Structured biological relationships data from
EmBiology

Authoritative, structured relationship data extracted against expert created ontologies, based on full-text, cross-publisher content.

Improve data-driven drug target and biomarker identification and prioritization. Perform high-throughput querying and identify broad or narrow relationships from the full-text.

The data are normalized and structured with Elsevier's GeneTree ontology (most concept types) and UniProt (proteins, protein classes and complexes). Multiple public identifiers are supported to enable integration with other datasets.

Example use cases

- Predict effective drug targets
- Improve biomarker identification and prioritization
- Disease exploration
- Custom reports and dashboards on specific targets or diseases
- Augment internal datasets with robust biological relationships
- Empower machine learning and AI with structured, interconnected data

Dataset content

The dataset currently includes 16M relations such as protein-protein interactions, effects of proteins, compounds and cells on diseases and cell processes. The information comes from full-text literature (Elsevier's ScienceDirect and other high impact publishers) and public and proprietary databases.

- Elsevier full-text: 5M+
- MEDLINE abstracts: 36M
- Non-Elsevier full-text: 2.5M
- Small molecule protein interactions from Reaxys: 1.3M
- Clinicaltrial.gov: 430K+
- 600K+ relations from public databases of protein-protein, small molecule protein interactions, MiRNA effects, SNP annotations and more.

XML Data

We offer the biology data as RDF/XML, with the option of additional instructions or consultancy on how to load it into a graph database (neo4j) or a relational database (PostgreSQL).

We also offer additional SciBite powered semantic enrichment if using the data to enrich custom proprietary content. Our Expert Services team can also provide custom reports, analysis and subsets of data, dependent on need.

For more information, talk to your
Elsevier Sales Representative today, or visit:
www.elsevier.com/solutions/datasets

ELSEVIER