

GENESIS MISSION DATASETS

BY CATEGORY: LIKELY DATASETS FROM 17 LABS,
WITH UNFILTERED TOKEN COUNT ESTIMATES

DNA, Genomics

JGI Genomics (LBNL),
20Qa

ATOM (LLNL),
1.25Qa

MVP (ORNL),
3.75Qa

Viral Genomics (LANL),
500T

Nuclear physics

Stockpile Stewardship
(NNSA), 250Qa

RHIC Heavy Ion (BNL),
45Qa

CEBAF (TJNAF),
2.5Qa

Stealth/top secret,
1.25Qa

Plasma science

NIF Shot Data (LLNL),
12.5Qa

Z Pulsed Power
Facility (SNL), 12.5Qa

NSTX-U (PPPL),
3.75Qa

Materials science

APS-U X-Ray (ANL),
125Qa

LCLS-II 'Data Deluge'
(SLAC), 75Qa

Materials Data Facility
(ANL), 500T

Materials Project
(LBNL), 125T

Critical Materials
(Ames), 500T

Climate, Energy

ESGF CMIP6 (LBNL),
10Qa

CICE / E3SM (LANL),
5Qa

ARM Archive
(ORNL/PNNL), 1.5Qa

NSRDB Solar (NREL),
1.25Qa

WIND Toolkit (NREL),
5Qa

EDX (NETL),
4Qa

Environmental DB
(SRNL), 500T

Particle, Astro, AI

LHC CMS Tier-1
(FNAL), 87.5Qa

Summit/Frontier I/O
(ORNL), 175Qa

LSST Rubin (SLAC),
15Qa

ALCF Cosmology
(ANL), 25Qa

DUNE Neutrino (FNAL),
7.5Qa

EMSL Data (PNNL),
2.5Qa

NSLS-II Imaging (BNL),
15Qa

Likely datasets and estimates only. Alan D. Thompson. December 2025. <https://lifearchitect.ai/>

