable, particularly in cases where personal data sets may be replicated or moved, or where the end use application is safety critical. One of the three GLACIATION use cases could require the development of new AI technology which reasons about human behaviours, and if so, that development should follow the recommended guidelines.

1. Introduction

GLACIATION addresses the problem of high energy consumption and carbon emissions in the management of data analytics in cloud-edge systems. It proposes the use of a Distributed Knowledge Graph to maintain the state of the various components for managing and processing data across the networked architecture. Reductions in energy consumption will be achieved by applying Artificial Intelligence techniques to enhance the management and movement of data, and will be demonstrated in three use cases.

There has been significant and growing worldwide concern over the increasing use of Artificial Intelligence, prompted by multiple demonstrated cases of bias in such systems, and on concerns over fairness, safety and diminishing autonomy. As a result, it is now important to consider the impact of any proposed development or deployment of AI methods, and to establish processes to minimize the risk to society of such developments.

This document is an intermediate review of the use of AI in GLACIATION, with a final report to be completed at the end of the project. We first review the guidelines for the ethical and trustworthy use of AI, and the forthcoming European AI Act. We then review the proposed use of Artificial Intelligence in GLACIATION, establishing when and in what context it is needed, and establishing which aspects are relevant to the ethical use of AI. We highlight the areas of GLACIATION for which action needs to be taken, to ensure ethical and trustworthy AI. We finish with an outline of planned work in this area for the remainder of the GLACIATION project.

2. Ethical and Trustworthy AI

The High-Level Expert Group on Artificial Intelligence (HLEG-AI) was established by the European Commission in 2018, as part of the Commission's preparation for anticipated socio-economic changes that will be caused by increased deployment of AI, and to begin to prepare an ethical and legal framework for these developments consistent with European values. The HLEG-AI has produced four main deliverables: (i) the Ethics Guidelines for Trustworthy AI (EGTAI) [EC19a]; (ii) Policy and Investment Recommendations for Trustworthy AI [EC19b]; (iii) the Assessment List for Trustworthy AI (ALTAI) [EC20a], and (iv) Sectoral Considerations on the Policy and Investment Recommendations [EC20b]. This has led most recently to the AI Act, approved by the European Parliament on 13th March 2024, and expected to come into force in May 2024. We focus predominantly on the Ethics Guidelines for Trustworthy AI.

The Guidelines describe the foundations of Trustworthy AI, establishing the four fundamental principles that should be adhered to:

- **respect for human autonomy**, recognising that humans interacting with AI systems must maintain full self determination, and should not be coerced or manipulated into action or inaction;
- **prevention of harm,** requiring that AI systems should not cause harm or exacerbate it, at the level of individual people, of society and of the environment;
- **fairness**, ensuring that AI systems should be deployed in support of equality and diversity, and should not support unwarranted bias or discrimination;
- **explicability**, requiring transparency and traceability, and open communication of the decisions and the reasons that led to them.

Based on these principles, there are seven key requirements:

1. **Human agency and oversight**. AI systems should support human autonomy and decision making, rather than replace it. The system should make it easier for a human to reach a good decision. There should be clear governance mechanisms which support human oversight of deployment, of any automated decision making, and of repeated review and evaluation.
2. **Technical robustness and safety**. AI systems should undergo the same rigorous design and verification processes as any safety critical software system. Competence should be clearly established, and the systems should be accurate and reliable, and secure and resilient to attack.
3. **Privacy and data governance.** AI systems must respect an individual's right to privacy, and should not exploit private data for actions that may discriminate against the person. Clear procedures are needed for management of data, to ensure correctness and integrity, to limit access to those entitled to it, and to ensure that behaviour cannot be compromised by the insertion of malicious data.
4. **Transparency**. AI systems must be able to explain their decisions or recommendations, and their consequences, in languages and concepts understandable to those who use them or who are affected by them. Further, to ensure correctness and to enable oversight, the process and decisions must be traceable.
5. **Diversity, non-discrimination and fairness**. AI systems must respect inclusion and diversity in their operation and in their recommendations. They should avoid bias

towards or against individuals or subgroups, that could lead to unfairness or discrimination. End-user and subjects of AI systems should be involved at all stages in the life cycle.

6. **Societal and environmental wellbeing**. AI systems should be designed to respect wider social and environmental goals, ensuring that their operation is sustainable and that is has minimal adverse effect on society and on future generations.
7. **Accountability**. AI systems must have mechanisms and audit process that support post-evaluation, reporting of harms, and avenues for redress when something does go wrong.

Following these guidelines, the European Commission proposed the Artificial Intelligence Act, which was passed by the European Parliament. In addition to supporting and clarifying the guidelines, the AI Act is designed to provide legal certainty around the design and deployment of AI systems, to support a unified market in AI technology, and to strengthen governance. The AI Act is based on a classification of risk with associated regulatory requirements, ranging from no specific obligations for low risk, transparency requirements for limited risk, significant regulation for high-risk systems, and prohibition for systems with unacceptable risk.

3. Proposed use of AI in GLACIATION

As discussed in the HLEG-AI Ethics Guidelines, different opportunities and challenges arise from AI systems used in different sectors or use cases, and the implementation of these Guidelines needs to be adapted to the particular AI application.

In GLACIATION, the use of Artificial Intelligence is proposed for the core technical research and innovation. This core technology will be deployed in the three industrial use cases which provide context and demonstrations for the innovations. These individual use cases may reveal specific challenges for the AI deployed in the core; in addition, the use cases themselves may deploy AI methods specific to the intended application. For the analysis and discussion that follows, we will make a distinction between core GLACIATION and the individual Use Cases.

3.1 Core GLACIATION

The main use of AI is developed in WP3 – AI-enabled Data Movement Engine. This includes the design and use of a Distributed Knowledge Graph, for representing the state of the GLACIATION platform, the infrastructure and surrounding environment, and the metadata and work requests; deep learning, reinforcement learning and other AI/ML methods for data and workload movement; AI/ML for predict future requests and future states of the platform and infrastructure; and Swarm Intelligence, to infer relevant knowledge from the distributed platform and to assist with data and workload movement. The AI technologies developed in WP3 will be situated in the overall architecture and infrastructure designed in WP2 and implemented in WP6, which will be exploited in WP5, and deployed and evaluated in use cases in WP7. The use of AI is intended to be largely mechanical, with little human interaction and little direct relevance to human operators or human subjects. In fact, these AI-based methods are intended to enhance or replace systems that are already automated. Thus, many of the most serious concerns of the guidelines in EGTAI do not apply (as acknowledged in EGTAI [EC19a]: “While most requirements apply to all AI systems, special attention is given to those directly or indirectly affecting individuals. Therefore, for some applications (for instance in industrial settings), they may be of lesser relevance”). Nevertheless, it is still important for developers and deployers of this technology to be aware of the potential use in different applications, and of any potential ramifications for human users or subjects.

A full understanding of the possible impact on human participants requires an analysis of the different stakeholder roles for GLACIATION. We envisage the following:

GLACIATION platform manager – responsible for efficient, reliable and safe operation of an instance of the GLACIATION platform;

Service providers – providers of computer hardware or storage;

Data providers – providers of data sets, to be used by clients, and moved by GLACIATION as appropriate;

End users – agents responsible for service requests placed on the platform;

Subjects – individuals whose data may be represented in the system, or who may be interacting with software or hardware whose operation makes use of the GLACIATION platform.

3.2 GLACIATION Use Cases

There are three use cases being explored in the GLACIATION project, and one possible addition to the project under review.

UC1 - Edge Decentralised data management (MEF/Sogei): Data management and processing of attendance, performance and payroll services for public administration staff. The data that is being moved and processed refers to individual human subjects, and must be protected. Individual human subjects may enter personal data, and receive personal information from the end-to-end system. However, the core GLACIATION AI technology does not process that data, and does not interact with the individual. GDPR adherence and privacy aspects that might arise from GLACIATION are handled in other work packages separately from the AI systems.

Specific challenges for core AI: protection and responsible management of data sets that contain personal human data.

AI in the use case: there is no specific use of AI in UC1.

UC2 – Data driven energy-efficient manufacturing (DELL): Data management and processing for manufacturing systems, including automated inspection and test, autonomous mobile robots, and cobots interacting with human operators. Some of the data may be personal data on human workers interacting with the cobots, and so must be protected. The core AI technology developed in GLACIATION does not process those individual data records, and does not interact directly with the human operators. GDPR adherence and privacy aspects that might arise from GLACIATION are handled in other work packages separately from the AI systems. However, efficient operation of the data analytics managed by GLACIATION is essential, since computation delays could negatively affect the safe environment for human operators in the manufacturing environment. The end use applications (for example, operation of cobots and autonomous mobile robots) could involve separate and distinct use of AI.

Specific challenges for core AI: protection and responsible management of data sets that contain personal human data; efficient management of workload, to ensure end-use safety.

AI in the use case: possible development of AI methods for cobots and mobile robots.

UC3: – Privacy preserving cross-company analytics (SAP) Data management in multi-company data analytics, and specifically distributed collaborative analytics of business data. The focus of the use case is on privacy-preserving computation. The AI technology developed in GLACIATION does not process the individual data records, and there is no interface with human operators.

Specific challenges for core AI: None

AI in the use case: there is no specific use of AI in UC3.

Glacipto *(proposed addition to GLACIATION, under review)**:*** data management and processing for electrical power generation and transmission. The AI technology developed in GLACIATION does not process the individual data records, and there is no interface with human operators. However, efficient operation of the data analytics managed by GLACIATION is essential, since computation delays or inappropriate workload placement could potentially affect power transmission.

Specific challenges for core AI: efficient management of workload, to ensure end-use safety.

AI in the use case: there is no specific use of AI in Glacipto.

3.3 GLACIATION and the key requirements of EGTAI

In this section, we review each of the seven key requirements in the light of the discussion above, identifying if and when the requirements translate into actions for GLACIATION. As noted above, UC2 is the only use case in which specific AI technology might be developed, and so the discussion of each requirement below refers to either Core or UC2.

3.3.1 Human agency and oversight

Core GLACIATION is intended to replace or enhance systems that are already automated, and no reduction in human agency is envisaged. Oversight requires traceability and explanations of data and workload placement decisions, for the GLACIATION Manager.

UC2: It is possible that AI technology for cobots and mobile robots affects Human Autonomy; if this technology is developed in GLACIATION, it needs to be monitored and assessed.

3.3.2 Technical Robustness and Safety

Resilience and security are central concerns for core GLACIATION. Software and systems engineering aspects are addressed in the non-AI work packages. But there is a specific need to assess the resilience of systems that use AI for data movement and workload placement, and to assess the reliability and recency of information available in the DKG, since there are potential safety issues for UC2 if results are delayed. The competence of the AI systems must be expressed clearly.

UC2: Any development of AI systems which interact with human subjects must be assessed for robustness and safety.

3.3.3 Privacy and Governance

Privacy of data is a central concern for GLACIATION, but it is handled in non-AI work packages. The AI technologies do not affect the privacy, since no personal data is accessed. Data governance is also a central concern of GLACIATION - it is important to ensure that data sets are only replicated or relocated to systems, providers and locations that respect data

privacy, ownership and sovereignty. This may require external guards on the output of AI-based data and workload movement engines to ensure that all data policies are respected. All copies of data created by AI movement must be secure, recoverable and able to be deleted.

UC2: Possible AI development may generate human data, so should respect privacy.

3.3.4 Transparency

The GLACIATION platform should provide traces of data movement, and provide explanations of decisions that were made. This will need traces of prior states of the platform, prior states of DKG, predicted demand and resource availability, in order to explain data movement decisions. The AI technology for deciding on data movement needs to be explainable, including explanations of any predictions of future states. These explanations will be needed to support the claims of technical robustness, and to support human oversight of GLACIATION operations. Some examples of possible requests for explanation by different stakeholders include:

> Manager: why was the data replicated onto that server? Why was that processing delayed? Why was that prediction made?
>
> Service provider: why was my hardware over/under utilised? What changes to the operating conditions could I make to ensure better utilisation?
>
> Data provider: why was my data set replicated to that server? What changes can I make to policy rules to stop that replication?
>
> End user: why was my worked delayed? How can I change my requests to get better performance?
>
> Subject: why did my personal data end up on that server?

UC2: if AI technology is developed for the end use application, transparency and explanations will be required, particularly for applications that involve human workers.

3.3.5 Diversity, non-discrimination and fairness

Core GLACIATION does not provide user interfaces and does not model human subjects. There is not expected to be any impact on human diversity. There may be possible risks in discrimination and unfairness to service providers or to end-user requests, based on geographic location, availability of trace data for predictions, and the accuracy of energy cost modelling. Impacts of sporadic activity or long latency in meta data updates should be modelled.

UC2: if AI technology is developed for the end use application, potential impacts on diversity, non-discrimination and fairness will need to be assessed, particularly for applications that involve human workers.

3.3.6 Societal and Environmental well-being

Environmental concerns, and in particular sustainable energy use, are the core concern of GLACIATION, and all technology is being developed to respect those concerns. However, the operation and effects of the AI systems need to be monitored to check for unexpected or unwanted behaviour (including excessive energy consumption in the training and operation of the AI systems themselves).

UC2: if AI technology is developed for the end use application, potential environmental impact of the AI systems themselves will need to be assessed.

3.3.7 Accountability

Each action or behaviour of the core AI modules for data or work placement should be accountable by some responsible authority. The traceability required for transparency will support this, providing data on the context for the decisions and the outcomes. This will be particularly important for uses cases where GLACIATION is managing personal data, or managing services that support safety critical functions.

UC2: if AI technology is developed for the end use application, then auditability and accountability will be required as above, and particularly for any functions that process personal data, or which maintain worker safety in the facility.

4. Recommended Actions

4.1 The Assessment List for Trustworthy AI web tool

During the preliminary analysis, it was noted that the ALTAI tool [ALTAI21] is running on outmoded software and hardware configurations. The configurations should be refactored to ensure compliance with recommended policies. Further, the general ALTAI tool should be reviewed to account for changes in the wider context over four years since the Assessment List was initially published. Refactoring of the existing ALTAI tool

4.2 Specialised ALTAI for GLACIATION

The use of AI in core GLACIATION is not expected to interface with human users, nor should it access personal data. The main concerns are to ensure reliability, and to provide explanations and audit trails for technical oversight. A specialised version of the ALTAI tool should be considered for GLACIATION, and a deeper review of the context for AI conducted.

4.3 Algorithm Development in WP3

The main AI development is in WP3. The developed algorithms should be augmented with explanation facilities, to allow for verification of the software, and regular monitoring and improvement. Explanations should be centred around the stakeholder scenarios, including platform manager, infrastructure providers, data provider, end user and human subjects of the data. To support safety critical applications that use GLACIATION, the competence of the AI systems should be established and explicitly stated. Actions arising from AI-based decisions should conform to policies on data movement.

4.4 Infrastructure Development

The AI systems developed in GLACIATION are intended to operate in the context of the wider GLACIATION platform and infrastructure. That platform should be able to record snapshots of the system state, so that AI-based decisions can be explained and audited.

4.5 Management of Use Cases

The use cases provide context for and demonstrations of the innovations of GLACIATION. Further assessment of the requirements of these use cases should be conducted, as part of the extended review of ethical and trustworthy AI in GLACIATION, to ensure that downstream requirements for AI systems are highlighted. For development of technology in the use case (i.e. in the end use application, and not the GLACIATION platform), the only area where specific AI development is envisaged is in UC2. It needs to be established whether or not

separate AI modules will be developed, and if so, then they should be included in the deeper review of ethical and trustworthy AI.

4.6 Final recommendations for cloud-edge management

The final report (D3.4) on Ethical and trustworthy AI for GLACIATION should contain general recommendations for trustworthy AI for future research, development and deployment in cloud-edge management systems.

5. Conclusions

GLACIATION is not expected to develop AI systems that interact with human end users, nor is it expected to process personal data. Thus, the main concerns for ethical and trustworthy AI are to ensure the protection of any data sets that involve private data, to ensure that GLACIATION has sufficiently low latency to support safety-critical end-use applications, and that the AI methods are reliable, traceable, auditable and explainable. If AI is to be developed in any of the use cases, they should also adhere to the recommended guidelines.

The existing ALTAI web tool should be updated and a modified version specific to GLACIATION should be considered. The tool should be used to conduct a final assessment of the use of AI, and to generate final recommendations for future research in the management of edge-cloud systems.

References and webography

[ALTAI21] https://altai.insight-centre.org/

[EC19a] European Commission, Directorate-General for Communications Networks, Content and Technology, *Ethics guidelines for trustworthy AI*, Publications Office, 2019, https://data.europa.eu/doi/10.2759/346720

[EC19b] European Commission, Directorate-General for Communications Networks, Content and Technology, *Policy and investment recommendations for trustworthy AI*, Publications Office of the European Union, 2019, https://data.europa.eu/doi/10.2759/465913

[EC20a] European Commission, Directorate-General for Communications Networks, Content and Technology, *The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment*, Publications Office, 2020, https://data.europa.eu/doi/10.2759/002360

[EC20b] European Commission, Directorate-General for Communications Networks, Content and Technology, *Sectoral Considerations on the Policy and Investment Recommendations for Trustworthy Artificial Intelligence*, Publications Office, 2020, https://data.europa.eu/doi/10.2759/733662