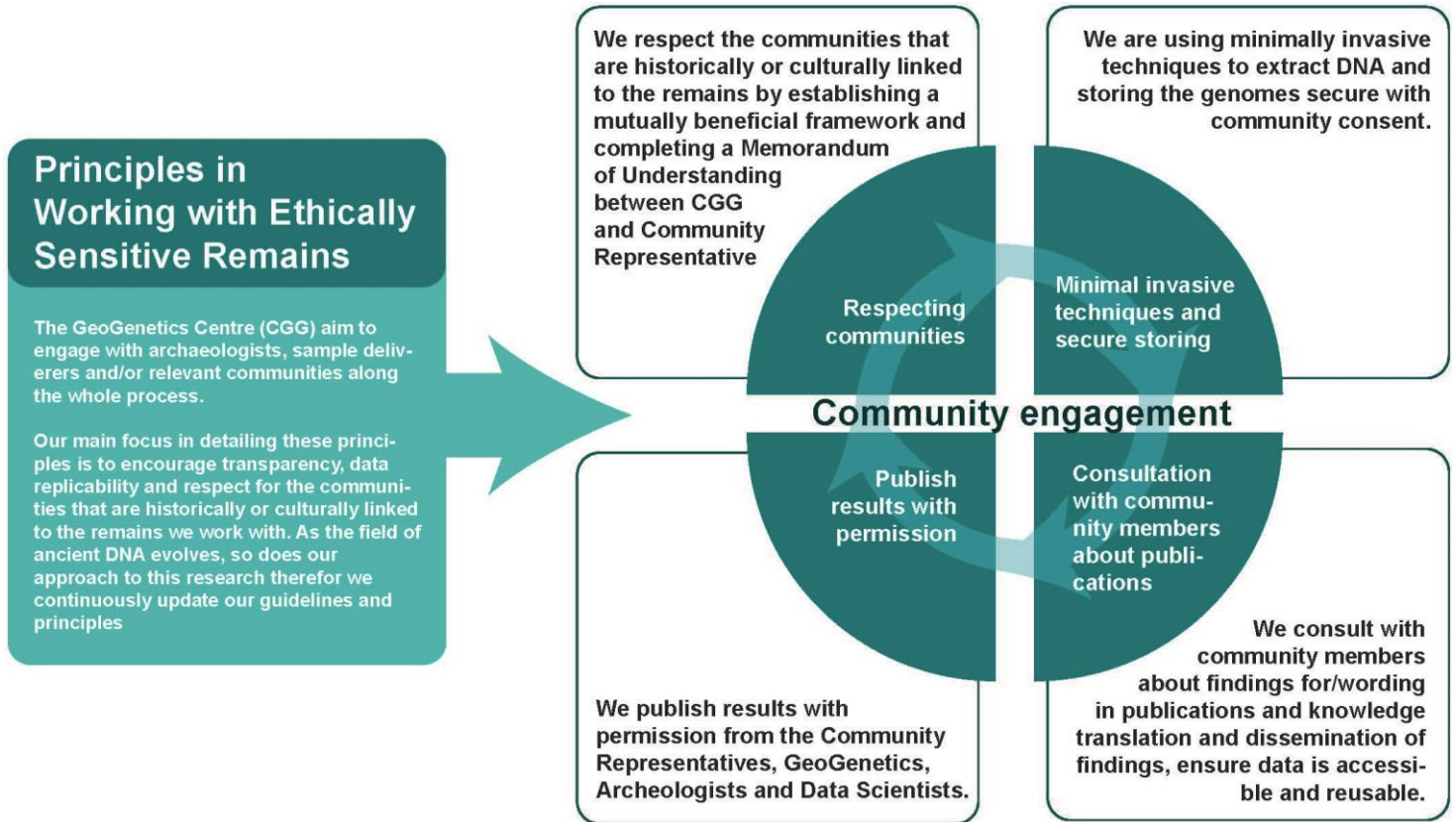


GLOBE Ethics Statement



Introduction

At the Globe Institute and the GeoGenetics Centre (CGG), we are heavily involved in many steps of the research data lifecycle, including collection and processing of samples, and bioinformatic analysis. We, therefore, believe it is important to set up clear and open principles of ethical data collection and management, to which we aim to adhere as part of the research process. Our main aim in detailing these principles is to encourage transparency, data replicability and respect for the communities that are historically or culturally linked to the remains we work with.

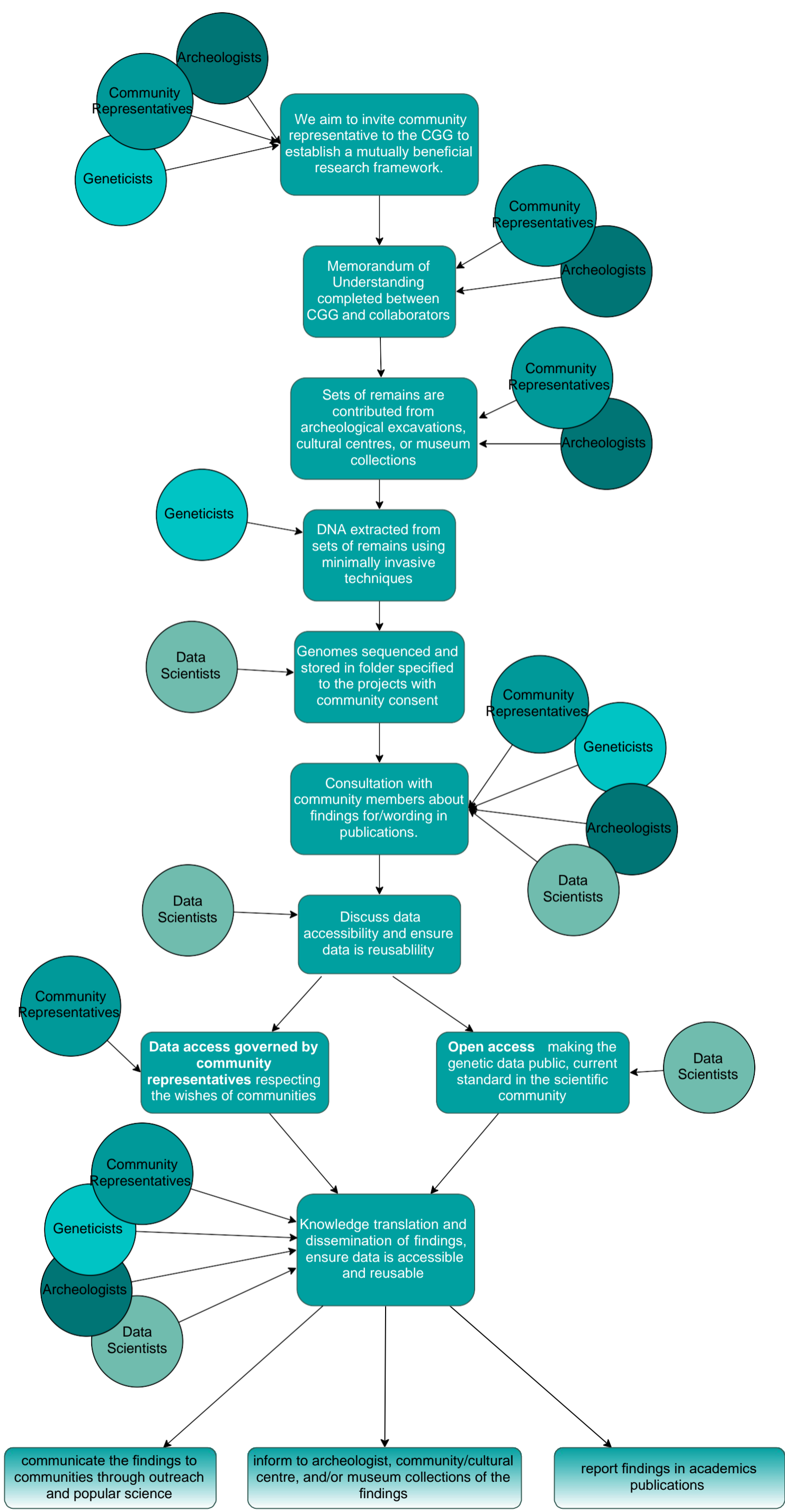
FIGURE

Below, we list our guiding principles for data collection and management in more detail.

1. Ethical review process

At GLOBE and CGG, we follow the ethical review guidelines by the [Research Ethics Committee](#) (REC) of the Faculty of Health and Medical Sciences at the University of Copenhagen and [Horizon 2020](#). The REC reviews research proposals and/or publications involving human study subjects and/or data from human study subjects or archaeological remains. The REC does not review projects with a medical focus. Where applicable projects with a medical focus are reviewed by the regional [Health Research Ethics Committee](#).

Approach to Working with Ethically Sensitive Ancient Human Remains at the Centre for GeoGenetics (CGG)



As the field of ancient DNA evolves, so does our approach to this research; we aim to continuously update our guidelines and principles.
 Last Update: September 30, 2020



2. Principles of engagement and support of Indigenous and First People's rights

We adhere to the CARE principles for indigenous data governance: <https://www.gida-global.org/care>

This means that we believe that:

- “Data ecosystems shall be designed and function in ways that enable Indigenous Peoples to derive benefit from the data.”
- “Indigenous Peoples’ rights and interests in Indigenous data must be recognised and their authority to control such data be empowered. Indigenous data governance enables Indigenous Peoples and governing bodies to determine how Indigenous Peoples, as well as Indigenous lands, territories, resources, knowledges and geographical indicators, are represented and identified within data.”
- “Those working with Indigenous data have a responsibility to share how those data are used to support Indigenous Peoples’ self-determination and collective benefit. Accountability requires meaningful and openly available evidence of these efforts and the benefits accruing to Indigenous Peoples.”
- “Indigenous Peoples’ rights and wellbeing should be the primary concern at all stages of the data life cycle and across the data ecosystem.”

3. Principles of data collection for ancient samples

We adhere to the following principles of data collection:

- All incoming samples at GeoGenetics are covered by an agreement letter (Memorandum of Understanding) signed by the signing parties, generally the sample provider and the researcher. This details our respective obligations. This MoU is signed before any sample is provided. At the request of the sample provider, the leftover of the samples can be returned.
- We make an effort to keep careful records of previous agreements that have been made in the past at our institution, to ensure these adhere to our current data handling principles to the best of our ability. Documents are digitized and saved on our server at KU.
- Our database curator prepares, receives and organizes these agreement documents, so that every time a sample arrives, they know which agreement covers it. If there is no agreement yet, they have to prepare one.
- We respect wishes of local and descendant communities as to how ancestral remains should be handled, and work with the communities to ensure that cultural and religious ceremonies and rituals are respected, both before and/or after sampling.

4. Principles of sequencing and storage of ancient data

At GLOBE/CGG when we process an ancient sample, we store its genomic and associated contextual data at different levels:

- We understand ancient DNA extraction is a destructive sampling process. For this reason, we aim to maximize the information we obtain when performing this process, and generally carry out whole-genome shotgun sequencing on the DNA extracts, for this reason.



- We screen the microbial communities associated with the ancient samples, identifying pathogenic and non-pathogenic microorganisms. We believe these are important sources of information that may reflect the life and habits of the individuals whose genomes we are sequencing.
- In addition to genetic data, we aim to store as much contextual data and archaeological context information as possible, so as to maximize the insights we can obtain. This includes radiocarbon dating and isotopic analyses, wherever possible. Sometimes however, a sample may not lend itself for radiometric or isotopic analyses (e.g. due to poor preservation), so this cannot be guaranteed for every sample we collect.

Our workflow is designed to assure the [FAIRness](#) of the data produced at GLOBE/GCC. Making our data findable and accessible is a top priority, and it is reflected in our active role in the development and adoption of standards for reporting ancient genomic data developed by the [Genomics Standards Consortium](#). We are leading the development of an extension of the [MixS](#) checklist for reporting ancient sequences ([MInAS](#)) in collaboration with other members of the ancient DNA community. With this initiative we aim to provide a set of best practices for sharing and archiving any ancient sequence data, that will be crucial to cope with the ever-increasing number of ancient datasets produced and to facilitate the design of new sampling campaigns.

5. Principles of data storage and analysis of sensitive present-day human data

When working with sensitive present-day human data, we adhere to both the university review process detailed above and the principles and regulations established by the corresponding study that collected the data. This might include, among other things, the use of restricted computer servers (separate from the general-use servers, and accessible only to project-specific researchers) for storing and analyzing this type of data. We work closely with the University of Copenhagen's [IT support](#) to ensure the necessary computational infrastructure is set up for each project.

Additionally, when working with identifiable present-day human genetic and/or phenotypic data, we are bound by the EU General Data Protection Regulation (GDPR), which places an additional set of obligations to our Centre researchers, in order to be accountable for data protection. For more information, see: <https://gdpr.eu/>