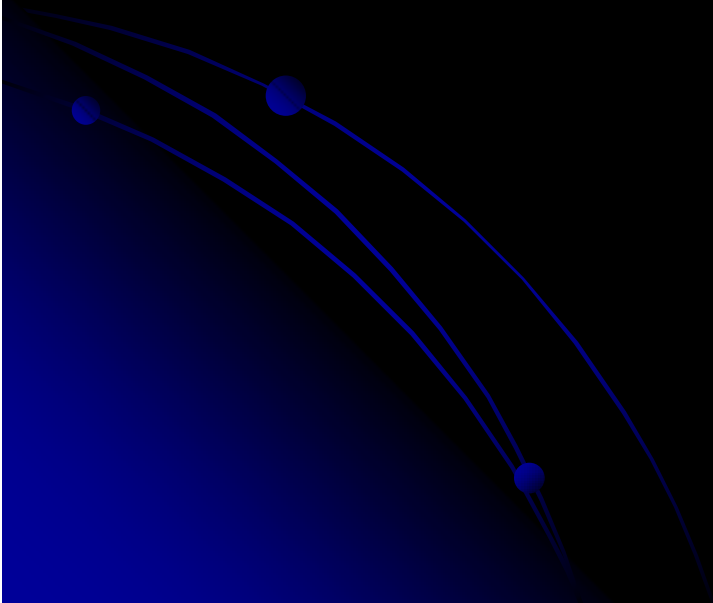
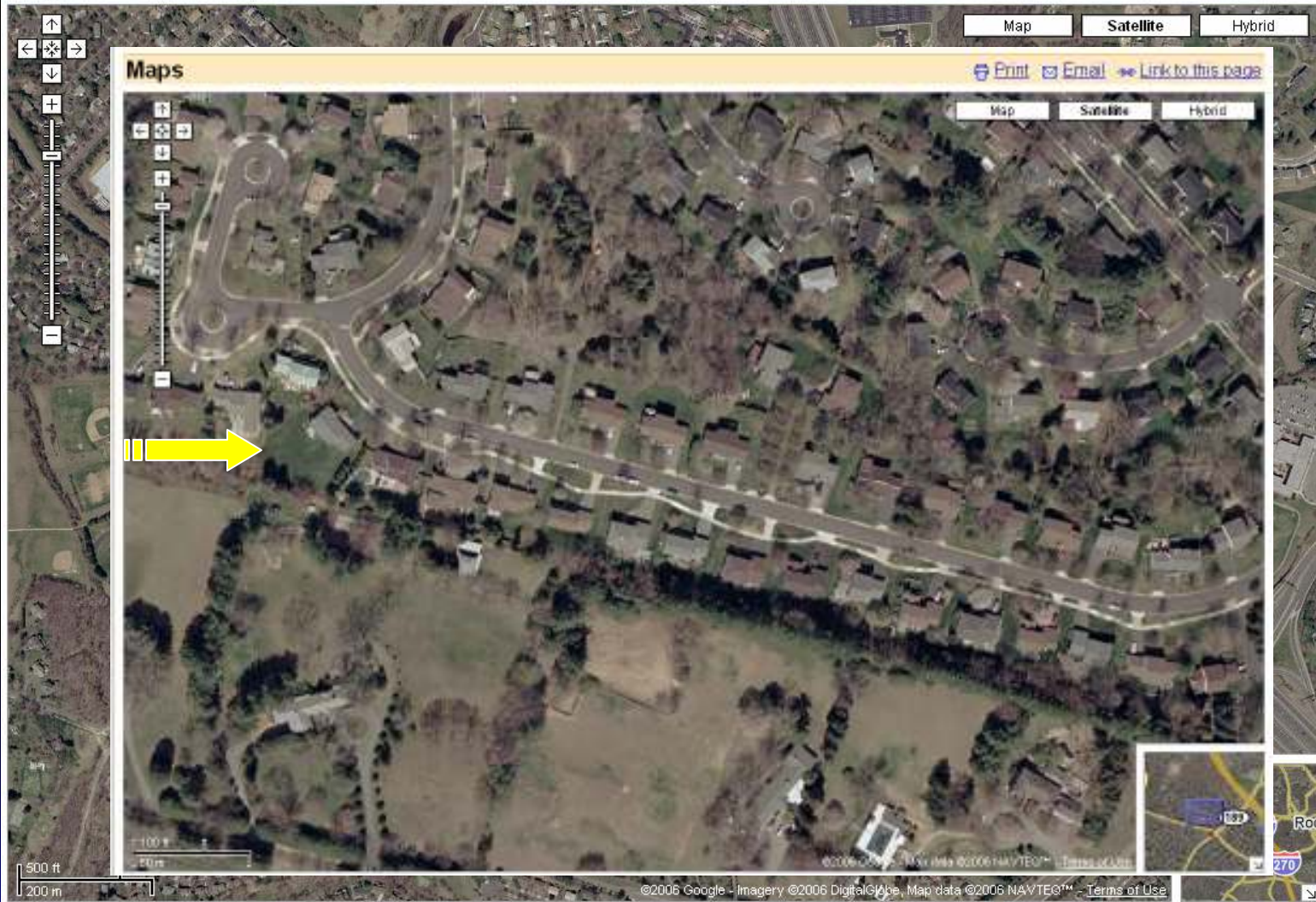


# Bioinformatics and the NCBI



## Maps

[Print](#) [Email](#) [Link to this page](#)



NIH is down here

# National Institutes of Health



# Historical context

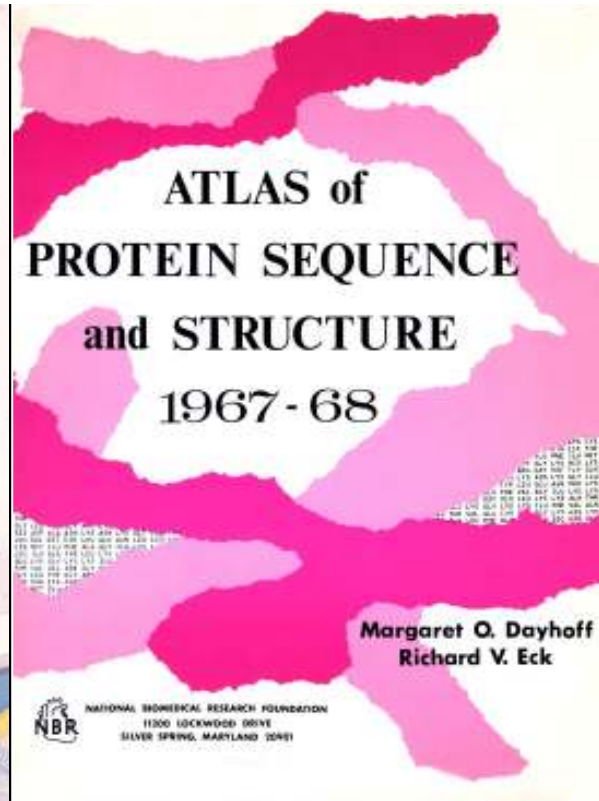
## Earliest estimates of number of protein families ~1000

- Emile Zuckerkandl

- Accomplissement et perspectives de la paleogenetique chimique.  
In: **Ecole de Roscoff (1974)** p. 69. Paris:CNRS.
- “The appearance of new structures and functions in proteins during evolution”,  
**Journal of Molecular Evolution (1975)** 7, 1-57

- Margaret Dayhoff

- Federation Proceedings 33, 2314. “The origin and evolution of protein superfamilies”,  
**Federation Proceedings (1976)** 35, 2132-2138



**Margaret Dayhoff**

<http://www.dayhoff.cc/>

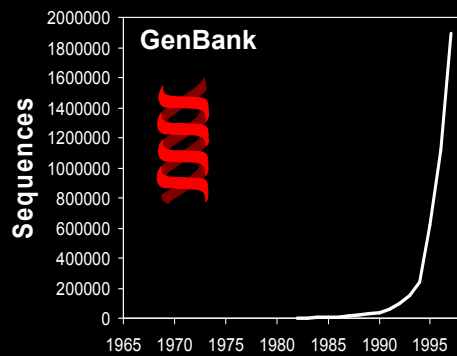
**Emile Zuckerkandl**

Molecular clock (with Pauling)

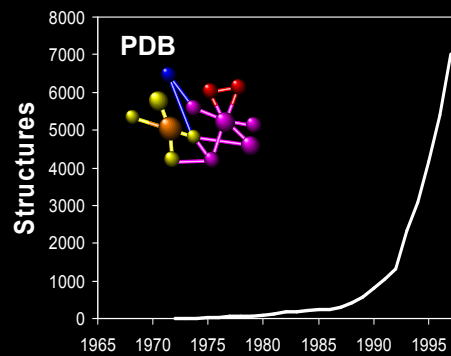


# Biological information growth

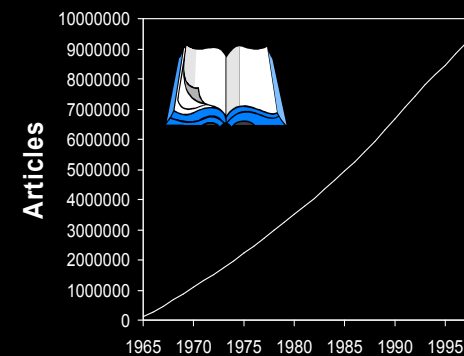
## NUCLEOTIDE SEQUENCES



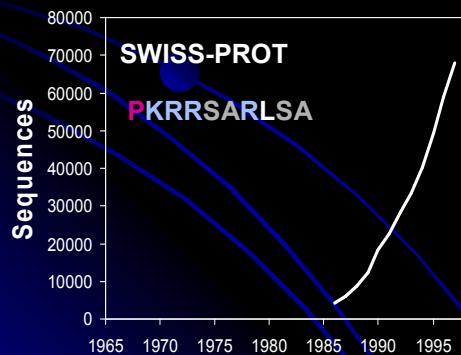
## PROTEIN 3D STRUCTURES



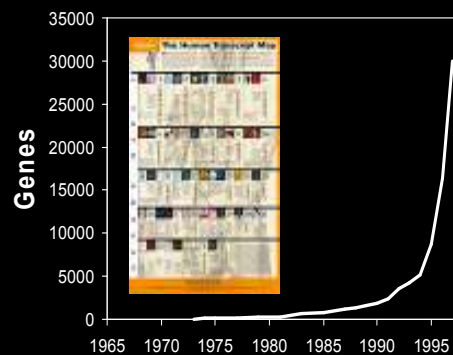
## BIBLIOGRAPHY



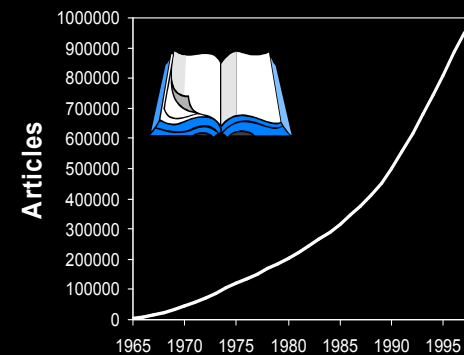
## PROTEIN SEQUENCES



## MAPPED HUMAN GENE

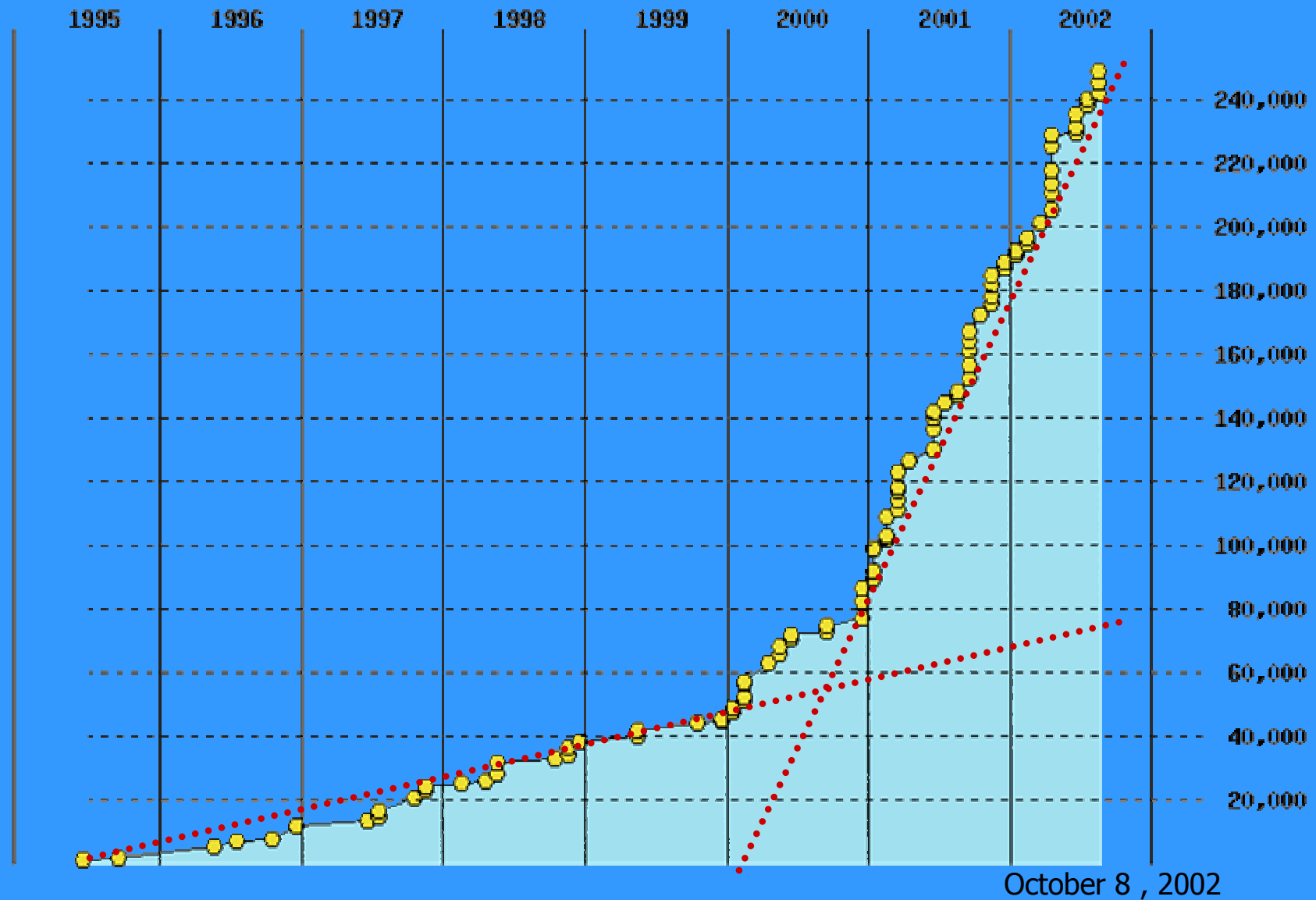


## BIBLIOGRAPHY-GENETICS



# Complete microbial genomes

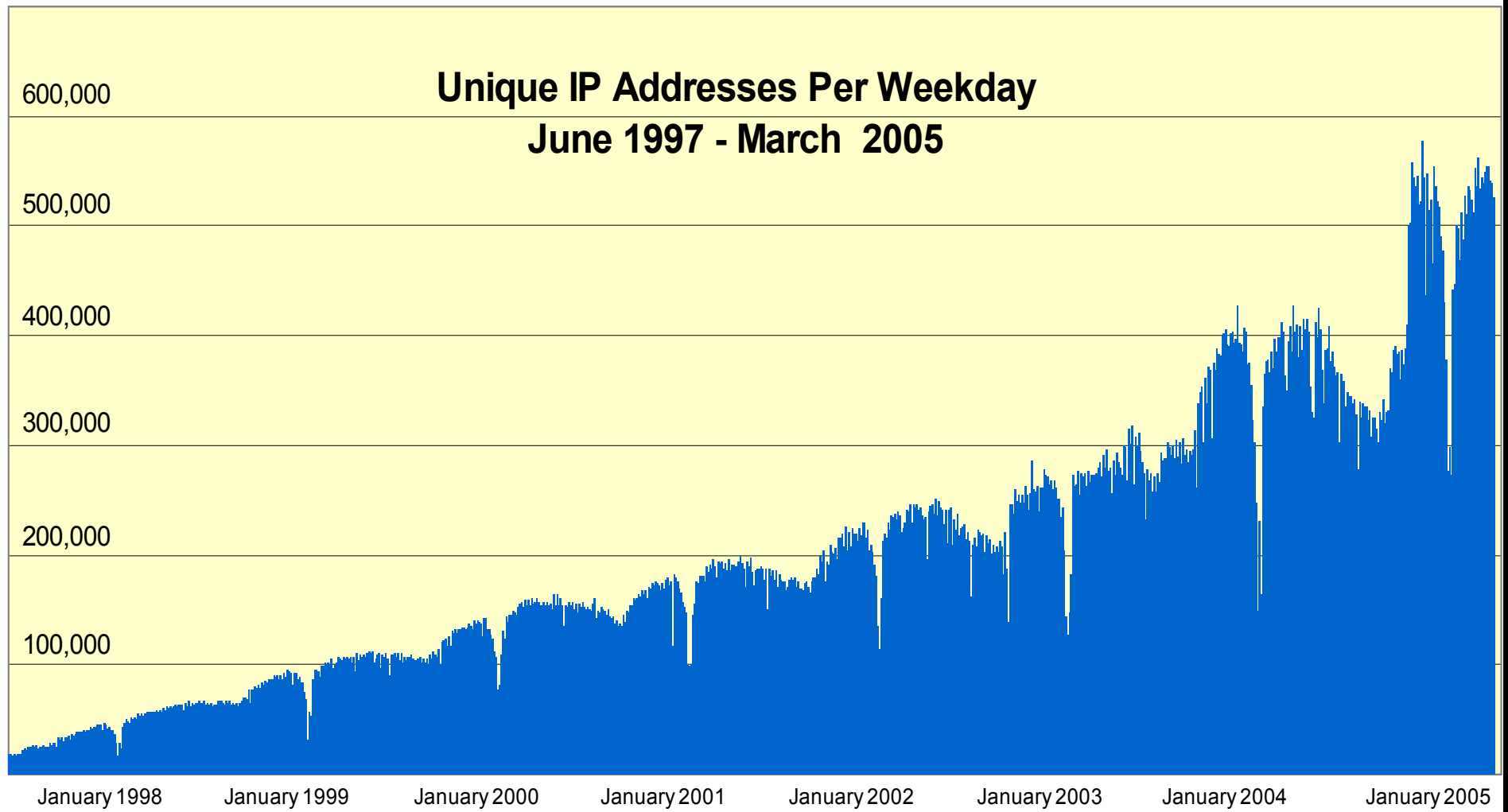
248,910 proteins in 90 complete genomes



# NCBI Usage Statistics

- Unique users at NCBI per day
  - Intellectual queries per day at all NCBI web sites
    - Including
      - PubMed
      - GenBank
      - Genomes
      - Structure
      - PubChem
      - .....
      - BLAST
- Static interactions
- Requires CPU cycles
- 
- The diagram features a list of NCBI services on the left, including PubMed, GenBank, Genomes, Structure, PubChem, and BLAST. A yellow bracket groups the first five services under the label 'Static interactions'. A yellow arrow points from a yellow box labeled 'Requires CPU cycles' to the BLAST service. In the background, there is a faint graph with a blue area under a curve and several blue dots.

# Unique users at NCBI per day



## Daily Downloads of NCBI Data

