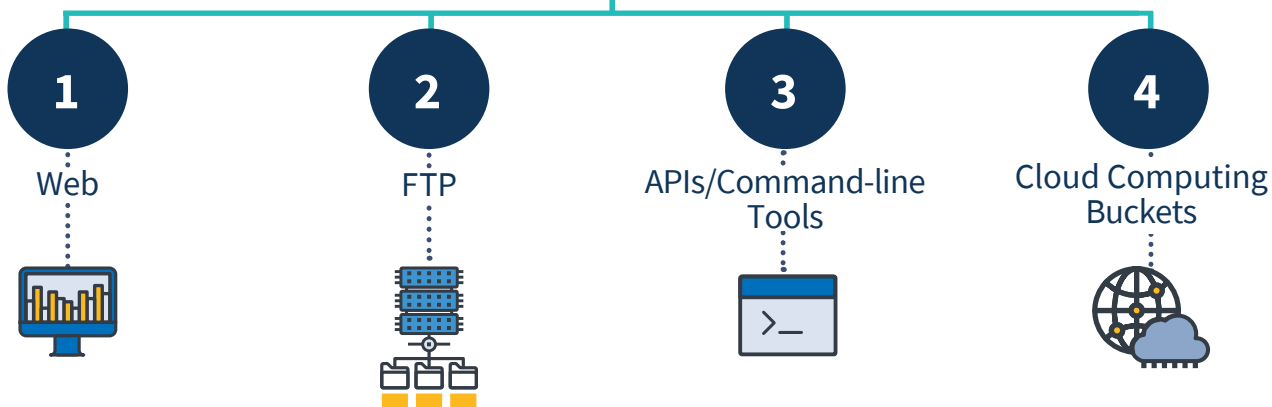


NCBI Data Access

How to find NCBI Data for your work

4 WAYS TO ACCESS NCBI DATA



Web

1. Textbox searching with text/terms

Resources have "send to" links and download buttons to facilitate downloading data.

All Database Search page

<https://www.ncbi.nlm.nih.gov/search/>

Molecular Databases Searching (Entrez) HelpDoc

<https://www.ncbi.nlm.nih.gov/books/NBK3837/>

PubMed User Guide

<https://pubmed.ncbi.nlm.nih.gov/help/>

2. BLAST Searching with sequences

BLAST Homepage

<https://blast.ncbi.nlm.nih.gov/Blast.cgi>

BLAST Help page

https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs

3. VAST Searching with 3D biomolecular structures

VAST Homepage

<https://www.ncbi.nlm.nih.gov/Structure/VAST/vast.shtml>

VAST Help page

<https://www.ncbi.nlm.nih.gov/Structure/VAST/vasthelp.html>

4. PubChem Searching with chemical information

PubChem Homepage

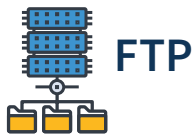
<https://pubchem.ncbi.nlm.nih.gov/>

PubChem Help page

<https://pubchemdocs.ncbi.nlm.nih.gov/about>

NCBI Data Access

How to find NCBI Data for your work



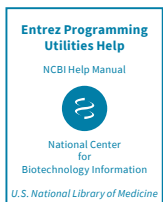
special_requests	6.62 Mb	2 files	2020-04-17 14:45
README_ensembl	15.94 kb		2021-05-24 20:35
gene2accession.gz	2.15 Gb		2021-05-30 20:39
gene2ensembl.gz	63.32 Mb		2021-05-30 20:39
gene2go.gz	23.21 Mb		2021-05-30 20:39
gene2pubmed.gz	49.08 Mb		2021-05-30 20:39

FTP Sites

Traditional FTP Site	https://ftp.ncbi.nlm.nih.gov
Fast Aspera FTP Download Site	https://www.ncbi.nlm.nih.gov/public/

APIs/Command-line Tools

Programmatic/scripting access to data in NCBI servers



1. NCBI-wide tools	
Entrez Programming Utilities, a.k.a. EUtils (API)	https://www.ncbi.nlm.nih.gov/books/NBK25501/
Entrez Direct, a.k.a. EDirect (Command-line tool)	https://www.ncbi.nlm.nih.gov/books/NBK179288/



2. NCBI Datasets resource	
Datasets APIs	https://www.ncbi.nlm.nih.gov/datasets/docs/v1/reference-docs/rest-api/
Command-line tool	https://www.ncbi.nlm.nih.gov/datasets/docs/v1/quickstarts/command-line-tools/

3. Other Resource APIs	
Genetic/Genomic Variation Service (APIs)	https://api.ncbi.nlm.nih.gov/variation/v0/
PubChem APIs	https://pubchemdocs.ncbi.nlm.nih.gov/programmatic-access



Cloud Computing Buckets

NCBI's SRA and PMC Article datasets as well as some analysis tools are available on Amazon Web Services (AWS) or Google Cloud Platform (GCP). More to be added soon...

Cloud Computing Sites:	
SRA in the Cloud homepage	https://www.ncbi.nlm.nih.gov/sra/docs/sra-cloud/
BLAST in the Cloud information	https://bit.ly/CloudBLAST
RAPT GitHub site - available for cloud use to assemble & annotate bacterial genome sequences	https://github.com/ncbi/rapt
PMC Article Datasets in the Cloud information	https://www.ncbi.nlm.nih.gov/pmc/tools/pmcaaws/