Ethical Analysis of a Biomedical Reasoning System: Hetionet / Project Rephetio

Biomedical AI & Ethics Ethical Report
Munongedzi Mabhoko

Overview

What the system is. Hetionet is a large, heterogeneous biomedical knowledge graph integrated from curated sources. Project Rephetio learns path-based features (metapaths) over Hetionet to prioritize drug–disease “treats” links for repurposing. The graph is implemented as a property graph (Neo4j) with typed nodes/edges and per-edge provenance in the public release.

Representational Logic

Paradigm. Property graph with typed nodes (metanodes) and edges (metaedges). Semantics arise from the schema and the ontological identifiers imported from source databases. While not OWL DL-native, Hetionet concepts can be aligned to OWL/RDF exports where description logic reasoning is needed. Common metanodes include Drug, Disease, Gene, Pathway, Phenotype, and ClinicalTrial, relations include treats, associates, binds, up/down-regulates, and participates_in.

Reasoning mechanics. Rephetio engineers metapath features that summarize typed connectivity between drug and disease nodes; a regularized logistic regression then predicts probabilities for a “treats” edge. This is statistical relational learning not logical entailment, so outputs are hypotheses requiring downstream validation.

Executive summary

Hetionet with Project Rephetio is a heterogeneous biomedical knowledge graph used to generate hypotheses about potential drug–disease treatments. It aggregates curated databases and literature into a property graph and learns path-based features to predict whether a drug may treat a disease. The system is valuable for research triage and for prioritizing repurposing candidates. It is not suitable for any clinical decision or patient allocation task without a formal program of prospective evaluation, transparency artifacts, and post-deployment monitoring. The remainder of this report explains the representational logic, evidentiary basis, foreseeable ethical risks, and the safeguards that would be required before any expansion of scope.

System described in plain terms

Hetionet represents biomedical entities such as drugs, diseases, genes, pathways, phenotypes, and clinical trials as nodes. Relationships such as treats, associates, binds, and regulates are edges. Project Rephetio computes metapaths that summarize typed connectivity between two nodes and uses a regularized model to estimate the probability that a drug treats a disease. The output is a ranked list of candidates with scores that reflect statistical association in the integrated graph. These scores are not mechanistic explanations and they are not clinical evidence. They are research signals that can guide literature review and experimental planning.

Representational logic and its implications

The underlying representation is a property graph with typed nodes and edges. Many node identifiers are aligned to established ontologies such as MeSH, Disease Ontology, ChEBI, Gene Ontology, and HGNC. The graph is not natively expressed in OWL. It can, however, be exported to RDF or constrained with SHACL or similar profiles if stronger schema guarantees are needed. In the current form, reasoning is statistical rather than deductive. The model infers likely links based on patterns in the graph. That choice brings scalability and empirical performance, but it weakens formal guarantees and relies on the quality and balance of the sources that feed the graph.

What counts as evidence here

Evidence in Hetionet is the set of edges and their provenance. Each edge should be traceable to a source database release or a publication. Model predictions reflect the distribution of evidence across that graph. Since biomedical knowledge is not uniformly curated, the evidence base is uneven. Disease areas that attract more research funding and publication will be overrepresented. Rare diseases, underfunded conditions, and research communities outside dominant publication venues can be underrepresented. These imbalances directly affect the model’s likelihood scores, which means the model can amplify historical attention rather than true clinical potential.

Ethical analysis

Justice and bias

There are two layers of bias risk. The first is coverage bias. Overrepresentation of certain diseases, drug classes, and protein targets will lead to systematic elevation of candidates in those areas. The second is historical and structural bias. If the evidence base reflects inequities in who is studied or which populations are included in trials, the model’s outputs will mirror those inequities. A common pattern in population health tools is the use of utilization or cost as a proxy for need. If any future extensions to this system incorporate such proxies, the outputs will recommend fewer resources for populations who historically face barriers to access. Even without those proxies, the graph inherits the research community's focus and omissions.

Transparency and interpretability

Metapaths are more interpretable than deep graph embeddings but are still non-intuitive to many clinical stakeholders. Without a model card, per-prediction rationales, and confidence information, end users can easily over-trust a score or read it as causal evidence. The system must present the supporting subgraph and the specific metapaths that contribute most to each score. It must also present uncertainty and calibration so that users understand how the score should be read.

Accountability and misuse

There is a foreseeable misuse risk in treating a ranked list as a clinical recommendation or as a substitute for systematic review. The outputs are hypotheses that require expert adjudication and, where appropriate, experimental or trial evidence. Any transfer of this system into decision support without guardrails would shift accountability from human judgment to an opaque statistical artifact. Accountability requires documented roles, decision logs, and a clear statement that the tool cannot by itself justify therapeutic action.

Privacy and data protection

Hetionet as published relies on curated public sources rather than identifiable patient data. If future versions ingest cohort or EHR derivatives, privacy risks increase. There is a real risk of linkage attacks when identifiers from different sources can be connected. Responsible use requires data minimization, purpose binding, and a privacy impact assessment before any integration of patient-level data. Access should be role-based, and outputs that could reveal membership in a sensitive cohort must be suppressed or differentially protected.

Security and integrity of the graph

Knowledge graphs are vulnerable to poisoning through incorrect or malicious edges as well as to supply chain risks in the ingestion pipeline. Since model predictions rely on graph topology, even small targeted changes can perturb ranked outputs. The ingestion pipeline should sign and version each release, log transformations immutably, and quarantine auto-extracted relations until they are reviewed. Model releases should be tied to frozen graph snapshots to ensure reproducibility and to enable rollback if integrity checks fail.

Validity and formal guarantees

There are no formal guarantees of correctness, safety, or generalizability to clinical settings. The model supports statistical generalization on retrospective data splits. This is appropriate for research ranking, not for decisions that affect patient care. If the scope ever changes, assurance must be built around prospective evaluation, calibration, and stability under graph updates. Time-based splits are essential because biomedical knowledge changes over time, and performance that depends on leakage from future knowledge will not translate.

Ethical acceptability under defined scope

Under a research scope that includes literature triage and candidate prioritization for laboratory study, the system is ethically acceptable with the following conditions. Each prediction must include provenance and a human-readable rationale. Users must be trained to treat scores as hypotheses. Releases must be versioned and reproducible. Reports must include subgroup coverage analyses so that research agendas are not skewed toward already well-studied domains. Any use that influences clinical decisions, resource allocation, or patient stratification is not acceptable without a formal program of prospective trials, bias auditing, and continuous monitoring.

Risk Controls & Recommendations

Data governance. Maintain per-edge provenance, licenses, and versioning, periodically refresh identifier mappings and document coverage/gaps.
Transparency. Publish metapath definitions, coefficients, and example rationales for top predictions; include confidence and stability analyses.
Fairness safeguards. Audit performance across disease areas and populations, avoid problematic proxies, document demographic coverage of sources.
Shift & reliability. Stress-test against KG updates, ontology changes, and node/edge dropout, report OOD behavior.
Clinical evaluation. For any patient-facing use, follow CONSORT-AI for prospective trials, document human–AI interaction and failure modes.

Recommendations that follow from the analysis

The system owners should publish a data card that describes sources, licenses, coverage, update cadence, and known gaps, and a model card that documents features, training procedure, calibration, and failure modes. The serving interface should display the top contributing metapaths and the supporting subgraph for each prediction. The evaluation protocol should use time-split folds and should report calibration and stability, not only ranking metrics. Coverage analyses should compare evidence density to burden of disease so that users can see where the tool is likely to underperform. Security controls should include signed releases, immutable ingest logs, and poison-resilience tests. If the team contemplates any clinical or operational use, they should adopt a safety case that links identified hazards to specific controls and to evidence that those controls are effective.

Conclusion

Hetionet and Project Rephetio provide a disciplined way to surface promising drug repurposing hypotheses from disparate biomedical sources. The strengths are explainable path-based reasoning, rich provenance, and the ability to unify many curated datasets. The risks arise from uneven coverage, inherited bias, and the temptation to read associative predictions as causal or clinical. Within a research program that treats outputs as hypotheses and that invests in transparency, provenance, and evaluation, the system is ethically defensible and practically useful. Outside that scope, particularly in any clinical or allocation context, the system would require a level of assurance and oversight that is not yet present and must be developed before deployment.

Biomedical Knowledge Graph — Core Components