DATA MANAGEMENT PLAN

PROJECT	
Project number:	[redacted]
Project acronym:	[redacted]
Project name:	[redacted]

DATA MANAGEMENT PLAN	
Date:	27/09/2024
Version:	1.0

Table of Contents

1. FAIR data

1.1. Data Summary

- **Re-use of data:** [redacted] will partially re-use existing data from previous research in diamond growth and processing, specifically in optimizing HPHT (High-Pressure High-Temperature) processes and X-ray characterization data. This data will be augmented with new experimental results to enhance the performance and scale of diamond substrates for quantum and X-ray optics applications.
- **Types and formats of data:** The project will generate multiple types of data, including:
  - X-ray topographic images and reports (JPEG/PNG, PDF)
  - Experimental results from diamond growth (CSV, XLSX)
  - Surface characterization data, such as interferometry results (TIFF, CSV)
  - Process documentation and analysis reports (DOCX, PDF)
- **Purpose of the data generation and re-use:** The data is essential for optimizing diamond production processes, refining surface quality, and assessing crystal quality and orientation. These insights will be critical for scaling up high-quality diamond substrates.
- **Expected size of the data generated or re-used:** The data size is estimated at around 500 GB, depending on the number of experimental runs, imaging, and characterization processes.
- **Origin/provenance of the data:** Data will originate from [redacted]'s know-how and R&D processes, subcontractor analysis (e.g., [redacted]), and existing published research.
- **Data utility outside the project:** The data may be useful to the broader scientific community, especially researchers in quantum technologies, X-ray optics, and material sciences. It could also benefit industrial stakeholders involved in synthetic diamond production or the quantum sensing, computing, or communication technology industries.

1.2. Making data findable, including provisions for metadata

All datasets will be assigned a DOI (Digital Object Identifier) when deposited in a trusted repository. Findings that can be protected by IP will be patented or protected via other instruments. Data that cannot be protected by IP and that may give a competitor a significant advantage will not be published but treated as an industrial secret. Data about experimental results such as X-ray topographic images, surface characterization data, interferometry results, and analysis reports may be published in international peer-reviewed scientific journals together with partners, ex. [redacted], universities, and scientific laboratories.

Rich metadata will be provided following established standards in materials science and optics. Metadata will include experiment type, data collection date, techniques used, and relevant parameters (e.g., surface roughness, crystal orientation).

Keywords such as "synthetic diamonds," "HPHT," "quantum sensing," "X-ray optics," and "crystal orientation" will be used for discovery.

Metadata will comply with standards such as Dublin Core or Materials Data Curation System (MDCS), allowing it to be harvested and indexed by repositories.

1.3. Making data accessible

Repository: Data will be deposited in a trusted repository such as Zenodo or a suitable material science database. The repository will ensure a persistent identifier for all datasets.

Data:

- Most of the data will be made openly available, except for data with commercial or IP sensitivities. For proprietary datasets, access will be restricted or embargoed until patents are filed.
- Data will be accessible via open, standardized protocols (e.g., HTTPS, OAI-PMH for metadata harvesting).
- Access to restricted data will be managed through the repository's access control, requiring user authentication and permission from a data access committee.

Metadata:

- Metadata will be made openly available and licensed under a public domain dedication CC0. This ensures that the metadata can be freely accessed and reused by any interested parties without restriction.
- Metadata will include all necessary information such as dataset identifiers, data descriptions, methodology, and access links, enabling users to access the data easily.
- The data will remain available and findable for a minimum of 10 years after the project's completion. This ensures long-term access for scientific and industrial users.
- Metadata will remain accessible even if the underlying data is no longer available, allowing users to trace the origin and methodology of the data.
- Documentation and references for any software needed to access or read the data will be provided. Where applicable, open-source code will be included to ensure data accessibility and reproducibility.

1.4. Making data interoperable

Data formats (e.g., CSV, JSON, TIFF) and metadata standards (e.g., Dublin Core, MDCS) will be used to ensure interoperability within the materials science and optics communities.

Data will include references to previous research data or datasets from prior studies where applicable.

1.5. Increase data re-use

Detailed documentation, including methodologies, codebooks, variable definitions, and unit descriptions, will accompany each dataset to ensure transparency and reproducibility.

Where possible, data will be made freely available under a Creative Commons Attribution (CC BY) license.

All the non-sensitive data will be shared in reusable formats with appropriate documentation, ensuring third-party usability.

All data generated during the project will undergo rigorous quality checks to ensure accuracy, consistency, and completeness. This includes:

- Validation and verification of experimental data through peer review by internal experts and partners.
- Data cleaning to remove any errors or inconsistencies in the datasets.
- Use of standardized protocols and procedures for data collection, ensuring that all measurements and results are repeatable.
- Regular audits to ensure data integrity, proper version control, and adherence to project-specific guidelines.
- Metadata checks to ensure all necessary descriptive information is accurate and complete for discoverability and reuse.

Further to the FAIR principles, DMPs must also address research outputs other than data and carefully consider aspects related to the allocation of resources, data security, and ethical aspects.

1.6. Other research outputs

In addition to data, the project will generate:

- Software for data analysis of X-ray topographic images and surface characterization data, such as interferometry results. The software will be shared as open-source via GitHub and Kaggle.
- Protocols for diamond surface finishing and defect mapping. The data generated from these protocols will be treated as proprietary initially, with access restricted to project partners and subcontractors. After the project completes the IP protection process, the non-sensitive protocols and associated data will be made available through a publication in an open-access journal, and the data will be shared under a Creative Commons Attribution (CC BY) license.
- Reports on the viability of large-scale HPHT diamond production. The reports will be treated as confidential during the project’s duration to protect any competitive advantage. However, summary data and non-sensitive information will be published in industry reports and shared with stakeholders. After the project's conclusion, the reports will be made publicly available via the Almax easyLab website and Zenodo repository.
- Diamond plates tested by the subcontractors ([redacted]). These diamond plates will remain with the subcontractors and may be used for further testing, experimentation, and scientific publications.

2. Allocation of resources

The costs for storage, archiving, and data management are estimated at €50,000, covering both direct (storage, repositories) and indirect (personnel, software) costs.

[redacted], our ICT expert, will oversee data management with support from external data specialists. [redacted] is appointed as the Data Protection Officer (DPO) for the project under EU General Data Protection Regulation.

Long-term preservation will be ensured by using trusted repositories that guarantee data availability for at least 10 years.

3. Data security

Data will be stored on encrypted servers with regular backups, and sensitive data will be access-controlled as needed, as specified in the "GDPR Compliance" section of this document.

All data will be securely stored in trusted repositories such as Zenodo and institutional repositories that comply with long-term preservation standards. These repositories offer robust security measures, including encryption, regular audits, and disaster recovery protocols to ensure the integrity and accessibility of the data over time.

4. GDPR Compliance

4.1. Overview of GDPR Applicability

The **General Data Protection Regulation (GDPR)** applies to this project as it involves the processing of personal data within the European Union (EU) and for EU-based data subjects. This regulation mandates strict guidelines for handling personal data, prioritizing individuals' rights to privacy and data protection. GDPR compliance is essential for ensuring that all data processing activities within the project respect the privacy and rights of individuals while maintaining transparency and accountability.

This project, "[redacted]" adheres to GDPR requirements in the following key areas:

- **Data Protection and Privacy by Design:** The project has embedded data protection principles into its processes from the outset, ensuring that data privacy and security are prioritized in all stages of data handling.
- **Compliance with EU Regulations:** All data-related activities are designed and executed following EU data protection regulations, ensuring that the project aligns with GDPR's legal and ethical standards.
- **Minimization of Data Collection:** Only essential data is collected, and all data processing is aligned with the project's specified objectives in accordance with the principles of data minimization and purpose limitation.

4.2. Lawful Basis for Data Processing

In accordance with GDPR Article 6, the lawful basis for processing personal data within this project is established based on **legitimate interest** and **contractual necessity**. This basis has been determined to ensure that all data handling aligns with GDPR requirements, minimizing any unnecessary or intrusive processing of personal data.

- **Legitimate Interest:** Personal data may be processed under the basis of legitimate interest where necessary for the operation and goals of the project. The project justifies this interest by ensuring that processing activities do not override the rights and freedoms of data subjects and are directly related to the research and innovation goals.

- **Contractual Necessity:** Where data processing is essential for fulfilling the contractual obligations of this project, the lawful basis of contractual necessity will apply. This includes data processing activities required for research and development partnerships, subcontracting, and any data sharing required under EU project guidelines.

The project strictly adheres to the GDPR principles of **purpose limitation** and **data minimization**, ensuring that data is collected and processed only for specified, legitimate purposes directly related to the project's objectives. Personal data will not be processed in any manner incompatible with these purposes, and any data processing activities unrelated to these goals are avoided to uphold GDPR's requirement for transparency and privacy.

All personal data collected within the project will be processed fairly, transparently, and lawfully, with clear documentation maintained for each lawful basis.

4.3. Data Subject Rights

In compliance with GDPR, the project ensures that data subjects have the right to exercise control over their personal data. The following rights are fully supported, with processes in place to address any requests promptly and efficiently:

- **Right to Access:** Data subjects have the right to request access to their personal data processed within the project. Upon request, individuals will be provided with information on the types of data collected, the purpose of processing, and any relevant details about data sharing or storage locations.
- **Right to Rectification:** Data subjects have the right to request corrections or updates to any inaccurate or incomplete personal data. Procedures are established to update data records efficiently and ensure accuracy.
- **Right to Erasure ("Right to be Forgotten"):** Data subjects may request the deletion of their personal data when it is no longer necessary for the purposes for which it was collected or if they withdraw their consent (where applicable). The project team will assess each request to confirm that it meets GDPR requirements and ensure secure deletion.
- **Right to Restriction of Processing:** Data subjects may request that the project restricts processing of their personal data under certain conditions, such as pending data correction or if the individual objects to the processing. Restricted data will be stored securely and excluded from further processing unless required for legal purposes.
- **Right to Data Portability:** When technically feasible and legally applicable, data subjects may request a copy of their personal data in a structured, commonly used, and machine-readable format. This allows data subjects to transfer their data to another organization if desired.
- **Right to Object:** Data subjects have the right to object to data processing based on legitimate interest grounds, particularly if they feel their rights and freedoms are being infringed. In such cases, the project will halt processing unless there are compelling legitimate grounds for continuation.

All requests from data subjects are documented and processed in accordance with GDPR guidelines. To facilitate efficient handling, designated personnel are trained to address each request while upholding transparency and protecting individual rights. A record of all requests, actions taken, and responses is maintained for accountability and compliance purposes.

4.4. Data Minimization and Purpose Limitation

In alignment with GDPR's principles of data minimization and purpose limitation, the project strictly limits the collection and processing of personal data to what is necessary to achieve its objectives. The project implements the following measures to uphold these principles:

- **Data Minimization:** Only data that is essential for the research and operational goals of the project is collected. This ensures that no unnecessary personal data is processed, reducing the risk of exposure and misuse. Project teams regularly review data collection practices to confirm that they align with the principle of minimization, removing or avoiding collection of non-essential data.
- **Purpose Limitation:** Personal data is processed solely for the specific purposes outlined in the Data Management Plan, directly related to the objectives of the project. Data will not be reused or repurposed in any manner incompatible with the original purpose unless further consent is obtained from data subjects or a legal basis is identified.
- **Periodic Review:** The project periodically reviews the relevance of the data being processed to ensure that only necessary data is retained. Data that no longer serves a relevant purpose is securely deleted or anonymized, minimizing storage and processing costs while enhancing compliance.

By strictly limiting data collection to only what is necessary and ensuring that personal data is only processed for clearly defined purposes, the project adheres to GDPR's requirements, protecting data subjects' privacy and minimizing exposure to data-related risks.

4.5. Data Security Measures

To protect personal data and comply with GDPR's data security requirements, the project has implemented robust security measures that safeguard data throughout its lifecycle. These measures ensure confidentiality, integrity, and availability of data, minimizing the risk of unauthorized access or data breaches.

- **Encryption:** Personal data is encrypted both in transit and at rest to prevent unauthorized access. Encryption standards meet GDPR requirements, ensuring that data remains secure even if accessed outside of controlled environments.
- **Access Control:** Access to personal data is restricted to authorized personnel only. Role-based access control (RBAC) ensures that team members can only access data necessary for their specific tasks, reducing the risk of unauthorized handling. Regular audits are conducted to verify compliance with access protocols.
- **Secure Storage and Backups:** Personal data is stored on secure servers that include regular backups to prevent data loss in case of hardware failure or accidental deletion. Backup data is also encrypted and stored in trusted repositories that comply with GDPR standards.
- **Pseudonymization and Anonymization:** Where feasible, personal data is pseudonymized or anonymized to protect individual identities. Pseudonymization is

applied to data that may require re-identification, while anonymization is used when data does not require further linkage to individual data subjects.

- **Incident Response Plan:** An incident response plan is in place to detect, contain, and mitigate any security breaches. If a data breach occurs, it will be reported to the relevant Data Protection Authority (DPA) within 72 hours, as required by GDPR. The project team will also notify affected data subjects if the breach poses a high risk to their rights and freedoms.
- **Monitoring and Regular Testing:** Data security measures, including firewalls, intrusion detection systems, and data access logs, are regularly monitored and tested to ensure ongoing protection. Vulnerability assessments and penetration testing are conducted periodically to identify and address any potential weaknesses.

These data security practices ensure that personal data is securely handled, stored, and processed throughout the project. By implementing these measures, the project adheres to GDPR standards and prioritizes the protection of individuals' personal information.

4.6. Third-Party Processing and Data Transfers

The project occasionally involves third-party processors and, where applicable, data transfers outside the European Union. To ensure GDPR compliance in these cases, the following measures are implemented:

- **Data Processing Agreements (DPAs):** All third-party processors handling personal data for the project are required to sign Data Processing Agreements (DPAs). These agreements outline each party's responsibilities for data protection and require third-party processors to uphold GDPR standards, including data security, confidentiality, and data subject rights.
- **Assessment of Third-Party Compliance:** Before engaging third-party service providers, the project conducts an assessment to confirm that the provider adheres to GDPR requirements. This includes evaluating their data protection policies, security measures, and ability to comply with data handling standards.
- **Restricted Transfers Outside the EU:** If data transfers outside the EU or European Economic Area (EEA) are necessary, the project ensures that adequate safeguards are in place. This may include using Standard Contractual Clauses (SCCs), Binding Corporate Rules (BCRs), or other GDPR-approved transfer mechanisms to protect data in non-EU jurisdictions.
- **Accountability for Data Transfers:** Records are maintained for all data transfers, including the nature of the data, purpose of transfer, and destination country. This documentation demonstrates compliance with GDPR requirements for international data transfers and ensures that all parties involved are accountable.
- **Continuous Monitoring and Audits:** The project periodically monitors third-party processors' compliance with the agreed-upon terms in the DPAs. Audits are conducted where feasible to confirm ongoing compliance, and corrective actions are taken if any issues are identified.

By ensuring that third-party processors and international data transfers meet GDPR standards, the project maintains robust data protection practices even when involving external parties. These measures ensure that personal data is protected throughout the processing lifecycle, regardless of location.

4.7. Documentation and Accountability

To demonstrate GDPR compliance and uphold accountability, the project maintains comprehensive records of data processing activities and ensures that all data management practices align with regulatory standards. This documentation serves as a record of compliance and is essential for transparency with regulatory authorities.

- **Record of Processing Activities (ROPA):** A detailed Record of Processing Activities (ROPA) is maintained, documenting all data processing activities within the project. This record includes information on data categories, processing purposes, lawful bases, data retention periods, and any third-party processors involved.
- **Data Protection Impact Assessments (DPIAs):** For high-risk processing activities, the project conducts Data Protection Impact Assessments (DPIAs) to evaluate potential risks to data subjects. DPIAs assess the nature, scope, context, and purposes of processing, and outline measures to mitigate any identified risks. This ensures that high-risk activities are managed with adequate safeguards.
- **Regular Compliance Audits:** Internal audits are conducted periodically to verify GDPR compliance across all project operations. These audits assess data handling procedures, data security measures, and documentation practices, identifying any areas needing improvement.
- **Staff Training and Awareness:** All team members involved in data processing activities receive GDPR training to ensure they understand their responsibilities. Training sessions cover key principles, rights of data subjects, and security practices, fostering a culture of data protection and compliance.
- **Ongoing Documentation Updates:** The project's data management documentation, including this GDPR compliance section, is regularly updated to reflect any changes in data processing activities or regulatory requirements. This ensures that records are accurate, complete, and reflective of current practices.

By maintaining thorough records and proactively managing GDPR compliance, the project demonstrates accountability and readiness for potential inspections or audits by Data Protection Authorities (DPAs). This approach strengthens the project's commitment to data protection and builds trust with data subjects and stakeholders.

4.8. Data Breach Response Plan

To address potential data breaches effectively and comply with GDPR's strict reporting requirements, the project has established a comprehensive Data Breach Response Plan. This plan ensures prompt identification, documentation, and mitigation of any incidents involving personal data.

- **Incident Detection and Reporting:** The project has systems in place to detect and report potential data breaches. All team members are trained to recognize indicators of a breach and understand the importance of immediate reporting to the designated Data Protection Officer (DPO) or project lead responsible for data protection.
- **Containment and Assessment:** Upon detecting a potential breach, the project team initiates containment measures to prevent further data exposure. The breach is assessed to determine the nature, cause, affected data, and severity of the incident, as well as the potential impact on data subjects.

- **Notification to Data Protection Authority (DPA):** In accordance with GDPR requirements, any data breach that poses a risk to data subjects' rights and freedoms will be reported to the relevant Data Protection Authority (DPA) within 72 hours of discovery. This notification includes details on the nature of the breach, affected data types, likely consequences, and measures taken to mitigate the impact.
- **Notification to Data Subjects:** If the breach presents a high risk to the rights and freedoms of individuals, affected data subjects will be informed promptly. This notification provides clear information about the breach, its potential impact, and recommendations for protecting themselves against possible adverse effects.
- **Investigation and Corrective Measures:** Following containment, the project team conducts a thorough investigation to identify the root cause of the breach and implement corrective actions to prevent recurrence. This may include strengthening security measures, revising data handling practices, or providing additional training to staff.
- **Documentation and Record-Keeping:** All data breaches are documented, including details on the breach's cause, the response actions taken, notifications issued, and corrective measures implemented. This documentation serves as a record for accountability and may be required during any DPA inquiries or audits.

By establishing a structured Data Breach Response Plan, the project ensures compliance with GDPR's breach notification requirements and prioritizes the security and privacy of data subjects' personal information.

4.9. Training and Awareness

To maintain GDPR compliance and ensure a high standard of data protection, the project has implemented a training and awareness program focused on educating all team members involved in data handling and processing activities. This program is essential for fostering a culture of data protection and ensuring that GDPR requirements are consistently met throughout the project.

- **GDPR Training Program:** All team members receive initial GDPR training that covers the core principles of data protection, data subject rights, lawful bases for processing, and data security measures. Training also includes specific project guidelines, emphasizing the importance of compliance in the context of the project's objectives.
- **Ongoing Awareness Sessions:** To reinforce compliance and keep up with evolving data protection practices, the project organizes regular awareness sessions. These sessions provide updates on GDPR requirements, address any changes in data handling procedures, and offer refreshers on best practices for data security.
- **Specialized Training for Key Roles:** Team members with significant data protection responsibilities, such as those managing access controls, incident response, or third-party processing, receive specialized training to ensure they are equipped with the knowledge to carry out their duties effectively and in compliance with GDPR.
- **Documentation of Training Activities:** All training sessions and awareness activities are documented, including participant lists, training materials, and session summaries. This documentation demonstrates the project's commitment to GDPR compliance and may be referenced during internal audits or inquiries from Data Protection Authorities (DPAs).

- **Continuous Improvement and Feedback:** The project encourages team members to provide feedback on the training program and share insights on data handling practices. This feedback is used to continuously improve the program and address any areas where further clarification or support may be needed.

By implementing a structured training and awareness program, the project promotes a deep understanding of GDPR principles across the team, ensuring that all data protection obligations are upheld. This proactive approach to training strengthens the project's compliance framework and enhances accountability in data management.

5. Ethics

This project adheres to strict ethical standards for data handling, ensuring compliance with both EU and industry-specific guidelines. While no personal or sensitive personal data is expected to be generated during the project, with the exception of the data of the personnel involved, partners, suppliers, service providers, and customers, ethical considerations are fully integrated into data management practices, particularly regarding the handling and protection of proprietary industrial data.

- **Protection of Proprietary Industrial Data:** The project generates and handles proprietary industrial data, including experimental results, process documentation, and analysis reports. Ethical considerations ensure that this data is protected to prevent misuse or unauthorized access, preserving both the intellectual property (IP) rights and competitive advantage of the project's stakeholders.
- **Confidentiality Obligations:** All team members and collaborators are bound by confidentiality agreements to prevent unauthorized sharing or dissemination of sensitive project data. These agreements reinforce the ethical commitment to protecting proprietary information and ensure that data is only accessed by authorized individuals for legitimate project purposes.
- **Data Access Controls and Security Measures:** Access to proprietary data is restricted to authorized personnel and protected by stringent security protocols, including encryption, access controls, and secure storage solutions. These measures ensure that data handling meets ethical standards for confidentiality and aligns with regulatory requirements for data security.
- **Transparency and Accountability:** Ethical transparency is maintained by documenting data handling processes and adhering to accountability measures. This includes keeping accurate records of data access and processing activities, enabling the project to demonstrate compliance with ethical standards and respond to any inquiries from partners or regulatory authorities.
- **Compliance with Industry Guidelines:** In addition to general ethical practices, the project complies with industry-specific guidelines relevant to quantum technologies and materials science. This includes sector best practices for the handling of proprietary data and protection of industrial IP, ensuring ethical responsibility in data management throughout the project lifecycle.
- **Minimizing Ethical Risks:** To mitigate any potential ethical risks, data collection is strictly limited to essential information necessary for achieving project objectives. Non-essential data collection is avoided, reducing the risk of unnecessary exposure of proprietary data and reinforcing the ethical commitment to data minimization.

By integrating these ethical considerations, the project ensures that proprietary industrial data is handled responsibly and securely, maintaining trust among stakeholders and aligning with high standards for data protection and IP integrity.

6. Other issues

[redacted] will adhere to the following data management procedures:

- **European Union Open Data Directive**: Compliance with EU guidelines for data sharing and accessibility to ensure that research data is openly accessible unless there are legitimate reasons for confidentiality.
- [redacted] **National Data Protection Regulations**: Compliance with national regulations concerning the handling and protection of sensitive and personal data, in line with GDPR.
- **Sector-Specific Guidelines for Quantum Technologies and Materials Science**: [redacted] will follow industry-specific guidelines for data management, ensuring that sensitive data related to proprietary materials or processes is handled in accordance with sector best practices for intellectual property protection.