Charles Tapley Hoyt

cthoyt.com · in linkedin.com/in/cthoyt · ⌂ github.com/cthoyt · ⬤ @ cthoyt · ||||| Q47475003
ⓘ https://orcid.org/0000-0003-4423-4370 · Ⓖ PjrpzUIAAAAJ

Fields of Work

Bioinformatics, Pathway Analysis, Machine Learning, Natural Language Processing, Ontology, Knowledge Graph, Proteochemometrics, Drug Repositioning, Systems Biology, Cheminformatics

Software

SSSOM-py, *Python package*.
https://github.com/mapping-commons/sssom-py/, https://mapping-commons.github.io/sssom-py/

PyBEL, *Python library for Biological Expression Language*.
https://github.com/pybel/pybel, https://pybel.github.io

Integrated Network and Dynamical Reasoning Assembler, *Text mining and systems biology modeling software*.
https://github.com/sorgerlab/indra, https://www.indra.bio/

PyKEEN, *Learning, evaluation and applications of knowledge graph embeddings.*.
https://github.com/pykeen/pykeen, https://pykeen.github.io

Spoken Languages

English, German

Articles

2022

1. **Unifying the Identification of Biomedical Entities with the Bioregistry**.
   DOI:10.1101/2022.07.08.499378 bioRxiv:2022.07.08.499378

2. **Ontology Development Kit: a toolkit for building, maintaining, and standardising biomedical ontologies**.
   arXiv:2207.02056

3. **ProtSTonKGs: A Sophisticated Transformer Trained on Protein Sequences, Text, and Knowledge Graphs**.

4. **A Unified Framework for Rank-based Evaluation Metrics for Link Prediction in Knowledge Graphs**.
   arXiv:2203.07544

5. **PyBioPAX: biological pathway exchange in Python**.
   DOI:10.21105/joss.04136

6. **An Open Challenge for Inductive Link Prediction on Knowledge Graphs**.
   arXiv:2203.01520

7. **ChemicalX: A Deep Learning Library for Drug Pair Scoring**.
   arXiv:2202.05240

8. **STonKGs: A Sophisticated Transformer Trained on Biomedical Text and Knowledge Graphs**.
   PMID:34986221 DOI:10.1093/bioinformatics/btac001 bioRxiv:2021.08.17.456616

2021

1. **A Simple Standard for Sharing Ontological Mappings (SSSOM)**.
   arXiv:2112.07051

2. **Gilda: biomedical entity text normalization with machine-learned disambiguation as a service**.
   DOI:10.1101/2021.09.10.459803 bioRxiv:2021.09.10.459803

3. **The role of metadata in reproducible computational research**.
   PMID:34553169 DOI:10.1016/j.patter.2021.100322 arXiv:2006.08589

4. **Integrating multi-omics data reveals function and therapeutic potential of deubiquitinating enzymes**.
   DOI:10.1101/2021.08.06.455458 bioRxiv:2021.08.06.455458

5. **Wavelet-Packet Powered Deepfake Image Detection**.
   arXiv:2106.09369

6. **Understanding the Performance of Knowledge Graph Embeddings in Drug Discovery**.
   arXiv:2105.10488

7. **A Review of Biomedical Datasets Relating to Drug Discovery: A Knowledge Graph Perspective**.
   arXiv:2102.10062

8. **A Systems Biology Approach for Hypothesizing the Effect of Genetic Variants on Neuroimaging Features in Alzheimer's Disease**.
   PMID:33554913 DOI:10.3233/jad-201397

9. **PyKEEN 1.0: A Python Library for Training and Evaluating Knowledge Graph Embeddings**.

10. **Leveraging Structured Biological Knowledge for Counterfactual Inference: a Case Study of Viral Pathogenesis**.
    arXiv:2101.05136

2020

1. **CLEP: A Hybrid Data- and Knowledge- Driven Framework for Generating Patient Representations**.
   DOI:10.1101/2020.08.20.259226 bioRxiv:2020.08.20.259226

2. **Extension of Roles in the ChEBI Ontology**.
   DOI:10.26434/chemrxiv.12591221

3. **The Minimum Information about a Molecular Interaction Causal Statement (MI2CAST)**.
   PMID:32637990 DOI:10.1093/bioinformatics/btaa622

4. **Bringing Light Into the Dark: A Large-scale Evaluation of Knowledge Graph Embedding Models Under a Unified Framework**.
   arXiv:2006.13365

5. **GuiltyTargets: Prioritization of Novel Therapeutic Targets with Deep Network Representation Learning**.
   PMID:32750869 DOI:10.1109/tcbb.2020.3003830 bioRxiv:521161

6. **PS4DR: a multimodal workflow for identification and prioritization of drugs based on pathway signatures**.
   PMID:32503412 DOI:10.1186/s12859-020-03568-5

7. **A Computational Approach for Mapping Heme Biology in the Context of Hemolytic Disorders**.
   PMID:32211383 DOI:10.3389/fbioe.2020.00074 bioRxiv:804906

2019

1. **The Impact of Pathway Database Choice on Statistical Enrichment Analysis and Predictive Modeling**.
   PMID:31824580 DOI:10.3389/fgene.2019.01203 bioRxiv:654442

2. **Quantifying mechanisms in neurodegenerative diseases (NDDs) using candidate mechanism perturbation amplitude (CMPA) algorithm**.
   PMID:31604427 DOI:10.1186/s12859-019-3101-1

3. **The KEEN Universe**.
   DOI:10.1007/978-3-030-30796-7$_1$

4. **Predicting Missing Links Using PyKEEN**.

5. **RatVec: A General Approach for Low-dimensional Distributed Vector Representations via Domain-specific Rational Kernels**.

6. **BioKEEN: a library for learning and evaluating biological knowledge graph embeddings**.
   PMID:30768158 DOI:10.1093/bioinformatics/btz117 bioRxiv:475202

7. **PathMe: merging and exploring mechanistic pathway knowledge**.
   PMID:31092193 DOI:10.1186/s12859-019-2863-9 bioRxiv:451625

8. **Integration of Structured Biological Data Sources using Biological Expression Language**.
   DOI:10.1101/631812 bioRxiv:631812

9. **ComPath: an ecosystem for exploring, analyzing, and curating mappings across pathway databases**.
   PMID:30564458 DOI:10.1038/s41540-018-0078-8 bioRxiv:353235

10. **Re-curation and rational enrichment of knowledge graphs in Biological Expression Language**.
    PMID:31225582 DOI:10.1093/database/baz068 bioRxiv:536409

11. **Challenges of Integrative Disease Modeling in Alzheimer39;s Disease**.
    PMID:31993440 DOI:10.3389/fmolb.2019.00158

2018

1. **BEL Commons: an environment for exploration and analysis of networks encoded in Biological Expression Language**.
   PMID:30576488 DOI:10.1093/database/bay126 bioRxiv:288274

2. **BEL2ABM: agent-based simulation of static models in Biological Expression Language**.
   PMID:29949955 DOI:10.1093/bioinformatics/bty107

3. **PyBEL: a Computational Framework for Biological Expression Language**.
   PMID:29048466 DOI:10.1093/bioinformatics/btx660

4. **A systematic approach for identifying shared mechanisms in epilepsy and its comorbidities**.
   PMID:29873705 DOI:10.1093/database/bay050 bioRxiv:269860

2017

1. **A Case Study on the Parametric Occurrence of Multiple Steady States**.
   DOI:10.1145/3087604.3087622 arXiv:1704.08997

2014

1. **Repurposing human PDE4 inhibitors for neglected tropical diseases: design, synthesis and evaluation of cilomilast analogues as Trypanosoma brucei PDEB1 inhibitors**.
   PMID:25127163 DOI:10.1016/j.bmcl.2014.07.063

▬▬ Reviews

1. **The LOTUS initiative for open knowledge management in natural products research** (2022).
   PMID:35616633 DOI:10.7554/elife.70780

2. **New reasons for biologists to write with a formal language** (2022).
   DOI:10.1093/database/baac039

3. **PecanPy: a fast, efficient and parallelized Python implementation of node2vec** (2021).
   DOI:10.1093/bioinformatics/btab202

▬▬ Acknowledgements

1. **Automated assembly of molecular mechanisms at scale from text mining and curated databases** (2022).
   DOI:10.1101/2022.08.30.505688 bioRxiv:2022.08.30.505688

2. **PecanPy: a fast, efficient and parallelized Python implementation of node2vec** (2021).
   DOI:10.1093/bioinformatics/btab202

3. **The status of causality in biological databases: data resources and data retrieval possibilities to support logical modeling** (2020).
   PMID:33378765 DOI:10.1093/bib/bbaa390

4. **Multimodal Mechanistic Signatures for Neurodegenerative Diseases (NeuroMMSig): a web server for mechanism enrichment** (2017).
   PMID:28651363 DOI:10.1093/bioinformatics/btx399