DATA AND RESEARCH RESOURCES SHARING PLAN

The BioNORAD project aims to develop a platform for early identification and mitigation of emerging pandemic threats, particularly those caused by novel strains of influenza viruses. The project will generate several types of data and research resources that can be valuable to the scientific community. In this section, we describe how we plan to share these data and research resources with the research community.

Types of Data and Research Resources: The BioNORAD project will generate several types of data and research resources, including:

- **Genomic data:** The project will generate predicted genomic sequences of key viral proteins of Influenza A.
- **Machine learning models:** The project will develop machine learning models for predicting the emergence and spread of novel strains of influenza viruses. These models will be based on the genomic data acquired during the project from public sequence databases.
- **Software tools:** The project will develop software tools for analyzing genomic and clinical data, as well as for visualizing and interpreting the results of machine learning models. These tools will be valuable for researchers working on influenza viruses and other emerging infectious diseases.

Sharing Plan: We plan to share the data and research resources generated by the BioNORAD project with the research community as widely as possible, while safeguarding the privacy of participants and protecting confidential and proprietary data and third-party intellectual property.

Data sharing: We plan to deposit the genomic data generated by the project in publicly accessible repositories with long-term DOI access. These repositories will provide a persistent and citable location for the data and ensure that they are discoverable and accessible to the research community. We will use the Zenodo platform to deposit the data, which will provide long-term archiving and DOI access. We will also ensure that the data are compliant with data sharing policies of funding agencies such as the National Institutes of Health and the Department of Defense. Genomic sequence data will be also deposited to NCBI.

Model sharing: We plan to deposit the machine learning models developed during the project in public repositories such as GitHub, with open-source licenses, making them freely available to the research community. We will provide documentation and instructions for running the models, as well as sample data for testing and validation. We will also encourage the research community to contribute to the development of the models by providing feedback and suggesting improvements.

Software sharing: We plan to deposit the software tools developed during the project in public repositories such as GitHub, with open-source licenses, making them freely available to the research community. We will provide documentation and instructions for using the tools, as well as sample data for testing and validation. We will also encourage the research community to contribute to the development of the tools by providing feedback and suggesting improvements.

Feasibility and Appropriateness: The data generated from this project will be shared through various channels to ensure maximum accessibility to the research community. In addition to depositing the software in a GitHub repository, models and data will be deposited in Zenodo, an open-access repository for scientific data, with a long-term DOI access. The datasets generated will be organized and labeled appropriately to enable easy retrieval, interpretation, and reuse by other researchers. The datasets will be accompanied by comprehensive documentation, including descriptions of the methods and procedures used to generate the data, as well as any relevant metadata.

To further facilitate the sharing of data and resources, the project team will follow best practices for data management, including adhering to FAIR (Findable, Accessible, Interoperable, Reusable) data principles. These practices include using standardized data formats and metadata, creating clear documentation for data and code, and ensuring that data and code are appropriately versioned and stored in secure and reliable repositories.

The project team recognizes that data sharing carries certain risks, particularly regarding the protection of sensitive information. However such risk is minimal in this project, which will not involve any clinical information or human participants.

In summary, the BIONORAD project is committed to promoting data and research resource sharing to enhance collaboration and accelerate the translation of research results into practical applications. The project team will ensure that data and resources generated during the project's period of performance are

made widely available while safeguarding the any proprietary information that we may use (including sequence data that do not carry permission to be posted publicly). By sharing and leveraging data and resources, this project will contribute to a more expeditious translation of research results into knowledge, products, and procedures to improve military health and global health security.

Implementation in a Dual-Use Capacity: The knowledge, information, products, technologies, or applications gained from the proposed research could be implemented in a dual-use capacity to benefit the civilian population that also addresses a need related to military health. The research findings will be published in peer-reviewed journals, presented at scientific conferences, and disseminated to the research community. The findings will also be communicated to the wider community through various media, such as press releases, newsletters, and social media platforms. The knowledge generated by this research will have broad applications beyond the military, with potential benefits for the civilian population, including the development of vaccines, diagnostic tools, and treatments for infectious diseases.

In summary, we will make every effort to ensure that all data and research resources generated during the project's period of performance will be made publicly available while safeguarding the privacy of participants and protecting confidential and proprietary data and third-party intellectual property. We recognize that the unique data generated by our project will be valuable to the research community, and we will make every effort to share this data with the research community. The research resources generated during the project will also be made publicly available. The knowledge, information, products, technologies, or applications gained from the research could be implemented in a dual-use capacity to benefit the civilian population that also addresses a need related to military health.