

Research software projects of EpiGraphDB

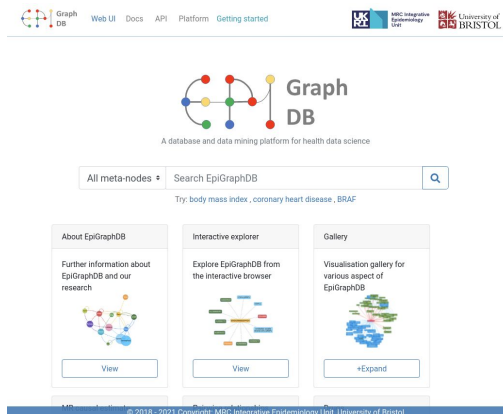
Yi Liu yi6240.liu@bristol.ac.uk

Senior Research Associate in Health Data Science

IEU software coordinators lightning talk
14 July 2021

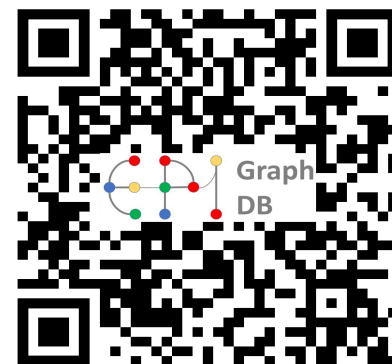
- The EpiGraphDB research project
- EpiGraphdb as a research software project
 - Architecture and components
 - Frameworks and technologies
 - Development and maintenance
- New features

<https://epigraphdb.org>

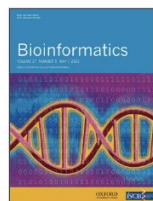


The screenshot shows the EpiGraphDB website homepage. At the top, there are navigation links: Web UI, Docs, API, Platform, and Getting started. The main header features the Graph DB logo and the text 'A database and data mining platform for health data science'. Below this is a search bar with the text 'Search EpiGraphDB' and a search icon. A suggestion below the search bar reads 'Try: body mass index, coronary heart disease, BRAF'. The main content area is divided into three columns: 'About EpiGraphDB' with a 'View' button, 'Interactive explorer' with a 'View' button, and 'Gallery' with an '+Expand' button. The footer contains copyright information: '© 2018 - 2021 Copyright MRC Integrative Epidemiology Unit, University of Bristol'.

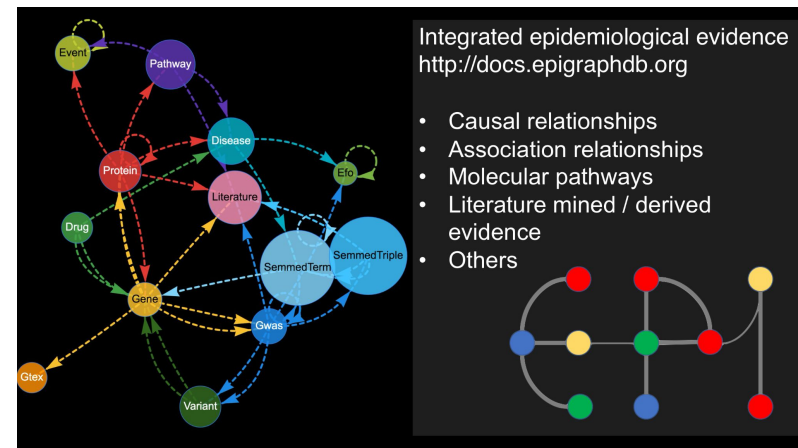
<https://docs.epigraphdb.org/slides/>



- Integration of evidence and knowledge to support data mining epidemiological relationships (PI Tom Gaunt)
 - further details on previous talks (<https://docs.epigraphdb.org/slides/>) and documentation
- Internal research
 - Platform (Liu, Elsworth, et al., Gaunt, 2021, Bioinformatics)
 - Evidence triangulation
 - Drug target prioritization
 - MR of molecular traits
 - Pathways, pleiotropy and mechanistic inference
 - Phenotype mapping
- External research studies start emerging
- **Get in touch if interested**



Volume 37, Issue 9
1 May 2021

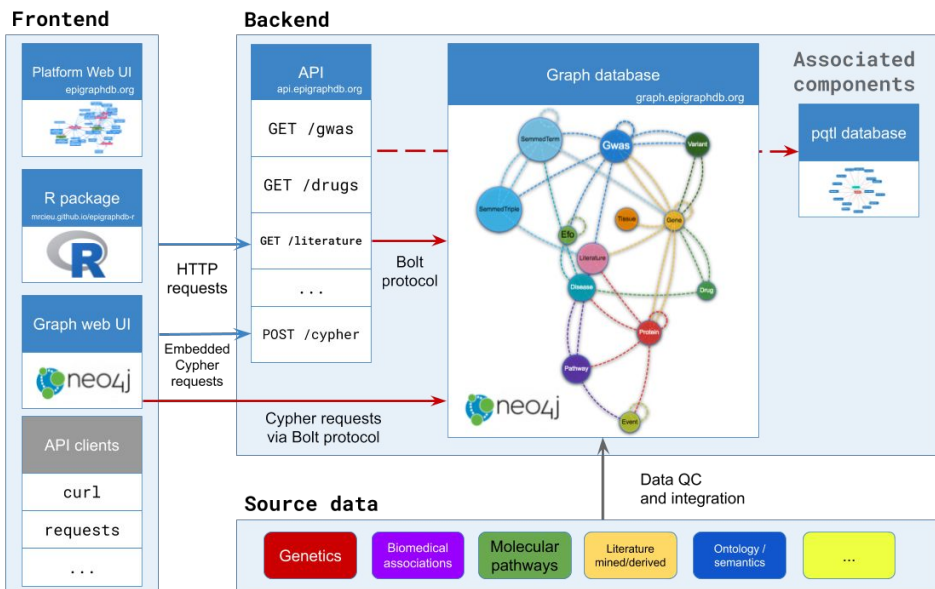


EpiGraphDB: a database and data mining platform for health data science

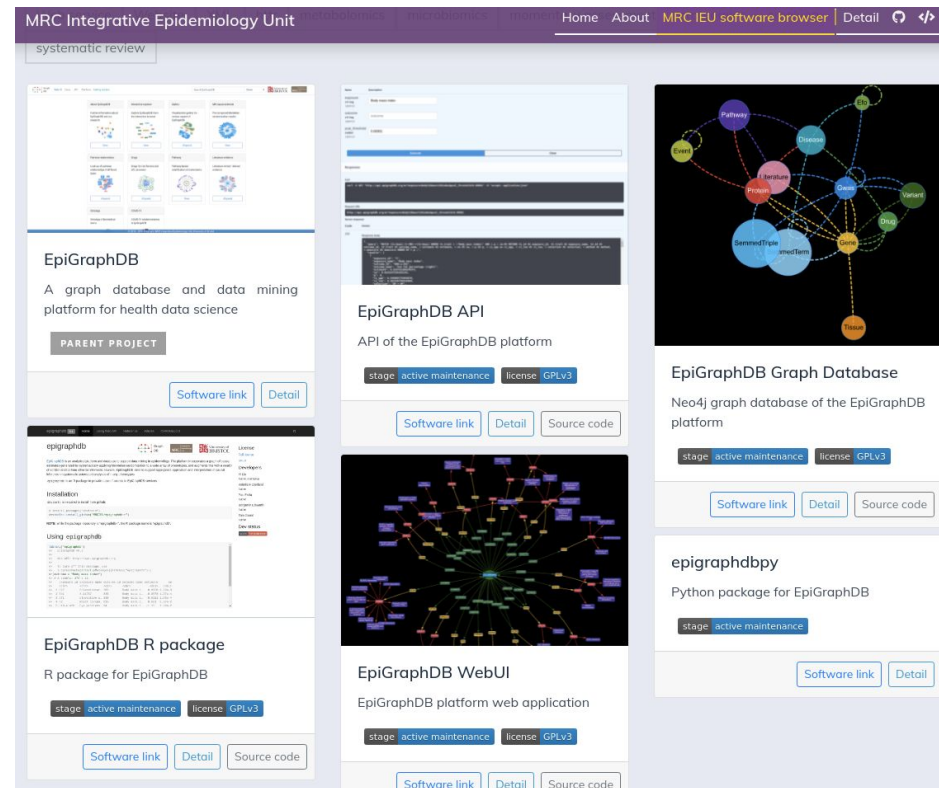
Yi Liu ✉, Benjamin Elsworth ✉, Pau Erola, Valeriia Haberland, Gibran Hemani, Matt Lyon, Jie Zheng, Oliver Lloyd, Marina Vabistsevits, Tom R Gaunt ✉
Author Notes

Bioinformatics, Volume 37, Issue 9, 1 May 2021, Pages 1304–1311,
<https://doi.org/10.1093/bioinformatics/btaa961>

Published: 24 November 2020 **Article history** ▼



- Platform
 - web <https://epigraphdb.org>
 - API <https://api.epigraphdb.org>
 - Docs <https://docs.epigraphdb.org>
- R package
 - CRAN <https://cloud.r-project.org/web/packages/epigraphdb/>
 - Docs <https://mrcieu.github.io/epigraphdb-r>
- Code
 - <https://github.com/mrcieu/epigraphdb>



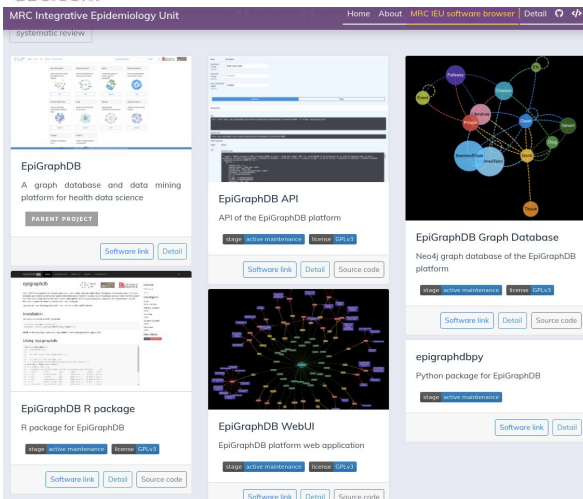
The screenshot shows the MRC Integrative Epidemiology Unit software browser interface. It features a navigation bar with 'Home', 'About', 'MRC IEU software browser', and 'Detail'. The main content area displays a grid of software components:

- EpiGraphDB:** A graph database and data mining platform for health data science. Includes 'PARENT PROJECT' and links for 'Software link' and 'Detail'.
- EpiGraphDB API:** API of the EpiGraphDB platform. Includes 'stage active maintenance' and 'license GPLv3' buttons, along with 'Software link', 'Detail', and 'Source code' links.
- EpiGraphDB Graph Database:** Neo4j graph database of the EpiGraphDB platform. Includes 'stage active maintenance' and 'license GPLv3' buttons, along with 'Software link', 'Detail', and 'Source code' links.
- epigraphdbpy:** Python package for EpiGraphDB. Includes 'stage active maintenance' and 'license GPLv3' buttons, along with 'Software link' and 'Detail' links.
- EpiGraphDB R package:** R package for EpiGraphDB. Includes 'stage active maintenance' and 'license GPLv3' buttons, along with 'Software link', 'Detail', and 'Source code' links.
- EpiGraphDB WebUI:** EpiGraphDB platform web application. Includes 'stage active maintenance' and 'license GPLv3' buttons, along with 'Software link', 'Detail', and 'Source code' links.

Search “**epigraphdb**” at MRC IEU github pages
<https://mrcieu.github.io>



WHO head: 'Our key message is: test, test, test' - B...
bbc.com



MRC Integrative Epidemiology Unit

systematic review

EpiGraphDB
A graph database and data mining platform for health data science
CURRENT PROJECT
[Software link] [Detail]

EpiGraphDB API
API of the EpiGraphDB platform
[stage] [active maintenance] [new] [GitHub]
[Software link] [Detail] [Source code]

EpiGraphDB Graph Database
Neo4j graph database of the EpiGraphDB platform
[stage] [active maintenance] [new] [GitHub]
[Software link] [Detail] [Source code]

epigraphdbpy
Python package for EpiGraphDB
[stage] [active maintenance]
[Software link] [Detail]

EpiGraphDB R package
R package for EpiGraphDB
[stage] [active maintenance] [new] [GitHub]
[Software link] [Detail] [Source code]

EpiGraphDB WebUI
EpiGraphDB platform web application
[stage] [active maintenance] [new] [GitHub]
[Software link] [Detail] [Source code]

Python:

- Unit-testing: **pytest**
- Static analysis: **mypy**, **flake8**
- Code formatting: **black**, **autoflake**, **isort**

R:

- Unit-testing: **testthat**
- Static analysis: **lintr**
- Code formatting: **styler**
- CI/CD: **GitHub Actions**

JS:

- Static analysis: **eslint**
- Code formatting: **prettier**

Jupyter:

- Automation: **papermill**
- Utilities: **nb_black**, **nbqa**

https://epigraphdb.org/entity?meta_node=Gwas&id=ieu-b-40

- V 1.0 release

- Data updates
- Codebase release
- Docs at P4 blog
(<https://biocompute.org.uk>)

- V 1.1 ongoing

- **web:** Accessibility improvement
 - Entity search
 - Functionality search
 - Overview page
of biomedical entities
- **platform:** Entity recommendation engine based on NLP and phenotype mapping
- More coming

EpiGraphDB node: (Gwas {_id: "ieu-b-40"})

Entity meta data

meta entity: (Gwas)
id: **ieu-b-40**
name: **body mass index**
source: ["OpenGWAS-2020-10-13"]

Entity properties

year: 2018.0
mr: 1
author: Yengo, L
consortium: GIANT
sex: Males and Females
pmid: 30124842.0
population: European
unit: SD
sample_size: 681275.0
nsnp: 2336260
build: HG19/GRCh37
trait: body mass index

EpiGraphDB platform resources

Resources that are associated with (Gwas {_id: "ieu-b-40"}) on the EpiGraphDB platform.

WebUI topic views

Topic views that are associated with (Gwas {_id: "ieu-b-40"}) on the EpiGraphDB WebUI.

literature_trait

param
Literature evidence regarding
GWAS traits

API endpoints

API endpoints that are associated with (Gwas {_id: "ieu-b-40"}) on the EpiGraphDB API.

GET /literature/gwas

param
GET /literature/gwas

R package functions

Functions that are associated with (Gwas {_id: "ieu-b-40"}) in the *epigraphdb* R package.

literature_gwas

param
literature_gwas

query_epigraphdb(route="/literature/gwas", ...)

param
query_epigraphdb: to API endpoint
"GET /literature/gwas"

Outline ^

EpiGraphDB platform resources

- WebUI topic views
- API endpoints
- R package functions

Connected entities

- Overview
- Details

Similar entities

- Entities with similar names
- Entities with similar semantic representation

NLP

<https://api.epigraphdb.org/#/NLP>

GET /nlp/query/text Get Nlp Query Text

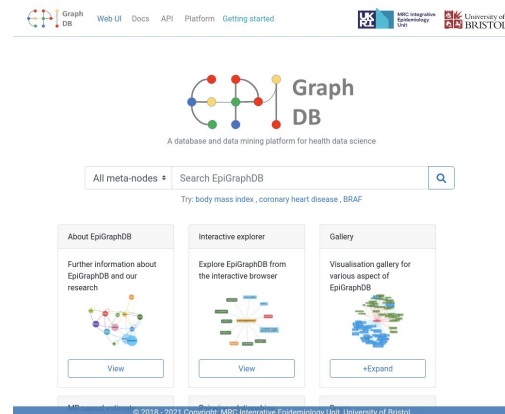
GET /nlp/query/entity Get Nlp Query Ent

GET /nlp/query/entity/encode Get Nlp Query Ent Encode

- Find documentation on
 - Platform and data <https://docs.epigraphdb.org>
 - Usage examples <https://github.com/MRCIEU/epigraphdb>
 - R package <https://mrcieu.epigraphdb.org/epigraphdb-r>

- Get in touch
 - Mailbox feedback@epigraphdb.org
 - Issues <https://github.com/MRCIEU/epigraphdb/issues>
 - Tom tom.gaunt@bristol.ac.uk

<https://epigraphdb.org>



The screenshot shows the EpiGraphDB website homepage. At the top, there are logos for Graph DB, MRC Integrative Epidemiology Unit, and University of Bristol. Below the logos is a navigation menu with links for Web UI, Docs, API, Platform, and Getting started. The main content area features the EpiGraphDB logo and a search bar with the text "Search EpiGraphDB" and a search icon. Below the search bar, there are three main sections: "About EpiGraphDB", "Interactive explorer", and "Gallery". Each section has a brief description and a "View" or "+Expand" button. The footer contains copyright information for 2018-2021.

<https://docs.epigraphdb.org/slides/>

