NCBI 2019 CLOUD SERVICES FEEDBACK SURVEY

OMB Control Number: 0925-0648
Expiration Date: 05/31/2021

Public reporting burden for this collection of information is estimated to average 2 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. An agency may not conduct or sponsor, and a person is not required to respond to, a collection of information unless it displays a current valid OMB control number. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to NIH, Project clearance Branch, 6705 Rockledge Drive, MSC 7974, Bethesda, MD 20892-7974, ATTN: PRA (0925-0648). Do not return the completed form to this address.

1. CHECK ALL THAT APPLY IN THE FOLLOWING STATEMENTS

   I (and my team) currently:

   ☐ Work with <u>non-human</u> data in the cloud on the following platforms:

   ☐ AWS ☐ GCP ☐ Azure other _____

   ☐ Work with <u>human</u> data in the cloud on the following platforms:

   ☐ AWS ☐ GCP ☐ Azure other _____

   ☐ Plan to start working in the cloud within the next 12 months.

2. NCBI is planning additional cloud-based data services to complement the Sequence Read Archive. Rank your likelihood to use the following data services in the next 12 months.

   (1=MOST INTERESTED, DELIVER IT NEXT, 2= SECOND PREFERENCE, …. , 9=LEAST INTERESTING.)

_____ **BLAST cloud service** *e.g.* with an opportunity to build and search private custom databases that include your local sequence data.

_____ **SRA sequence search** *e.g.* to find accessions similar to a query string perhaps 1+ kilobases in length.

_____ **SRA metadata query** *e.g.* using BigQuery or other cloud-native database query tools to generate worklists of SRA accessions for downstream processing.

_____ **GenBank / RefSeq** *e.g.* quickly retrieving sequence data and annotations for large datasets.

_____ **dbSNP annotations** *e.g.* to mark-up VCF genotype files containing genomic coordinates and alleles with dbSNP functional and allele frequency information.

_____ **ClinVar** *e.g.* to filter human called variants into subsets according to their clinical significance and validation properties.

_____ **dbGaP genotype data** *e.g.* as VCF file objects for use in cloud-based analysis workflows.

_____ **dbGaP phenotype data** *e.g.* queried as observations via FHIR service.

_____ Other (specify): _____