OPEN COMMONS CONSORTIUM
PRIVACY AND SECURITY AGREEMENT

ACKOWLEDGEMENTS

The privacy and security policies below are adapted in part from the GA4GH privacy and security policies.

OPEN COMMONS CONSORTIUM RECITALS

WHEREAS, the mission of the Open Commons Consortium (OCC) includes developing, managing and operating data commons and cloud computing infrastructure to support scientific, medical, health care and environmental research, and whereas the Open Commons Consortium is a consortium managed by the Center for Computational Science Research, Inc., which is an Illinois based 501(c)(3) not-for-profit corporation;

WHEREAS, various parties have developed an open source technology platform license under the Apache License, Version 2.0, (the "**Platform**"), including software, and other technologies, for managing, analyzing and sharing biomedical data;

WHEREAS as the biomedical data managed by the Platform is organized into one or more projects ("**Projects**"), and that project data ("**Project Data**") may consist of both open access and controlled access data;

WHEREAS, various parties have contributed biomedical data ("**Contributed Data**") to a Project, and permitted the OCC to provide researchers and others with access to the Contributed Data, subject to the restrictions set forth in the Open Commons Consortium (OCC) Data Contributor Agreement;

WHEREAS, the Platform allows Users to remove data from the Platform and to transfer the data to other systems, including cloud computing based systems, and to other geographic locations.

DATA STEWARD

Each Project should identify a Data Steward ("**Data Steward**"), who is responsible for: i) ensuring that Project Data is contributed compliant with OCC Data Contribution Agreement; ii) that anyone with access to data in the Project signs a OCC Data Use Agreement; iii) that any user of computing, analysis, or other services of the Platform signs an OCC Services Agreement; iv) that users of the Platform do not share Platform accounts or logins; and v) for ensuring that when Project Users remove data from the Platform, they follow the Privacy and Security Policies below.

PRIVACY

Primary Duty of Data Privacy Protection

- All Data should be safeguarded in accordance with applicable laws, norms, and guidelines, and should not be misused or wrongfully disclosed.

Consent

- Data should be used strictly in accordance with the Data Donor's (or his/her legal representative's) consent for collection, use and sharing, and/or the terms and conditions of authorization for use by competent bodies or institutions, and in compliance with national and international laws, general ethical principles, and best practice standards that respect restrictions on downstream uses.
- Data should only be made available in identifiable form for specific purposes according to the level of permission of the research participant Data Donor and Data Steward. If data are Coded or Anonymized, it should take place at the earliest opportunity consistent with use for the authorized purposes. Moreover, Data Stewards should provide a clear summary or description of the coding or anonymization process that was applied, so that prospective Data Users can judge for themselves whether or not they can use the Data by the ethical standards and legal rules of their own jurisdiction, and judge for themselves whether the use of the Data is a responsibility that they can or are prepared to accept. Such description should also make clear that if Data are Anonymized, further data linkage would not be impossible.

Ensuring Proportionate Safeguards

- Data privacy safeguards should be proportionate to the sensitivity, nature, and possible benefits, risks, and uses of the Data.
- Assessments of privacy risks should involve not only disclosure issues, but also reasonably likely harms, which may include individual or group discrimination or stigmatization. The reputational risks for entities or individuals of allowing particular uses of Data should also be considered.
- Where required, if Data Donors have consented to broad use or sharing of their Data, Data Stewards should conduct Privacy Impact Assessments to assess privacy risks before further use or sharing of the Data.

- Data Stewards should maintain an inventory that addresses the storage arrangements for Data, as well as the flows of such Data with appropriately defined sensitivity classes of the Data.

Re-identification

- Any attempt to re-identify individuals should be strictly prohibited, except where expressly authorized by the Data Donor or under the law. This obligation follows the Data through the data sharing chain. Data Stewards should monitor data usage on a regular basis to detect any such re-identification attempts.
- Organizations should take reasonable steps to prevent the identity of individuals being leaked or determined through covert means such as metadata, URLs, and message headers.

Data Quality

- In order to promote valuable sharing, Data Stewards should ensure that the Data and any associated metadata held are accurate, verifiable, unbiased and current, and stored in systems that enhance security, interoperability and
- Data Stewards should conduct regular quality assessments of data sets.
- To ensure that quality controls are kept up-to-date, and to improve interoperability, Data Stewards should develop and maintain simple feedback mechanisms that inform them on the quality of Data and their annotations, and how the quality of the Data might be improved.

Data Disclosure and Publication

- Identifiable Data may only be disclosed publicly in a publication or other format if the Data Steward ensures that either: 1) Data Donors have provided explicit consent to public disclosure of their Identifiable Data, or 2) Data Donors have made their Identifiable Data public by their own actions or permissions. Moreover, no action should unjustifiably interfere with the interests and rights of Data Donors.
- Commitments made to Data Donors, e.g., that their Identifiable Data will not be publicly disclosed, should be respected by Data Stewards even after Data Donors have died, unless the legal representative of the deceased Data Donor provides express proof of the Data Donor's wishes to the contrary.
- Where data are to be disclosed in a Coded or Anonymized form and prior to any public disclosure of Data, the risk and potential harm of re-identification or identification of individuals through public disclosure should be assessed and documented by a Data Steward. The Data Steward should maintain the risk assessment in a secure file. Any

public disclosure of Data should include an explanation of the basis by which the risk and potential harm of re-identification or identification was deemed sufficiently minimal.

Collection and Sustainability

- Primary Data collection and aggregation should comply with all legal and ethical requirements of the jurisdiction in which the Data were collected. Identifiable Data should be held and used for only so long as is foreseen necessary for the purposes of their use, unless exemptions apply in applicable laws.
- The collection, use, and sharing of Data should be limited to what is relevant and necessary to accomplish the research purposes of a Project.
- Where appropriate, Data Stewards should ensure that Data are sustained for future use and sharing, through both archiving and using appropriate identification and retrieval systems, and through critical appraisal of the mechanisms and systems used for sharing Data, whether Identifiable, Coded, or Anonymized.
- Data Stewards, in consultation with relevant entities or individuals, should establish a plan for the possible winding down of a database or Project, and in particular establish, if possible, that the Data will be archived or transferred to another database for use in future Projects. Such a policy should make clear that Data will continue to be shared with Data Users without undue restrictions and will remain in conformity with the terms of any original consent or approval and subject to ongoing governance oversight.

Access

- Requests by Data Users to Data Stewards for access to Data should demonstrate, at a minimum: (1) legitimate intended uses that are in the public interest (i.e., securing an objective commonly valued by society) and with regard to established human rights; (2) assurances that Data are being accessed only by authorized individuals, e.g., accredited persons accessing Data that will be held and used only in safe environments; (3) a legitimate and specified time period of access; and (4) secure disposal or return to the Data Steward of the Data after use and outside of any required retention period.

Data Breach

- A Data Breach by a Data User involving the potential disclosure of Identifiable Data should be disclosed without undue delay to the relevant Supervisory Authority and to the Data Steward. The Data Steward should then disclose the Data Breach following the Data Incident Management Plan associated with the Project.

- A Data Breach by a Data Steward involving the potential disclosure of Identifiable Data should be disclosed without undue delay following the Data Incident Management Plan.
- Mechanisms and procedures should be in place to maximize the likelihood of detection of Data Breaches, and to evaluate the re-identification risks in case such breaches occur. These should be kept under regular review.

Accountability

- Data Stewards should clearly identify the individuals within their organization or entity who are responsible for data privacy, data management, and reporting procedures (including a contact person or contact point for complaints). Appropriate and regular training for the identified individuals to discharge these duties should be provided.
- Data Stewards should track new regulations, policies, expectations, and best practices, sharing these with responsible individuals within their organization or entity.
- Data Stewards and Data Users must comply with applicable privacy regulations and ethical norms at every stage of data sharing, and be in a position to provide assurances that privacy interests are appropriately protected when Data are collected, stored, processed, and shared.

Transparency

- General information should be made openly available on an ongoing basis to Data Donors as a group about how the Data in a Project are being used and for what purposes, as far as is practicable.
- Data Stewards should provide individual Data Donors, if they so request, information about how their individual Data are being used and for what purposes, as far as is practicable.
- Data Stewards and Data Users should be open about their policies and practices with respect to the privacy and security management of Data and access arrangements. These policies and practices should be made openly available in a variety of formats (e.g. digital and hard copy) and be generally understandable.

Complaints or Inquiries

- Data Stewards should put procedures in place to receive and respond to complaints or inquiries about their policies and practices relating to the privacy and security of Data. The procedures should be easily accessible and simple to use, and should involve a commitment to deal with all complaints in a timely fashion.

Vulnerable Populations

- Entities or individuals that seek to use and share Data from Vulnerable Populations should consider conducting a population-specific Privacy Impact Assessment regarding the usage and sharing of such Data.
- Data Stewards should consider working with Vulnerable Populations whose Data are proposed to be used and shared to develop a confidentiality agreement to prevent unauthorized disclosure of Data, as well as a data access protocol that governs all requests by third parties for research requiring the use of such Data.
- Research findings that identify Vulnerable Persons or Populations should not be published or disclosed without the consent of the relevant persons or communities or their representatives.

SECURITY

- Each organization should implement a security risk-management program that objectively assesses risks, and implements appropriate safeguards to protect the sensitivity and integrity of the Data it holds and accesses, and the availability of its resources and services.
- Organizational, technological and physical measures appropriate to the data use and sharing and its objectives should be implemented in such a way as to protect the interests of the individuals, families and communities whose Data are being contributed, and the interests of the organization.
- Each organization should implement and maintain security policy, practices, and technical safeguards consistent with jurisdictional law and current best practices.

Organizational Measures

- There should be ongoing commitment to security and continued emphasis of its importance by all involved in the use and sharing of Data.
- As human errors are among the most difficult errors to control, Data Stewards and their organizations should, with ongoing commitment of adequate resources: (1) develop, monitor and enforce a policy (consistent with this Policy) to secure Data; (2) appoint a security officer responsible for the implementing and enforcing the security policy and practices; (3) implement internal and external security reviews and audits; and (4) implement and require ongoing training and education of personnel on privacy and security policy and best practices.

- Each organization should implement Identity and Access Management (IAM) policy, procedures, and technology to verify the identity of each individual to whom access rights are to be granted, and to ensure that each individual is given access to all of (and only) the data and services required for a specified period of time. IAM includes identity proofing, credential issuance, rights authorization, identity authentication, and rights revocation. As part of the IAM policy, organizations should maintain a list of persons having access to Data and the list should be reviewed regularly and authenticated.
- Organizations that agree to recognize and accept authenticated identities and security attributes issued by other organizations ('federated identity') have the responsibility of assuring the trustworthiness of the issuers, as well as the currency and authenticity of asserted identities.
- Sharing of Data, whether Identifiable, Coded, or Anonymized, should be limited to legitimate scientific purposes and on a realistic need-to-know basis.
- Consequences for data breaches and breach of Confidentiality should be clearly stipulated and enforced.

Technical Measures

- Physical and logical access to computer systems and networks should be restricted to authorized individuals, and access granted only for those information assets and functions required to perform the user's assigned duties.
- Data should be Coded or Anonymized at the earliest possible opportunity.
- Where Data are Coded, an organization may assign a key to enable Coded Data to be re-identified. The assigned key may not be derived from or related to the associated individual, should not be used for any other purpose, and should not disclose the mechanism used for re-identification. The direct identifiers associated with keys should be isolated on a separate dedicated server/network without external access.
- Emergency-management and disaster-recovery plans and safeguards should be implemented, including regular back-ups.
- Technical measures to secure Data should comply with the relevant guidance and regulations (e.g., for clinical trials) and should aim to be interoperable with data sharing systems and software.
- Every system that accesses, stores, or transmits Data should record an audit log of all record security-relevant events. Audit trails should be reviewed regularly, and all suspicious events should be investigated. Where possible, automated, enterprise-wide, audit trail monitoring, with alerts for misuse and algorithms to amend or terminate

access, should be implemented. Audit logs should be maintained for a minimum of one year and carefully protected.

- Configuration management of all hardware and software should be implemented. Every change should be reviewed for potential privacy and security impacts.
- Organizations should install security-critical upgrades as soon as they become available, should monitor sources of security threat information, and should take recommended actions to protect data and services from known and emerging threats.
- Organizations should routinely test their security systems, and periodically (e.g., yearly) engage an independent third-party to perform security assessment and penetration testing.

Physical Measures

- Computers, network equipment, media, and facilities used to collect, access, store, process, transport, or transmit Data must be continuously protected using appropriate physical, technical, and procedural safeguards that limit access to authorized individuals.
- Physical security measures should be in place to protect Data from natural hazards such as floods, fires, or earthquakes.
- Hardware used for sharing Data should be tamper-resistant.