Definitions

Annotations: Data describing the gene models, pathways, ontologies, and other relevant biological information. Annotations are always associated with a reference genome and are consumed during a command execution through the argument or input variables.

Argument(s): Parameter variables that are defined with a key-value pair (*Example: --Threads 10*). Arguments are used by the tools during command building.

Command: Fundamental unit of a tool that executes a specific job.

Data Store: Stores all data types uploaded into the user's account. The data store is subdivided (sequence data, references, annotations, and metadata) for easy access.

Execution Flow: Displays the list of tools in the execution order

FileList: List of files used by a tool/command

File: A single file used by a tool/command

Metadata: Data that describes an experiment. Metadata is used for data analysis through the options, argument, or input variables.

Merge Step: A step in a pipeline that merges all the input files into a single output file.

My Pipelines: Pipelines shown within a project. Pipelines can be copied from the Pipeline Library into a project.

Hub: Pipelines are grouped into categories (Hub) based on function.

Input(s): Any data files used in a command. These are either uploaded by the user or generated within an executed job.

Jobs: Each execution of a pipeline is designated with a job ID.

Option(s): Command parameters that are defined with a single word (*Example: --ignore*). These are used by the tools while command building.

Output(s): Data files created during pipeline execution.

Pipeline: A set of computational tools, which run either sequentially or parallelly in order to achieve a specific data analysis objective. Tools/commands are designated as steps in a pipeline.

Pipeline Library: The list of pre-configured pipelines available on the platform

Projects: Self-contained mini-workspaces, where sample sets can be analyzed independently without interference from other data, using multiple pipelines.

Reports: Results of a pipeline execution are aggregated into easily understandable formats for quick viewing.

Sample Set: A set of samples (cohort) selected within a project. Sample sets are used for data analysis.

Tool: A bioinformatics software program that solves a specific biological problem. A tool can contain multiple commands (Example: Samtools index, sort)

Revision #1
Created 24 January 2022 03:12:03 by Stanome
Updated 3 March 2022 03:02:49 by Stanome