MSc Microbiology Sem IV MIC 403c

MICROBIAL GENOMICS & PROTEOMICS

Lecture 4 I Biological Databases

Dr Shilpa Kaistha

Department of Microbiology

IBSBT, CSJMU, Kanpur

Databases

- Database: systematics collection of structured information or data
- It is supported by electronic storage and manipulation of data in a computer system
- Used for ease of Data management
- A database is controlled by DBMS: Database Management System (4 types)- (MySQL, Oracle, dBASE, Clipper, FoxPro)
- Database Server and Client: dB client if the software that connects with dB server and performs functions in the data. So different types of database clients can connect with database server (stores and manages data).
  - Hierarchical
  - Network
  - Relational
  - Object Oriented
  - Cloud Database

Database Access

- Individual database: Collection of integrated files used by an individual
- Shared Database: Database shared by organization in one location
- Distributed database: Database stored on different computers in different locations and connected by a client/server network
- Public Database/Databank: Compilation of data that s available to the public with no restrictions
- Available but with copyright
- Accessible but not downloadable
- Commercial

Curators of databases

- Large Public Institutions
  - NCBI (National Center for Biotechnology Information)
  - EMBL: European Molecular Biology Laboratory
  - DDBJ: Data Databank of Japan
- Quasi Academic Institutes
  - Swiss Institute of Bioinformatics
  - TIGR: The Institute of Genome Research
- Academic groups or Scientists
- Private Commercial Company

Biological database

- Collective term for Data compilation, organization, analysis, retrieval and dissemination of biological information
- Biological Databases are important component of bioinformatics
- Biological Databases is collection of data that is structured, searchable, updated and cross referenced
- Type of Data
  - Published Literature
  - High Through put Experiment OutPuts
  - Computational Analysis
- Information include Biological macromolecules sequence, structure, function, localization, expression data and its analysis

Published: March 1994

Biological databases: A new scientific literature

Robert J. Robbins

Publishing Research Quarterly **10**, 3–27(1994) | Cite this article

Abstract

Biology is entering a new era in which data are being generated that cannot be published in the traditional literature. Databases are taking the role of scientific literature in distributing this information to the community. The success of some major biological undertakings, such as the Human Genome Project, will depend upon the development of a system for electronic data publishing. Many biological databases began as secondary literature—reviews in which certain kinds of data were collected from the primary literature. Now these databases are becoming a new kind of primary literature with findings being submitted directly to the database and never being published in print form. Some databases are offering publishing on demand services, where users can identify subsets of the data that are of interest, then subscribe to periodic distributions of the requested data. New systems, such as fordable while offering a

Features of Biological Data

- High Volume Data
- Data heterogeneity- sequence, graphs, images, Xray data
- Large Scale Data Integration
- Searchable (Index)- table of content
- Data Sharing/ Cross referenced--- link with other databases
- Dynamic – Periodical updates– new editions
- Data Curation
- Exponential growth in biological data- fast increase in biological information requires measure for future expansion
- Data are no longer published in conventional manner but directly submitted to databases

Types of Data- based on Content

- Nucleotide Sequences-
  - raw reads, assembled genomes, whole genome sequence, single nucleotide polymorphism
  - Eg. NCBI GenBank
- Protein Sequences
- 3D structure
- Gene expression data
- Metabolic pathways
- Biological information about diseases, drugs, images etc
- Organism

Types of Databases- Source of Data

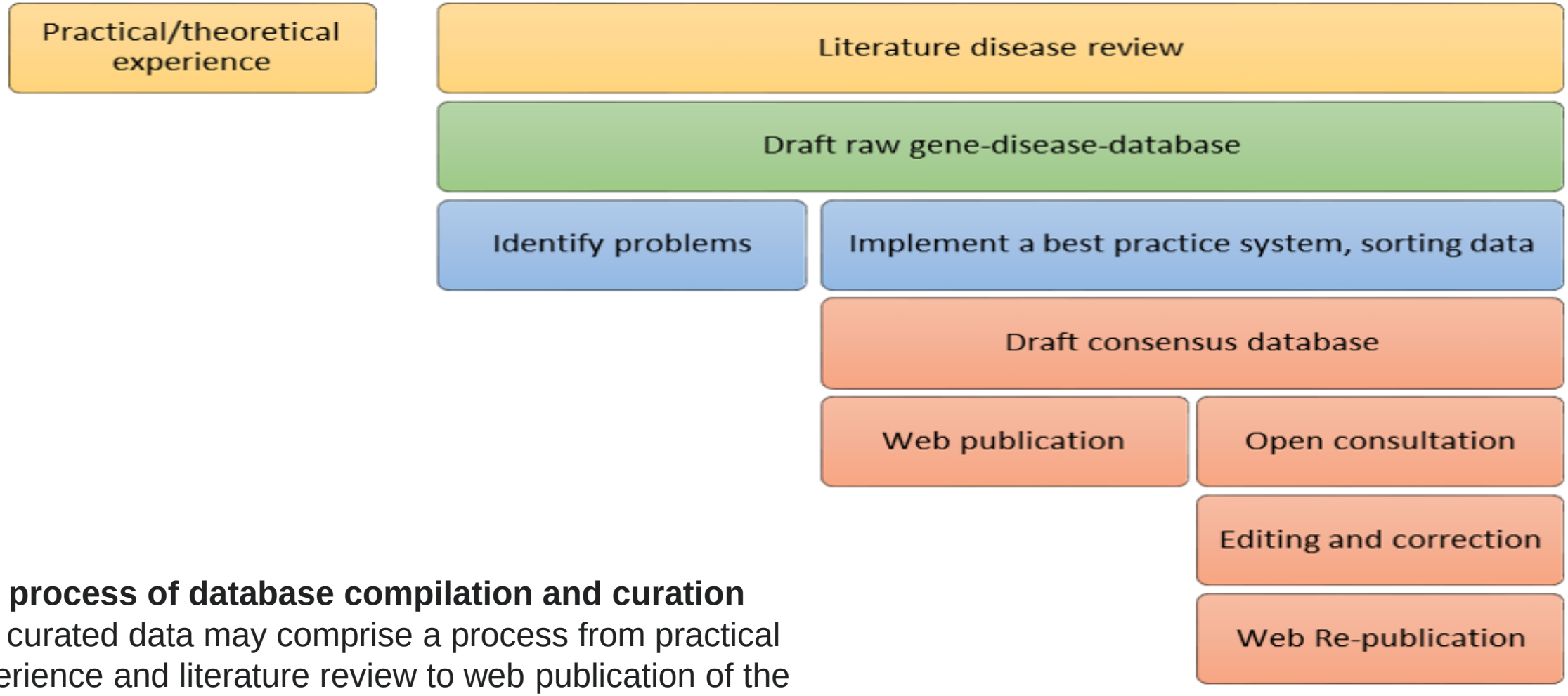
- **Primary database**
  - Original data submission by researcher
  - Example
    - Literature: Medline (PubMed, PubMed Central)
    - Nucleotide: GenBANK, EMBL, DDBJ
    - Protein: Protein Data Bank (PDB), UniProt

Secondary Database/ Derived Databases

- Derived form results of primary Database
- Manually Curated or by automated methods
- Content controlled by Third Party
- Enhanced with annotations: structures, images
- Examples:
  - NCBI Protein
  - PROSITE, Conserved Domain
  - PFAM
  - RefSeq
  - OMIM: Online Nutation Inheritance in Man
  - UniProt

Specialized Databases

- Focussed on a particular research interest of organisms
- Specific categories/ function of sequences
- Data generated by specific sequencing/ HTS technology
- Usually Curated or contain raw information
- Molecule specific
- Disease specific
  - Eg Flybase, Worm-base, Plant DB, Pseudomonas database
  - TRANSFAE- Transcription factor database
  - RFAM: RNA family database



The process of database compilation and curation
The curated data may comprise a process from practical experience and literature review to web publication of the database

K. Brown, "Online Predicted human Interaction Database," Bioinformatics, vol. 21, no. 9, pp. 2076-2082, 2005

COVID-19 Information

Public health information (CDC) | Research information (NIH) | SARS-CoV-2 data (NCBI) | Prevention and treatment information (HHS) | Español

UNITE

A new NIH initiative to end structural racism and achieve racial equity in the biomedical research enterprise.

LEARN MORE

NCBI Home
Resource List (A-Z)
All Resources
Chemicals & Bioassays
Data & Software
DNA & RNA
Domains & Structures
Genes & Expression
Genetics & Medicine
Genomes & Maps
Homology

Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

About the NCBI | Mission | Organization | NCBI News & Blog

Submit

Deposit data or manuscripts into NCBI databases

Download

Transfer NCBI data to your computer

Learn

Find help documents, attend a class or watch a tutorial

Popular Resources

PubMed
Bookshelf
PubMed Central
BLAST
Nucleotide
Genome
SNP
Gene
Protein
PubChem

Literature database- Pubmed interface used to search NLM Medline (biomedical literature)

PubMed

PubMed.gov

diabetes mellitus

Search

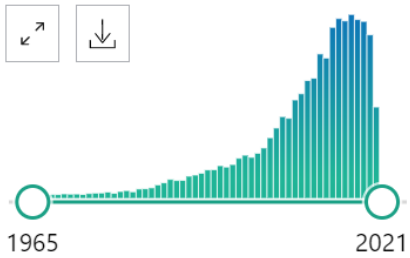
Advanced Create alert Create RSS User Guide

Save Email Send to Sorted by: Best match Display options

MY NCBI FILTERS

21,697 results

RESULTS BY YEAR



Filters applied: Randomized Controlled Trial. Clear all

TEXT AVAILABILITY

Abstract

1

Cite

Share

The effect of a novel probiotic on metabolic biomarkers in adults with prediabetes and recently diagnosed type 2 **diabetes mellitus**: study protocol for a randomized controlled trial.

Palacios T, Vitetta L, Coulson S, Madigan CD, Denyer GS, Caterson ID.

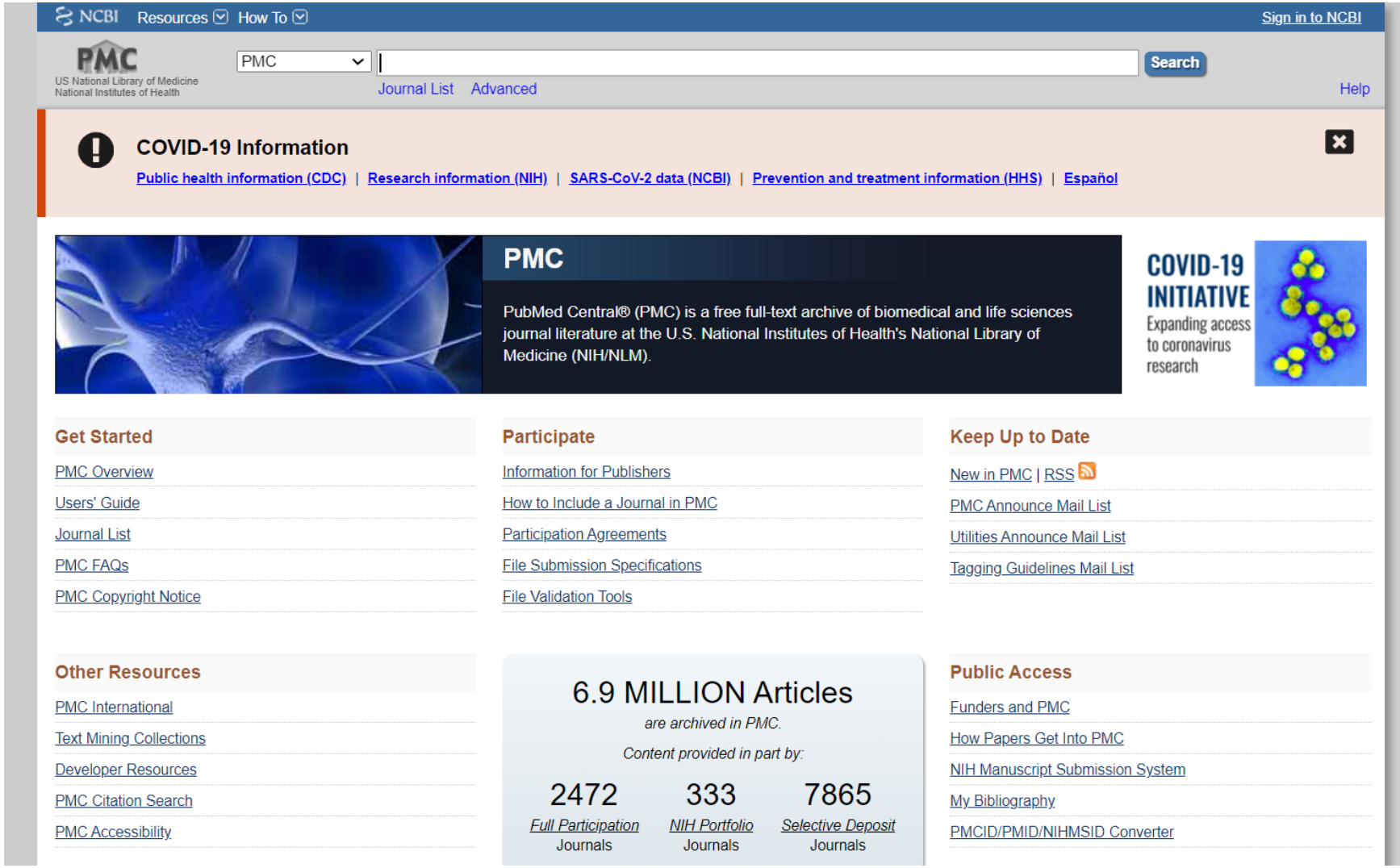
Trials. 2017 Jan 9;18(1):7. doi: 10.1186/s13063-016-1762-x.

PMID: 28069054 Free PMC article. Clinical Trial.

BACKGROUND: Shifts in the gastrointestinal microbiome have been shown to contribute to the progression of metabolic diseases including prediabetes and type 2 **diabetes mellitus**. Research suggests that in-vivo modulation of the gut microbiome by specific probiotic mic ...

Back to Top

PubMed Central



COVID-19 Information

Public health information (CDC) | Research information (NIH) | SARS-CoV-2 data (NCBI) | Prevention and treatment information (HHS) | Español

Article attributes
Associated Data
Author manuscripts
Digitized back issues
MEDLINE journals
Open access
Preprints
Retracted

Text availability
Include embargoed articles

Publication date
1 year
5 years
10 years
Custom range...

Research Funder
NIH
AHRQ
ACL
ASPR
CDC
DHS
EPA
FDA
NASA
NIST
VA
Customize ...

Clear all

Display Settings: Summary, 20 per page, Sorted by Default order

Send to:

Search results

Items: 1 to 20 of 135007

<< First < Prev Page 1 of 6751 Next > Last >>

1. **SARS-CoV-2** binds platelet ACE2 to enhance thrombosis in **COVID-19**
Si Zhang, Yangyang Liu, Xiaofang Wang, Li Yang, Haishan Li, Yuyan Wang, Mengduan Liu, Xiaoyan Zhao, Youhua Xie, Yan Yang, Shenghui Zhang, Zhichao Fan, Jianzeng Dong, Zhenghong Yuan, Zhongren Ding, Yi Zhang, Liang Hu
J Hematol Oncol. 2020; 13: 120. Published online 2020 Sep 4. doi: 10.1186/s13045-020-00954-7
PMCID: PMC7471641
Article PubReader PDF–10M Cite

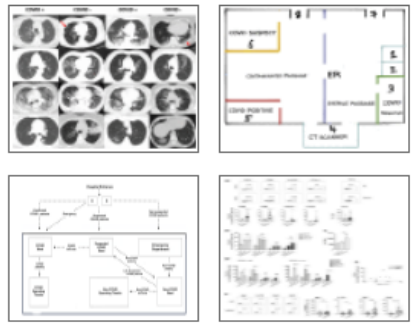
2. Immunity, endothelial injury and complement-induced coagulopathy in **COVID-19**
Luca Perico, Ariela Benigni, Federica Casiraghi, Lisa F. P. Ng, Laurent Renia, Giuseppe Remuzzi
Nat Rev Nephrol. 2020 Oct 19 : 1–19. doi: 10.1038/s41581-020-00357-4 [Epub ahead of print]
PMCID: PMC7570423
Article PubReader PDF–4.0M Cite

3. Coronavirus Disease 2019–**COVID-19**
Kuldeep Dhama, Sharun Khan, Ruchi Tiwari, Shubhankar Sircar, Sudipta Bhat, Yashpal Singh Malik, Karam Pal Singh, Wanpen Chaicumpa, D. Katterine Bonilla-Aldana, Alfonso J. Rodriguez-Morales
Clin Microbiol Rev. 2020 Oct; 33(4): e00028-20. Published online 2020 Jun 24. doi: 10.1128/CMR.00028-20
PMCID: PMC7405836
Article PubReader PDF–5.0M Cite

4. Antigen-Specific Adaptive Immunity to **SARS-CoV-2** in Acute **COVID-19** and Associations with Age and Disease Severity

Filter your results:
All (135007)
NIH grants (7645)
Embargoed (0)
Manage Filters

PMC Images search for *covid*



See more (3467)...

Find related data
Database: Select
Find items

Web of Science

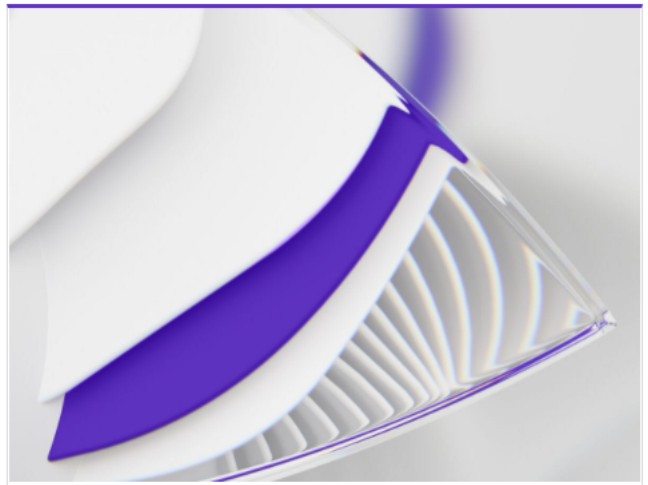
Confident research begins here.

Go to product Contact us

The Web of Science™ is the world's most trusted publisher-independent global citation database. Guided by the legacy of Dr Eugene Garfield, inventor of the world's first citation index, the Web of Science is the most powerful research engine, delivering your library with best-in-class publication and citation data for confident discovery, access and assessment.

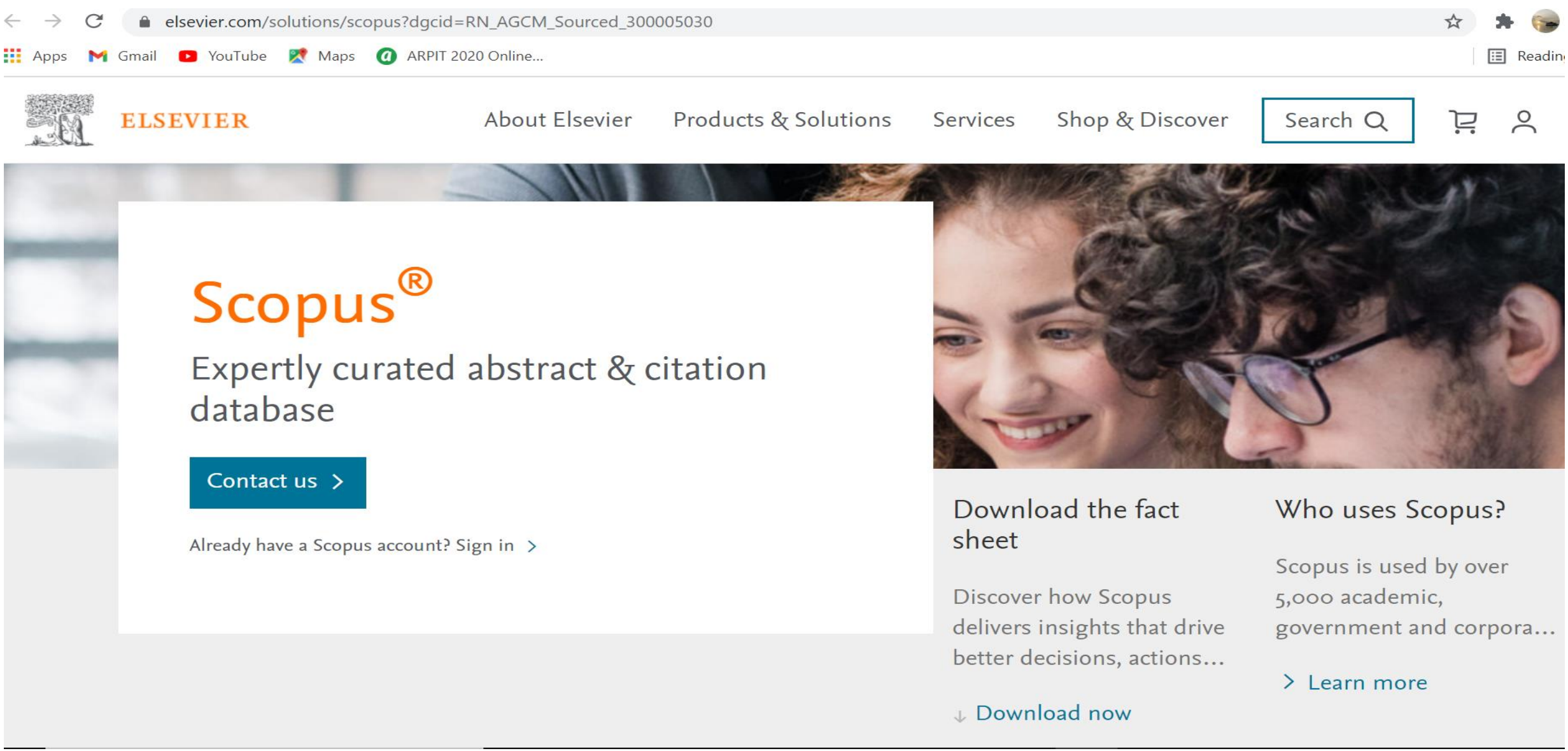
Our multidisciplinary platform connects regional, specialty, data and patent indexes to the Web of Science Core Collection™. Our comprehensive platform allows you to track ideas across disciplines and time from almost 1.9 billion cited references from over 171 million records.

Over 9,000 leading academic, corporate and government institutions and millions of researchers trust the Web of Science to produce high-quality research, gain insights and make more-informed decisions that guide the future of their institution and research strategy.



Web of Science My Research Assistant.
For wherever inspiration strikes.

Learn more



SCOPUS indexed Science Direct database of ELSEVIER publishers

sciencedirect.com

Apps Gmail YouTube Maps ARPIT 2020 Online... Reading li

ScienceDirect Journals & Books Register Sign in

Brought to you by:
Indian Institute of Technology Kanpur

Search for peer-reviewed journal articles and book chapters (including open access content)

Keywords | Author name | Journal/book title | Volume | Issue | Pag

Advanced search

FEEDBACK

Google Scholar

From Wikipedia, the free encyclopedia

Google Scholar is a freely accessible web search engine that indexes the full text or metadata of scholarly literature across an array of publishing formats and disciplines. Released in beta in November 2004, the Google Scholar index includes most peer-reviewed online academic journals and books, conference papers, theses and dissertations, preprints, abstracts, technical reports, and other scholarly literature, including court opinions and patents.[1] While Google does not publish the size of Google Scholar's database, scientometric researchers estimated it to contain roughly 389 million documents including articles, citations and patents making it the world's largest academic search engine in January 2018.[2] Previously, the size was estimated at 160 million documents as of May 2014.[3] An earlier statistical estimate published in PLOS ONE using a Mark and recapture method estimated approximately 80–90% coverage of all articles published in English with an estimate of 100 million.[4] This estimate also determined how many documents were freely available on the web.

Google Scholar has been criticized for not vetting journals and for including predatory journals in its index.[5]

Google Scholar

Type of site	Bibliographic database
Owner	Google
URL	scholar.google.com
Registration	Optional
Launched	November 20, 2004; 16 years ago
Current status	Active

Contents [hide]