bioCADDIE

The NIH BD2K biomedical and healthCAre Data Discovery Index Ecosystem

Do for data what PubMed did for literature

Main goals

Build an index of which data objects are available in major data repositories or aggregators, how these objects are associated with each other, and under which conditions they can be accessed.

Explore how to facilitate, monitor, and reward individuals who reuse and cite data, curators who annotate and package information, in addition to data producers.

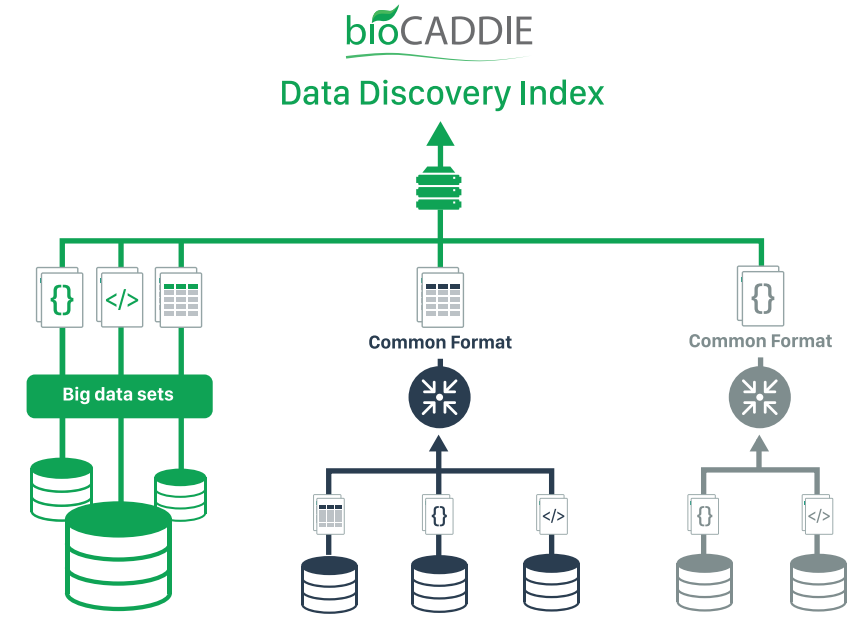
Help users find accessible data, and assist data producers on how to publish data for maximal discoverability.

Expected outcomes

Build a prototype to dock pilot products that will be build also with and the larger community.

Serve as a community hub, incubator and clearinghouse for examining issues, developing and testing innovative strategies, and establishing appropriate partnerships for creating a Data Discovery Index.

Contribute to the establishment of the NIH Commons, a conceptual framework for a digital environment that allows efficient storage, manipulation, and sharing of research objects.



Big data sets of particular interest to NIH and not covered by aggregators.

Major aggregator services (i.e., indices or repositories that use a common metadata format)

Potential use cases

Disease-based search across scales: Find all datasets from Alzheimer's patients that have RNA seq, behavioral, and imaging data available.

Molecular-based search across organisms and scales: Which proteomics and metabolomics datasets are related to the same biological process?

Molecular data/phenotype associations: Which datasets are available that have genome data about IDH1 and IDH2 in humans or other species for a particular phenotype of interest?

Behavioral and environmental data: What is the effect of stress on health? Could different components (family, work, neighborhood) have stronger associations with health?

biocaddie.org
2014 -2017

The Data Discovery Index Consortium supported by the

NIH NIH Big Data to Knowledge (BD2K)

under grant 1U24AI117966-01.

Executive Committee

University of California, San Diego
Lucila Ohno-Machado (PI)
MaryAnn Martone

University of Michigan
George Alter

University of Oxford
Susanna-Assunta Sansone

University of Texas Houston
Hua Xu