

Resource Summary Report

Generated by SPARC SAWG on Sep 15, 2026

BioBERT

RRID:SCR_017547

Type: Tool

Proper Citation

BioBERT (RRID:SCR_017547)

Resource Information

URL: <https://github.com/dmis-lab/biobert>

Proper Citation: BioBERT (RRID:SCR_017547)

Description: Pre-trained biomedical language representation model for biomedical text mining. This repository provides fine-tuning codes of BioBERT, language representation model for biomedical domain, especially designed for biomedical text mining tasks such as biomedical named entity recognition, relation extraction, question answering, etc.

Synonyms: Bidirectional Encoder Representations from Transformers for Biomedical Text Mining

Resource Type: software application, software resource

Defining Citation: [DOI:10.1093/bioinformatics/btz682](https://doi.org/10.1093/bioinformatics/btz682)

Keywords: Pretrained, biomedical, language, representation, model, text, mining, bio.tools

Funding: National Research Foundation of Korea

Availability: Free, Available for download, Freely available

Resource Name: BioBERT

Resource ID: SCR_017547

Alternate IDs: biotools:biobERT

Alternate URLs: <https://github.com/naver/biobert-pretrained>, <https://bio.tools/BioBERT>

License: Apache License

Record Creation Time: 20220129T080335+0000

Ratings and Alerts

No rating or validation information has been found for BioBERT.

No alerts have been found for BioBERT.

Data and Source Information

Source: [SciCrunch Registry](#)

Usage and Citation Metrics

We found 56 mentions in open access literature.

Listed below are recent publications. The full list is available at [SPARC SAWG](#) .

Liu X, et al. (2026) Maternal and Pediatric Precision in Therapeutic Knowledge Portal (MPRINT-KP): Landscape Analysis of Pharmacology Research in Maternal and Pediatric Patient Populations. *Pharmacotherapy*, 46(3), e70096.

Yang L, et al. (2026) Multimodal Collaborative Modeling of Molecular Structures and Biomedical Text for Accurate Drug-Drug Interaction Extraction. *Biomedicines*, 14(6).

Alhassan A, et al. (2026) Recognition and linking of discontinuous named entities in healthcare: a comparative performance analysis. *Frontiers in digital health*, 8, 1758921.

Alhabiti S, et al. (2025) Predicting Behavioral Determinants of Health from Clinical Text Using Transformer Models and BiLSTM. *medRxiv : the preprint server for health sciences*.

Xu Y, et al. (2025) A multimodal knowledge-enhanced whole-slide pathology foundation model. *Nature communications*, 16(1), 11406.

Hier DB, et al. (2025) A simplified retriever to improve accuracy of phenotype normalizations by large language models. *Frontiers in digital health*, 7, 1495040.

Saadh MJ, et al. (2025) Advanced machine learning framework for enhancing breast cancer diagnostics through transcriptomic profiling. *Discover oncology*, 16(1), 334.

Raciti D, et al. (2025) Characterization and automated classification of sentences in the biomedical literature: a case study for biocuration of gene expression and protein kinase activity. *Database : the journal of biological databases and curation*, 2025.

Schoenmaker L, et al. (2025) Toward Assay-Aware Bioactivity Model(er)s: Getting a Grip on Biological Context. *Journal of chemical information and modeling*, 65(13), 7013.

Zhang S, et al. (2025) SCATrans: semantic cross-attention transformer for drug-drug interaction predication through multimodal biomedical data. *BMC bioinformatics*, 26(1), 157.

Theodoropoulos C, et al. (2025) Reduction of supervision for biomedical knowledge discovery. BMC bioinformatics, 26(1), 225.

Jimenez-Maggiora GA, et al. (2025) Artificial intelligence-enabled safety monitoring in Alzheimer's disease clinical trials. The journal of prevention of Alzheimer's disease, 12(1), 100002.

Ghelfi A, et al. (2025) Hayai-Annotation: A functional gene prediction tool that integrates orthologs and gene ontology for network analysis in plant species. Computational and structural biotechnology journal, 27, 117.

Happi Happi BG, et al. (2025) GRU-SCANET: unleashing the power of GRU-based sinusoidal capture network for precision-driven named entity recognition. Bioinformatics advances, 5(1), vba096.

Rodin D, et al. (2025) Medical Cannabis Library: development of a curated database for research articles on cannabis therapeutic activity. Journal of cannabis research, 7(1), 39.

Wang H, et al. (2024) Mitophagy in Alzheimer's Disease: A Bibliometric Analysis from 2007 to 2022. Journal of Alzheimer's disease reports, 8(1), 101.

Huang DL, et al. (2024) A Combined Manual Annotation and Deep-Learning Natural Language Processing Study on Accurate Entity Extraction in Hereditary Disease Related Biomedical Literature. Interdisciplinary sciences, computational life sciences, 16(2), 333.

Rehana H, et al. (2024) Evaluating GPT and BERT models for protein-protein interaction identification in biomedical text. Bioinformatics advances, 4(1), vbae133.

Alamro H, et al. (2024) BioBBC: a multi-feature model that enhances the detection of biomedical entities. Scientific reports, 14(1), 7697.

Sharma T, et al. (2024) Exploring COVID-related relationship extraction: Contrasting data sources and analyzing misinformation. Heliyon, 10(5), e26973.