

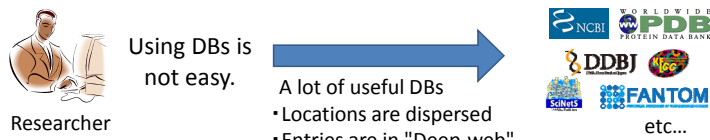
# The steps of open data integration by National Bioscience Database Center (NBDC) in Japan

Jun-ichi Onami<sup>1</sup>, Shigeru Yatsuzuka<sup>1</sup>, Tomoe Nobusada<sup>1</sup>, Katsuki Mori<sup>1</sup>, Hideki Hatanaka<sup>1</sup>, Toshihisa Takagi<sup>1,2,3</sup>  
<sup>1</sup> National Bioscience Database Center, Japan Science and Technology Agency <sup>2</sup> DDBJ Center, National Institute of Genetics  
<sup>3</sup> Department of Biological Sciences, Graduate School of Science, University of Tokyo

## National Bioscience Database Center (NBDC)

<https://biosciencedbc.jp/en/>

Databases are important for research, but...

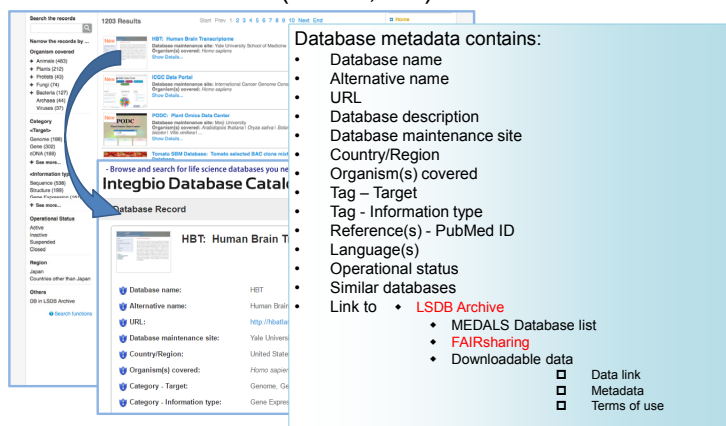


Integrated Database Project & NBDC solves those problems.

## Integbio Database Catalog

<https://integbio.jp/dbcatalog/?lang=en>

Integbio Database Catalog provides the information of life science databases (over 1,300).



Database metadata contains:

- Database name
- Alternative name
- URL
- Database description
- Database maintenance site
- Country/Region
- Organism(s) covered
- Tag - Target
- Tag - Information type
- Reference(s) - PubMed ID
- Language(s)
- Operational status
- Similar databases
- Link to
  - LSDB Archive
  - MEDALS Database list
  - FAIRsharing
  - Downloadable data

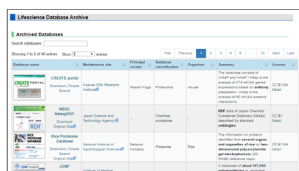
Other options: Data link, Metadata, Terms of use.

International collaboration with FAIRsharing.org started from 2018

 Metadata exchange standards, databases, policies

## Life Science Database Archive

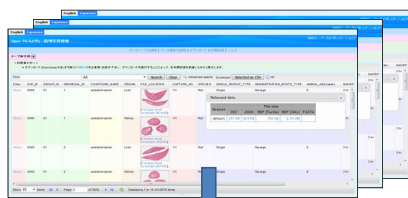
<https://dbarchive.biosciencedbc.jp/index-e.html>



SCIENTIFIC DATA

Recognized as "Recommended repository"

- Over 140 databases are **downloadable**.
- Curated **metadata** helps you to reuse data.
- Almost all databases under **CC BY-SA** license.
- All data have **DOI (Digital Object Identifier)**.
- RDFization of Life Science Database Archive



LSDB Archive SPARQL endpoint  
<https://dba-rdf.biosciencedbc.jp/sparql>

## Life Science Database Cross Search

<https://biosciencedbc.jp/dbsearch/index.php?lang=en>

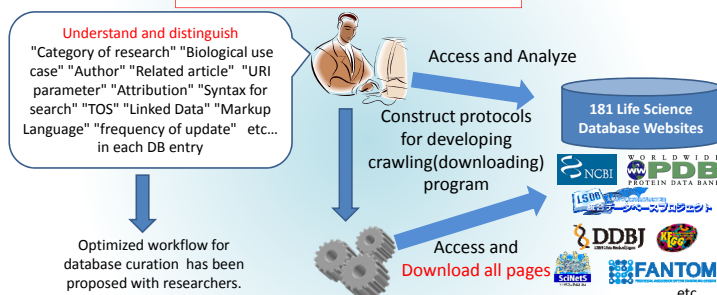
- Cross search among dispersed **181 DBs** (15 million entries).
- Deep-web entry search  
 (It is difficult for commercial search engines to find whole DB entries.)
- Specialized in Life science category



Cross search

Database information were collected by...

## Life Science Database Curator



## Toward Integrated Search

The NBDC RDF Portal provides a collection of life science datasets (Integbio Database Catalog is also included) in RDF



NBDC RDF Portal  
<https://integbio.jp/rdf/>

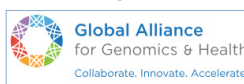
## Outreach activities

Educational activity for domestic researchers

Translation FAIR principle into Japanese  
 DOI:10.18908/a.2018041901

Referred to by the Japanese government in a guideline for establishing data policies for a kind of research organizations.  
<http://www8.cao.go.jp/cstp/stsonota/datapolicy/datapolicy.pdf>

Collaboration through international annual congresses



Discussion for standards and collaborative researches