



# AZTEC

---

NIH BD2K Center of Excellence at UCLA



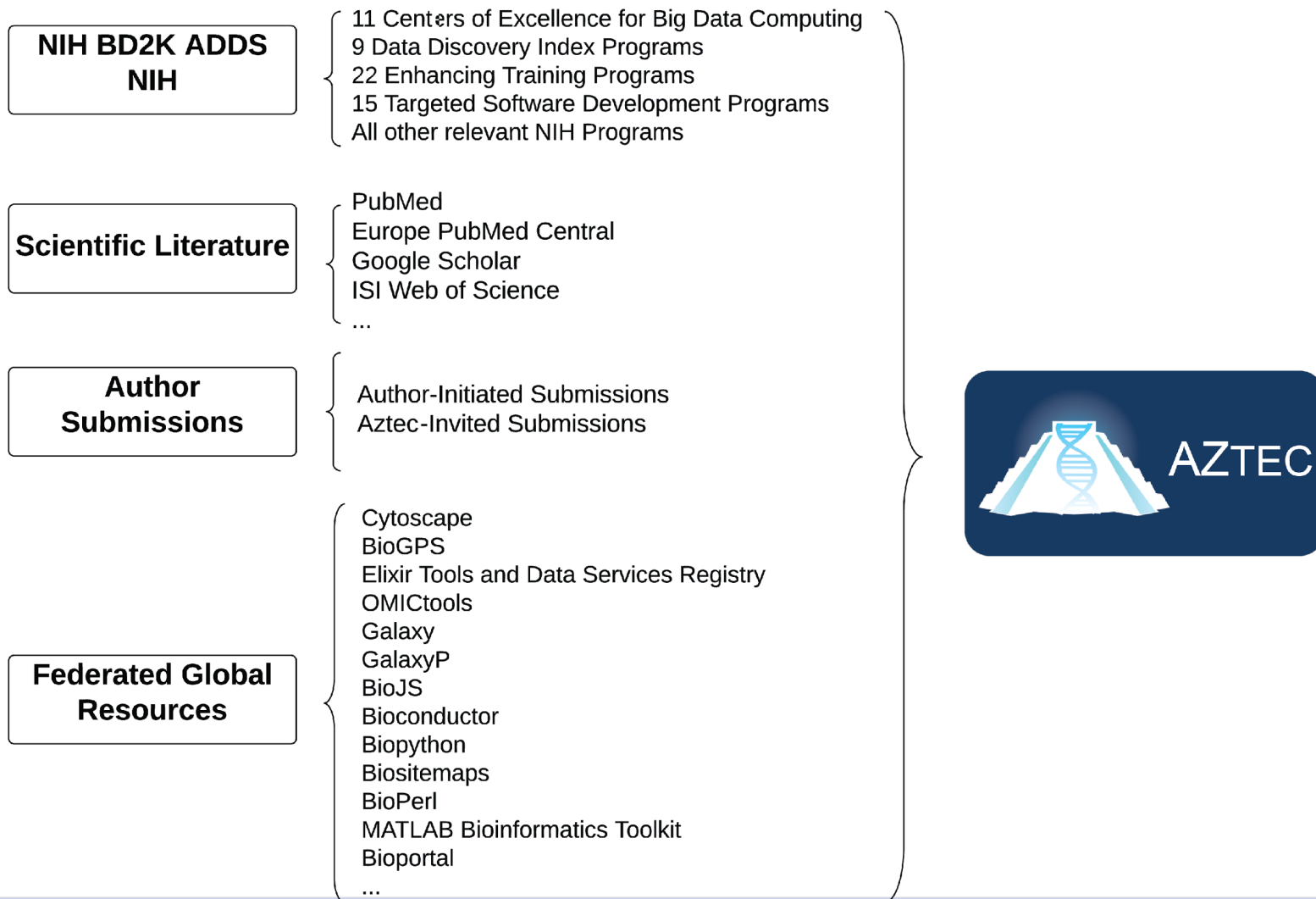
# What is BD2K AZTEC?

---

- A new global resource discovery index and metadata database for all forms of biomedical software
- A to Z TEChnology
- Aztec will serve the community by:
  - enhancing the findability of online resources
  - improving citability of the software tools
  - facilitating the reproducibility of experiments
- Aztec will interact with the nationwide BD2K programs and serve as one of the primary means to link the BD2K resource ecosystem together

# BD2K AZTEC

## Sources of BD2K Aztec Resources



# Overall Architecture and Workflow



# Identifiers and Ontologies

## Operations, Data, and Topics

**EDAM Ontology**

- Rigorously defines bioinformatics topics, operations, data types and data formats.
- Can be extended to more broadly encompass biomedical software.
- Incorporated into Aztec's Function, Biological Domain, and Input/Output Format metadata fields.

## Authors and Researchers

ORCID

## Publications

Crossref

- Uniquely identifies individual tool and publication authors.
- Incorporated into Aztec's Author and PI metadata fields.
- Uniquely identifies journal publications with DOIs.
- Incorporated into Aztec's Primary Publication and Other Publication metadata fields.

## Software Tools

University of California  
CDL  
California Digital Library

zenodo

- Uniquely identifies software programs and tools with DOIs.
- Pre-existing DOIs issued by Zenodo or other services will be tracked by Aztec. Tools without DOIs can have new DOIs minted by Aztec through our partnership with the California Digital Library.
- Incorporated into Aztec's DOI metadata field.

# Front page

[AZTEC](#) [Advanced Search](#) [Register New Resource](#) [Related Links](#) ▾



AZTEC is Technology |

Search Resources by Name, DOI, Description, Tags, or Authors.

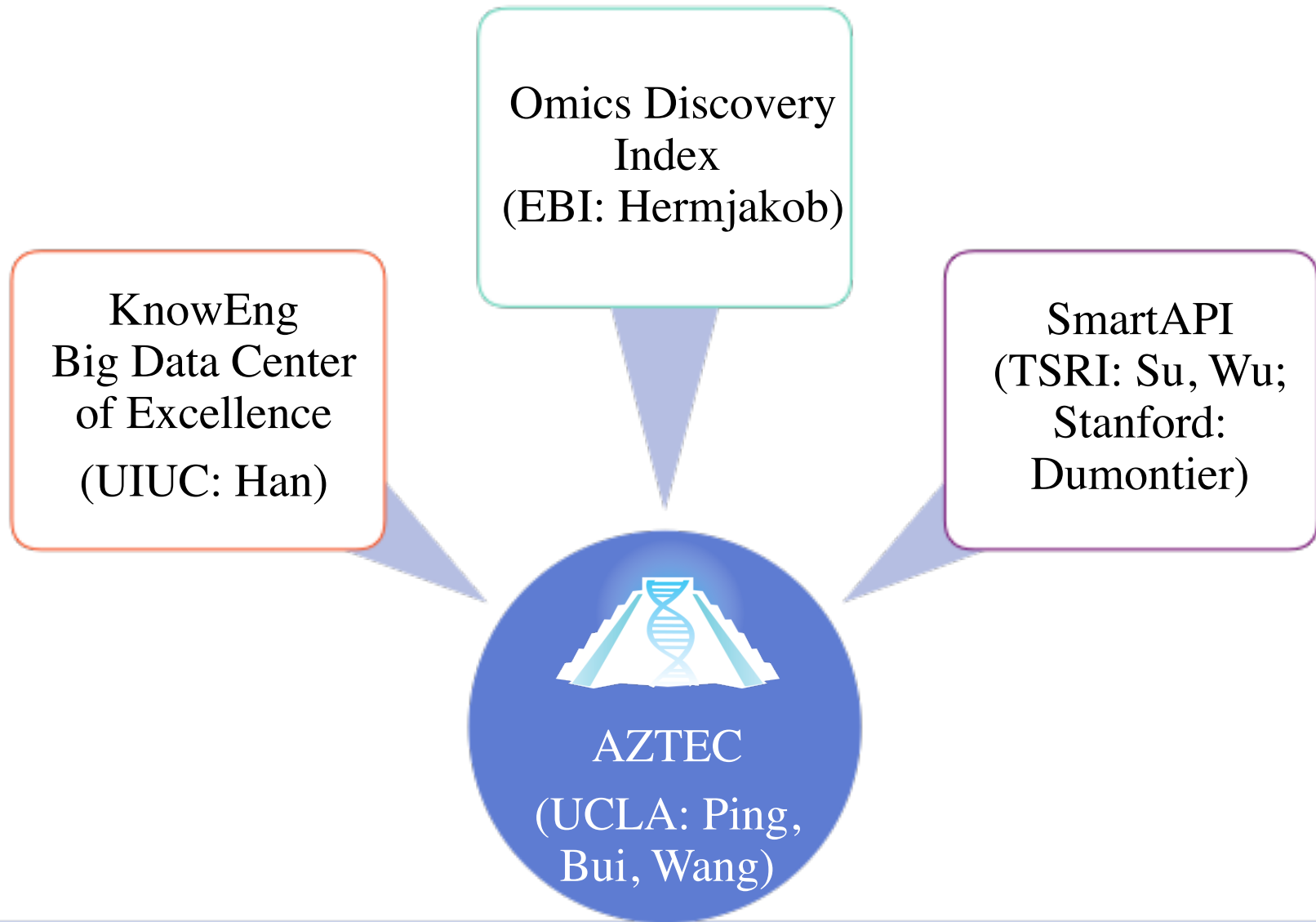


# Search “RNA”

The screenshot displays the AZTEC search interface. At the top left is the AZTEC logo, followed by a search bar containing the text "Search" and a magnifying glass icon. To the right of the search bar is a button labeled "Advanced" with a dropdown arrow. Below the search bar, the interface is divided into three main sections. On the left is a sidebar with filter options under the heading "Show results for". The first filter is "Tools (146)", which is selected. Below this are three filter categories: "Source", "Platforms", and "Types", each with a list of radio button options. The "Source" category includes "bioconductor (4)", "galaxy (9)", "elixir (1)", and "biocatalogue (2)". The "Platforms" category includes "Web UI (1)", "Command line (8)", "Linux (8)", "Mac (2)", and "API (1)". The "Types" category includes "Tool (analysis) (1)", "SOAP (2)", "Topic (1)", "Transcriptomics (1)", and "Data generation (1)". The middle section shows the search results. It starts with "Showing 1 - 10 of 155" and "Results Per Page:" with options for 5, 10, 25, 50, and 100. The first result is "macs", with a description "Model-based Analysis of ChIP-Seq", author "undefined", source "galaxy", and tag "ChIP-seq". The second result is "ccat", with a description "Control-based ChIP-seq Analysis Tool", author "undefined", source "galaxy", and tag "ChIP-seq". The third result is "macs21", with a description "MACS 2.1.0 (Model-based Analysis of ChIP-Seq) peak calling Galaxy tool used at ...", author "undefined", source "galaxy", and tags "Sequence Analysis" and "ChIP-seq". The fourth result is "seq-gen", with a description "Sequence-Generator", author "undefined", source "elixir", and no tags. On the right side of the interface is a "Search History" section with a search icon, the text "Resource: SEQ", and a circular refresh icon.

# Current and Future Collaborations

---





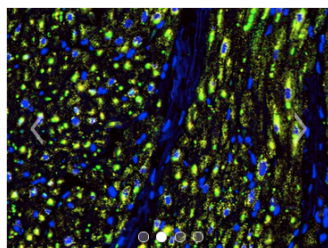
# Other UCLA Center Tools



Home Overview Functions Tutorials Help Desk Links

## Get started with COPaKB

- Protein Identifier Query
- Amino Acid Sequence Query
- MS Data File (DTA)
- MS Data File (COPaKB Client)
- MS Data File (mzML)
- Analysis of Multiple Datasets



## COPaKB Worldwide Usage Report

(as of June 2<sup>nd</sup>, 2015)



The color scale represents the number of visits to COPaKB from

## About COPaKB

The COPa Knowledgebase (COPaKB) Project is developed under NHLBI Proteomics Centers Program. COPaKB has been created as a unique resource to facilitate understanding of novel biological insights from proteomic datasets:

- COPaKB is a curated relational database of protein molecular and biomedical phenotype properties, interfaced to a website for public data retrieval.
- COPaKB enables investigators to process raw proteomic datasets without the need of high-end instrumentation, and returns a consistently annotated report of protein properties.
- COPaKB offers a wide range of informatics tools for investigators to analyze different studies in parallel and to conduct meta-analyses.

## Statistics and Updates of COPaKB

As of June 2015, Google Analytics reports the following data on COPaKB (Please see the Figure on the left):

- 157,975 pageviews;
- 8.81 pageviews per visit;
- 8,195 unique visitors;
- 117 countries are represented.

In May 2013, COPaKB implemented the following updates:

- COPaKB now offers ten modules in total;
- Information Hyperlinked over Proteins (IHOP) function is now part of the COPaKB reports;
- COPaKB is updated with Human Protein Atlas (HPA) version 11;

## Instructions to Users

Help Desk  
COPaKB Tools  
iCOPa iOS App  
Analyzing a mzML Data File  
Analyzing a DTA Data File  
Combining Parallel Analyses

## Links

NHLBI Proteomics @ UCLA  
NHLBI Proteomics Program  
The EMBL-EBI Team  
The KTH Team  
The ZJU Team  
The TSRI Team  
The UCLA Team

