

# ACQUISITION OF AND ACCESS TO RESEARCH OMICS DATA

Daniel C. Berrios<sup>1</sup>, Deepak S. Kulkarni<sup>1</sup>, Gregory D. Eley<sup>3</sup>, Ash Mooney<sup>4</sup>

<sup>1</sup>NASA Ames Research Center, Mountain View, CA

<sup>2</sup>NASA Johnson Space Center, Houston, TX

<sup>3</sup>KBR Wyle Services, LLC

<sup>4</sup>Venesco & SaiTech Joint Venture, LLC

## ABSTRACT

Omics data are essential for understanding the myriad and complex effects of space environments on humans. To assure maximum benefit from these kinds of data, the NASA Human Research Program Data Management Plan stipulates that human omics data should be archived within and accessed through the NASA Life Sciences Portal (NLSP). The NLSP has the capability to acquire and provision access to omics (and other kinds of) research results for individual and ad-hoc groups of subjects at the direction of institutional review boards, or other authorizing bodies or individuals, per institutional, program and investigation-specific policies and procedures. However, because some single-subject omics data, like CT scans and other kinds of large, complex biomedical data, could be used to identify heretofore unknown risks to the subject's health, or, in certain cases, be used to identify a subject, NASA Policy Directive 7170.1 describes various policies regarding the management of and access to "research genetic testing" data, which includes many kinds of omics data. For example, NPD 7170.1 prohibits access to human research genetic data by NASA personnel who make employment decisions for the subjects from whom the data were obtained. To meet the objective of acquiring research omics data for NLSP in compliance with the policies in NPD 7170.1 and other applicable NASA policies, we designed NOMADS (the NLSP Omics Multimodal Acquisition of Data System), a new component that supports the transfer of large research data files, including research genetic testing data, using one of several different transfer mechanisms. The choice of mechanism is made by the submitter of the data, with guiding information from the system, and is likely to often be determined in large part by the nature and source location of the data. For example, for small files where the source data files are not already stored in a cloud storage system, users are likely to prefer to transfer their data to the NLSP via a web browser. Conversely, for large sets of files already organized and stored in a cloud storage system, users may opt for NOMAD's cloud-to-cloud transfer method. All omics datasets targeted for the NASA Life Sciences Data Archive must pass a variety of quality checks to ensure data integrity and adherence to the standards defined by the LSDA Data Submission Guidelines (DSG) (see <https://nlsp.nasa.gov/explore/lสดahome/datasubmit>). These include requirements that data are consistent with open standards established by the omics community. Non-compliant data will not be accepted however archivists are available to advise submitters on how to revise data submissions and re-submit until compliance is achieved. Following compliance with the LSDA DSG, omics data next undergo a variety of additional quality checks to ensure the data meet omics community standards. Domain specific Omics data quality control tools and techniques are continually evolving and linked to the advancements in omics assays utilized and thus, the tools and techniques utilized by the LSDA for data quality control and validation will need to be sustained accordingly. All human omics data will be access controlled according to the policies described above, and requiring IRB approval for any additional access grants once the data are acquired (including access for analysis using the NLSP workspace tools).

