

The Hacking Tribunal

Future Event Prediction



theHackingTribunal:~\$

Future Event Prediction



Madison Conley

Computer Science, *MVSU*



Nadia Rapheal

Computer Science, *WSSU*



Josiah Johnson

Computer Science, *WSSU*



Usir Jackson

Computer Science, *Claflin*



THE HACKING TRIBUNAL
PREDICTING THE FUTURE THROUGH AI

theHackingTribunal:~\$

Future Event Prediction



THE HACKING TRIBUNAL
PREDICTING THE FUTURE THROUGH AI

Project Theme:

FutureVision: Predicting the Future of Biology through Knowledge Evolution

Biology has always been a critical field to **understanding and improving humankind**. As we evolve, so does the study of life, with studies macroscopic, microscopic, and everywhere in between. Being able to **predict the next-generation breakthroughs** can only happen when we understand the pillars upon it stands now, and the **foundational innovations it has gone through**.

Our theme in this hackathon is to **predict the future of biological studies** by **analyzing trends of innovation found in the past**. Through this process, we will create forecasted projections upon what we believe is the future of biology in developments, breakthroughs, and emerging fields.

Project Goals:

Develop an AI-assisted forecasting platform that uses encyclopedic knowledge of past trends in biology to produce forecasted emerging concepts and innovation opportunities.

- Build a **biological knowledge graph** based on pages from Wikipedia and WikiData
- Identify **historical patterns** of scientific development (when they occurred, what they impacted, where it has lead)
- Detect **convergence** between biological disciplines (when they meet, at what point in history, and why)
- Generate evidence-based future **biological trend forecasts.**
- Demonstrate HPC techniques for **graph processing and analysis**
- Deliver **fully documented** deliverables

Roles:

Usir - Data Collection

Data + Knowledge Graph Lead

- identifying **strong articles** from Wikipedia and WikiData
- **extracting** biological concepts, categories, related articles, and WikiData properties
- converting information into **nodes** [content] and **edges** [connections]
- ensuring **high data quality** (no duplicates, good relationships)

Nadia - Historical Analysis

Graph Analytics & Prediction Lead

- identify important and growing concepts and **important nodes** for past and future
- detecting **convergence** at the intersection of large concepts
- creating **scoring systems** for breakthroughs and keywords
- discover which **trends and concepts** are relevant to the projection

Roles:

Madison - AI Assisted Interpretation

AI Research & Insight Lead

- **utilizing LLMs** to summarize graph findings and generate insight reports
- **verifying** historical trends and forecast explanations with **confidence levels**
- generating insights in **natural language** sentences with multiple data input

Josiah - Visualization, Project Management

Visualization Development, Project Manager

- building **interactive graphs**, timelines, and forecast displays
- creating the **demo interface** (website, dashboard, etc)
- **preparing** the readme, poster, presentation, and demo
- ensuring the **project is organized** and tasks are delegated

Project Milestones and Task Breakdown I:

[1] Project Foundation and Repository Setup

- Create **GitHub repository**, populate README
- Set up **Google Colab** notebooks
- Create a shared config cell in notebooks with Wikipedia APIs for all to use

[2] Biological Seed Dataset Creation

- Select 5-10 Wikipedia articles to use for **starting nodes**
- Use WikiAPI to pull text, wikilinks, and categories recursively to **depth 3**
- Use "instance of", "part of", and other **relations**
- Feed through **token parser** to produce concept-relation pairs

[3] Knowledge Graph Construction

- Represent each biological concept as a **graph node** with important metadata
- Create **directed edges** from wikilinks and WikiData semantic relationships, add weight
- Store completed graph in **NetworkX** [maybe in interactive graph?]

Project Milestones and Task Breakdown II:

[4] Historical Evolution Analysis

- Pull Wikipedia API's **"first-edit" timestamp** for birthdate
- Build a timeline dataset, for each node record **first year of appearance** and **expansion**
- Compute graph centrality metrics (degree, betweenness, closeness?) across **historical snapshots**
- Identify **historical** innovation pathways (i.e. DNA to PCR to CRISPR)

[5] Convergence Detection System

- Apply community detection to segment graph into biological **sub-field clusters**
- Track inter-cluster edge growth, when **cross-links** start accelerating then flag it
- **Rank convergence hotspots** by rate of cross-cluster edge addition, save best as key findings
- Prompt LLM with top-5 convergence pair subgraph to input for **natural language** output

[6] AI Forecasting Engine - RAG Approach

- Adopt **RAG** to use the knowledge graph and Wiki article text for corpus
- Design **forecasting prompt**, with subgraph of top-converging nodes and growth metrics
- Create new keyword list, **concepts under-represented** but predicted to grow anyways
- Produce **Future Trend Report**, top 10 predicted breakthroughs (thoughts, reasoning, score)

Project Milestones and Task Breakdown II:

[7] Visualization and User Experience

- Build an **interactive knowledge graph explorer**, with nodes filterable by sub field cluster, time period, and convergence score, with hover-card details on each concept
- Create a **dashboard** showing concept emergence
- Design a convergence **heatmap** and a radar chart of predicted breakthrough zones

[8] Final Submission Package

- Finish updating **complete README file** with all goals, system architecture, environment setup, and reproduction guide
- Design **project poster** covering important information as outlined
- **Complete commenting** and annotation in Colab notebooks
- **Practice demo** with live-visualization and key findings before submit

Technologies Used:

Wikipedia API

Python

Pandas

NetworkX

HuggingFace

JSON

PyVis

GitHub

Google Colab

Matplotlib

Gantt Chart Timeline

4-Day Hackathon Gantt Chart (Condensed)

Workstream	Owner	Day 1	Day 2	Day 3	Day 4
Project Setup & Architecture	Lead/Integration				
Data Collection & JSON Pipeline	Data Engineer				
Knowledge Graph Construction	Graph/HPC Engineer				
Historical Evolution Analysis	Forecasting Engineer				
Convergence Detection	Forecasting Engineer				
Forecasting Engine	Forecasting Engineer				
HPC Optimization & Benchmarking	Graph/HPC Engineer				
Visualization & Dashboard	Forecasting Engineer				
Documentation & Repository	Lead/Integration				
Poster, Demo & Presentation	Entire Team				