HARD TO CACHE

Research Poster

A VISUALIZED SUMMARY OF OUR RESEARCH

ADDRESSED PROBLEM

Despite advances in AI/HPC, >70% of computational studies fail replication due to undocumented dependencies, environment drift, and fragmented workflows. This erodes scientific trust and impedes progress.

Our research asks:

- **How can we systematically quantify reproducibility risks in research artifacts?**
- **Can an automated framework bridge the gap between data collection, analysis, and public dissemination?**

HardToCache contributes:

- **A novel Scorecard methodology to audit reproducibility across 8 critical dimensions**
- **Integrated automation tools for scraping, scoring, and visualization**

GOALS

- **A Quantifiable Reproducibility Framework**
- **Seamless Automation Pipeline**
- **Accessible Science Dissemination**
- **Open Collaboration Infrastructure**

RESOURCE NEEDS/LIST

Resources used:

- **Google Colab - Python**
- **SGX3 Project Server - HTML/SSH**
- **Github**
- **Canva**
- **pandas, fitz, pathlib, matplotlib**
- **Google Sheets**

METHODOLOGY

1. **Collect**
   - **Automated scraping of research papers**
   - **Extract code/data from publications**
2. **Score**
   - **Apply 6-key Scorecard**
   - **Validate with expert feedback**
3. **Share**
   - **Generate website + poster**
   - **Publish tools/scores on GitHub**

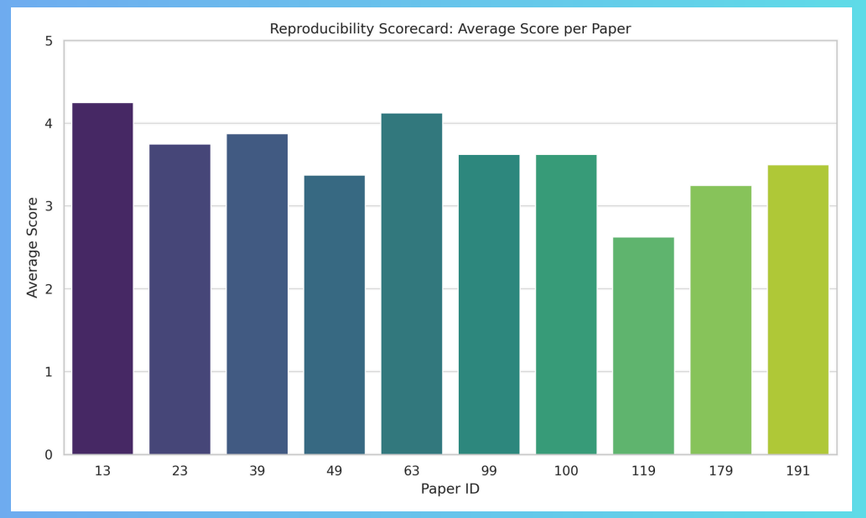
EXAMPLE PAPER SCORING

- **Only a few papers had fully available code and datasets that were able to be reproduced on any given environment. Most lacked full environment or hardware details**
- **This suggests that while some research is highly reproducible, many papers will still miss key details**
- **Each category was scored from 0 (not mentioned/accessible) to 5 (very clearly addressed) using keywords in the paper.**

POSSIBLE EXPANSIONS

- **Scale Across Disciplines**
  - **Biomedical research, Climate modeling, Social science**
- **Real-Time Monitoring**
  - **Develop a browser extension that shows the reproducibility score and live ratings.**
- **Enhanced Automation:**
  - **Auto-update scores when papers are revised**

SCORECARD EXPLANATION



Canva Tip: Double-tap on the sample charts to change the data

AUTHORS

Charli Brooks
Elizabeth City State University
cnbrooks04@gmail.com

Silas Erving
Morehouse College
silaserving815@gmail.com

Chante Ray
Mississippi Valley State University
chanteray3@gmail.com

Seth Austin Mack
Morehouse College
sethamack89@gmail.com

HACKHPC@ ADMI25 HACKATHON

Supported by: SGX3 projectEUREKA! ADMI

https://hackhpc.github.io/admi25