UNIVERSITÄT LIECHTENSTEIN

“WE PROVIDE OUR RESOURCES IN A DEDICATED REPOSITORY”

SURVEYING THE TRANSPARENCY OF HICSS PUBLICATIONS

HICSS2025

Irdin Pekaric and Giovanni Apruzzese

MOTIVATION

- Many **research papers** are **limited** by their **length** and cannot include all **necessary data**, **code**, and **artifacts** for replication.
- **External repositories**, such as **GitHub**, offer a solution by allowing researchers to **share additional resources** that enhance the **transparency** and **reproducibility** of their work.
- The motivation behind this study is to assess the current state of **repository usage in HICSS publications** and **identify trends and gaps** in **transparency practices.**

https://allvectorlogo.com/github-logo/

https://sue.eu/partners/gitlab/

https://seekvectorlogo.com/zenodo-vector-logo-svg//

https://logos-world.net/google-drive-logo/

RESEARCH GOAL

GOAL: This study aims to investigate to what extent HICSS publications utilize external repositories to provide supplementary materials, which are essential for understanding and replicating research outcomes.

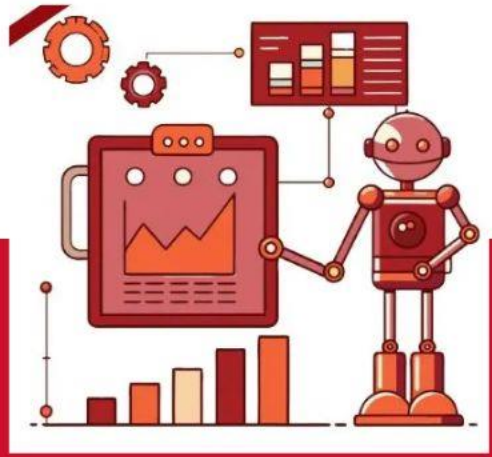
User Studies

Technical Papers

METHODOLOGY

https://miro.medium.com/v2/resize:fit:2000/1*f2-zeAOSNB4RGlqH9emTlQ.jpeg/

1. Data Collection



https://lazymonkey.in/blog/wp-content/uploads/2024/08/Patient-feedback-best-practices-3.webp

2. Preliminary (Automated) Paper Analysis

https://lazymonkey.in/blog/wp-content/uploads/2024/08/Patient-feedback-best-practices-3.webp

3. Manual Review and Validation

DATA COLLECTION

- **Goal: Extract papers** (pdfs) from **Scholar Space**
  Extract papers from **2017-2024**

- **How**: Develop a web scrapper
- Needs to be able to handle **four-layered structure**:
  0. **Community** (HICSS)
  1. **Subcommunity** (e.g. HICSS 57)
  2. **Track** (e.g. Software Technology)
  3. **Minitrack** (e.g. Software Sustainability: Research on Usability, Maintainability, and Reproducibility)
  4. **Publication** (e.g. Understanding Open-Source NLP Artifact Adoption Through Information Systems Success Factors)
- **Issues**: Unconsistent structure

AUTOMATED ANALYSIS

5579 HICSS papers

User study involvement? → user study, questionnaire, survey, interview, etc.

Technical implementation? → empirical, experiment, evaluation, source code, etc.

Link to external repository? → GitHub, GitLab, Zenodo, Figshare, etc.

objective → keywords

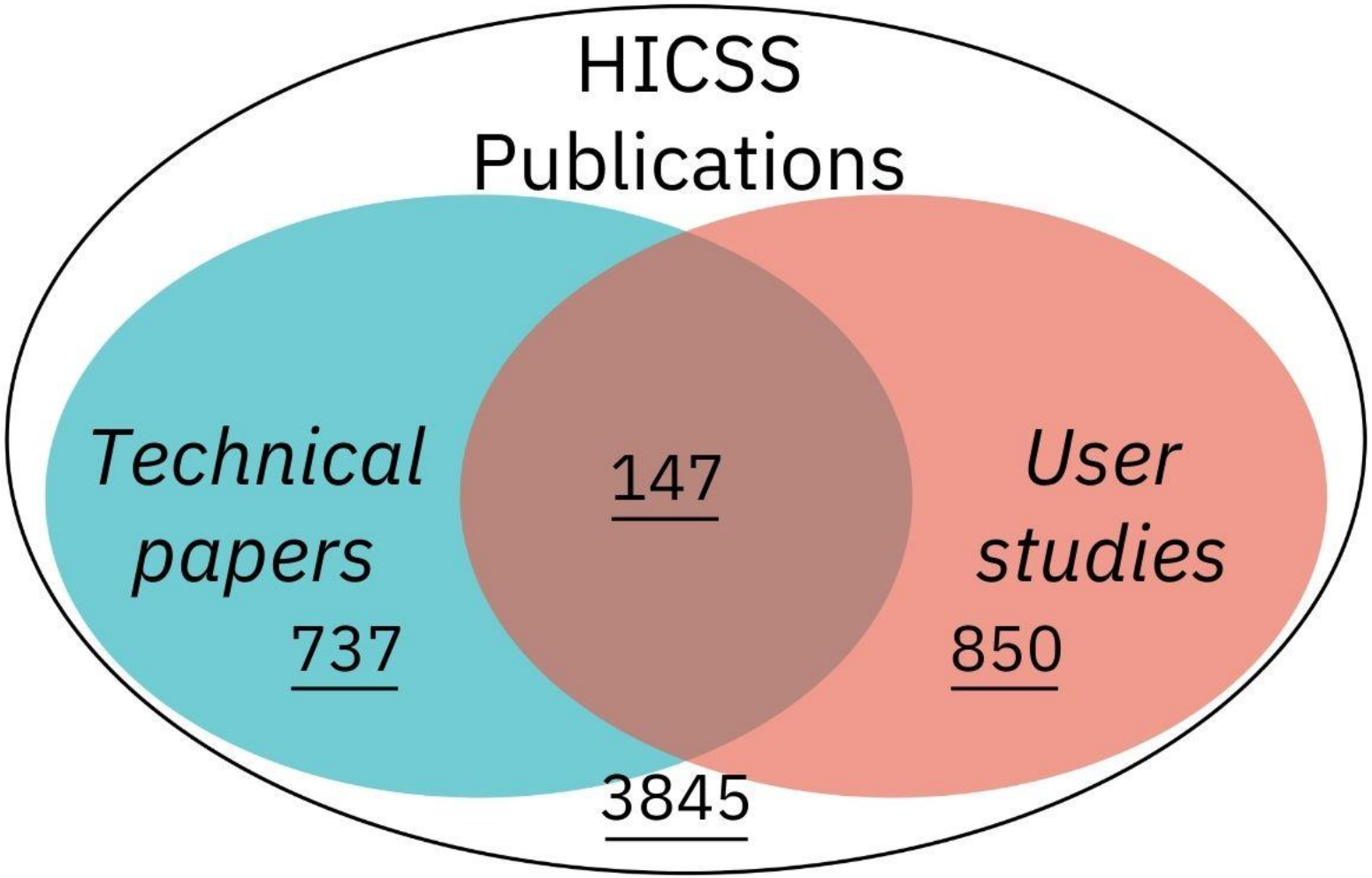
RESULTS

1949 technical papers
551 user studies
131 technical papers with a repository
15 user study papers with a repository

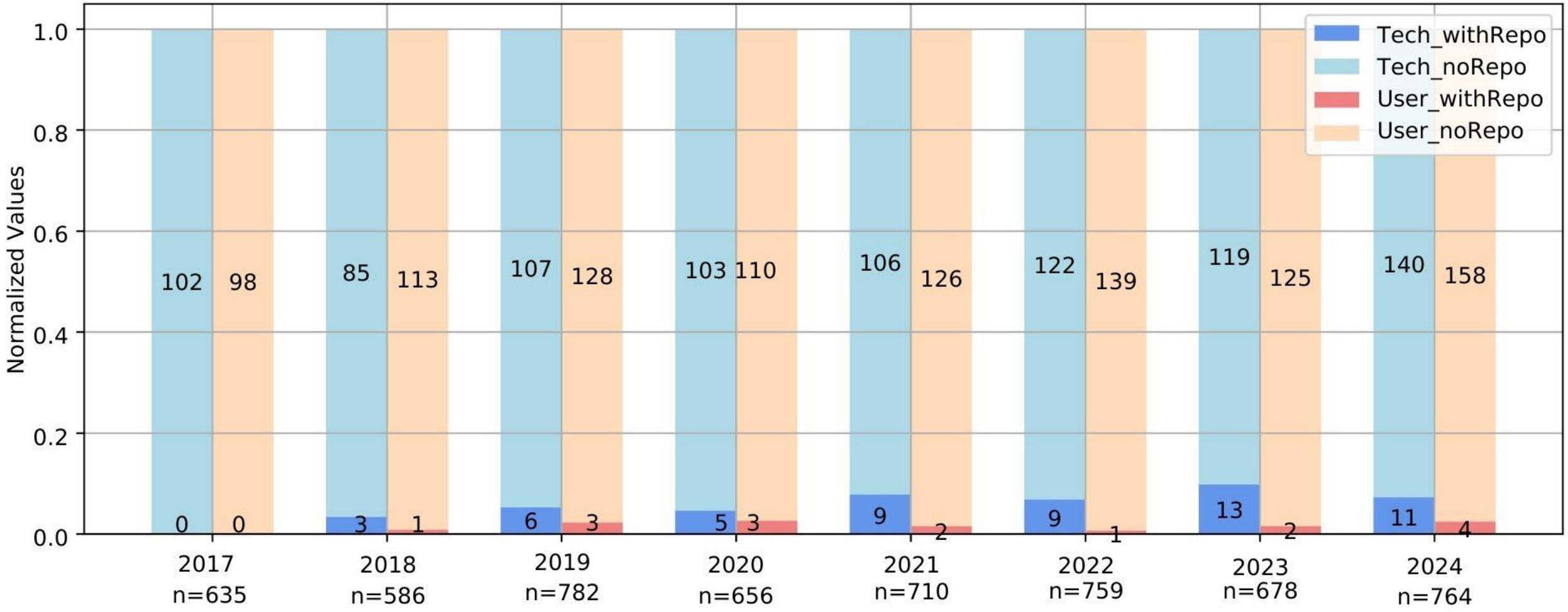
MANUAL ANALYSIS

- **Goal:** Validate automated findings and ensure accurate categorization.
- **How:** Manual review of candidate papers identified by the script**.**
- **Three** main verification steps:
  - Confirm the **correct category assignment**
  - **Check for links** to dedicated **repositories**
  - **Inspect** repository **content** (if available)
- **Method:**
  - Conducted in **pairs**
  - **Frequent discussions** to resolve uncertainties and reach consensus

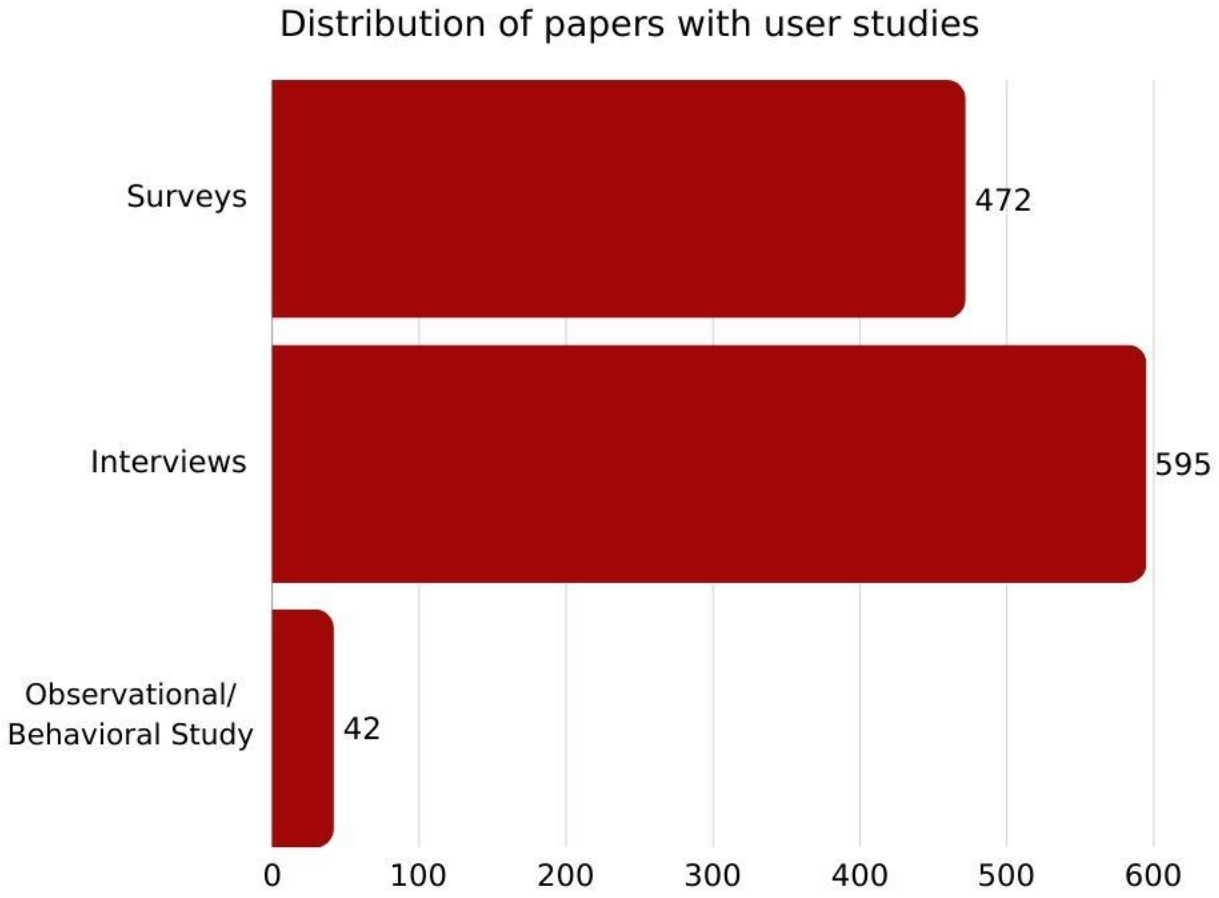
PUBLICATIONS AT HICSS

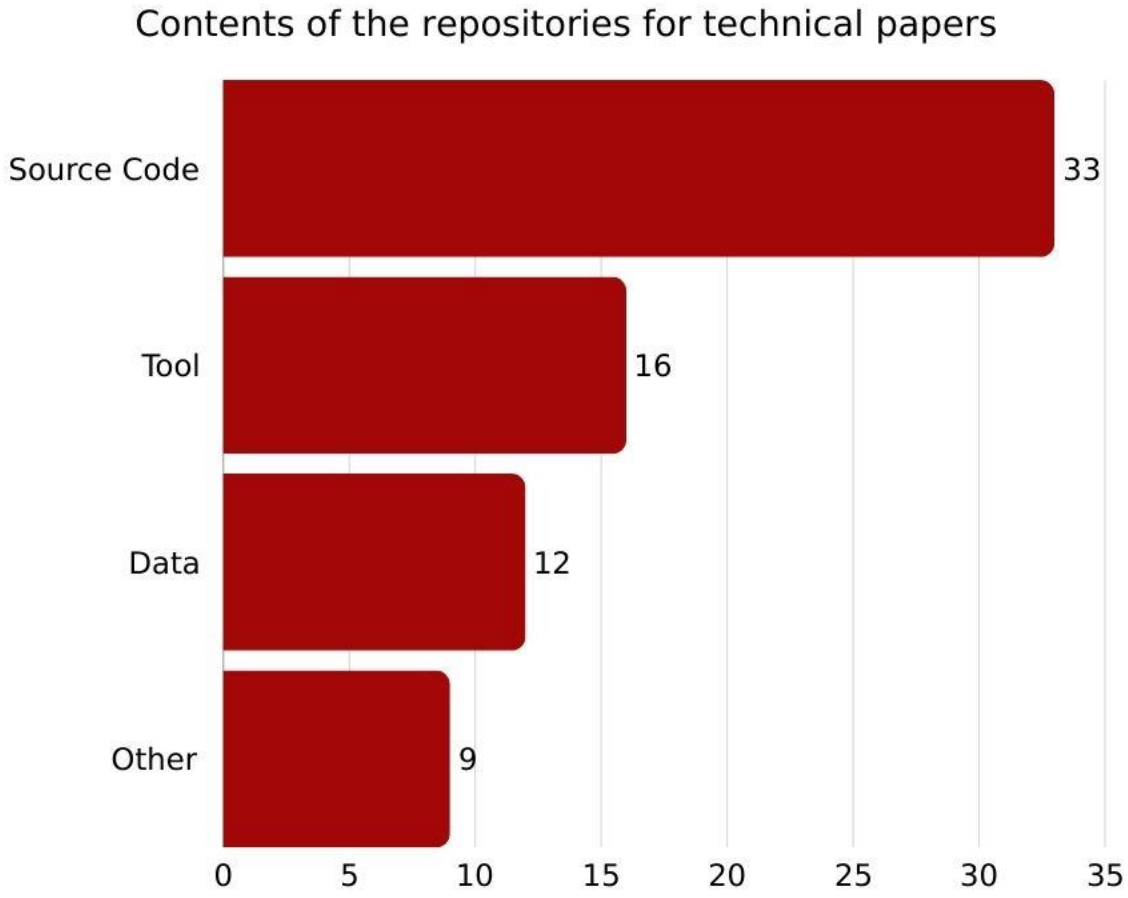


TEMPORAL DISTRIBUTION OF PAPERS



RESULTS (1)





Repositories: S=11, I=7, B=1

RESULTS (2)

Track Name	Technical Total	Technical Repository	User Study Total	User Study Repository
Decision Analytics and Service Science	141	10	103	1
Digital Social Media	64	5	83	1
Electric Energy Systems	55	4	1	0
Digital Government	58	3	90	2
Internet Digital Economy	90	1	130	2
Knowledge Innovation and Entrepreneurial Systems	29	2	69	2
Software Technology	140	16	47	2
Location Intelligence	20	0	4	0
Information Technology, Social Justice, and Marginalized Contexts	9	1	8	0
Collaboration Systems and Technologies	126	6	151	2
Information Technology in Healthcare	85	2	123	2
Software Engineering Education and Training	17	2	14	0
Organizational Systems and Technology	50	4	174	2

Distribution of papers over the tracks

FINDINGS AND IMPLICATIONS

Resource Transparency:

- Only 3% of the 5579 analyzed HICSS papers linked to external repositories

Limited Repository Links:

- A very small number of papers provided access to supplementary materials necessary for reproducibility.
- Many papers lacked essential resources, which impairs the ability of other researchers to validate and build upon the findings.

Need for Improved Resource Sharing:

- The findings underscore the critical need for researchers to make their artifacts publicly available to enhance transparency and facilitate reproducibility in research.

Encouragement for HICSS Organizers:

- HICSS organizers are encouraged to promote the release of supplementary resources and consider giving higher regard to papers that provide such artifacts.
- Implementing specific labels for papers that include external resources could help improve visibility and accessibility.

LIMITATIONS

Manual Analysis Constraints:

- The study relied on a manual review process, which is inherently time-consuming and may introduce human error in categorizing papers and identifying repository links.

Keyword Search Limitations:

- The preliminary filtering script used keyword searches, which may have missed relevant papers that did not include specific terms in their abstracts.

Potential Underestimation:

- The results may underestimate the number of user studies or technical papers due to the limitations of the keyword-based approach and the possibility of overlooked links.

Lack of Ground Truth:

- Without a definitive benchmark for comparison, it is challenging to ascertain the accuracy of the findings, leading to uncertainty in the results.

Variability in Repository Quality:

- The quality and usability of the repositories linked in the papers were not assessed, which may affect the overall effectiveness of the shared resources.

CONCLUSIONS

Limited Resource Sharing:

- The analysis revealed that only 3% of HICSS papers from 2017 to 2024 provided links to external repositories, highlighting a significant gap in resource transparency.

Impact on Reproducibility:

- The scarcity of supplementary materials restricts the ability of researchers to validate findings and reproduce studies, which is essential for scientific integrity.

Need for Cultural Shift:

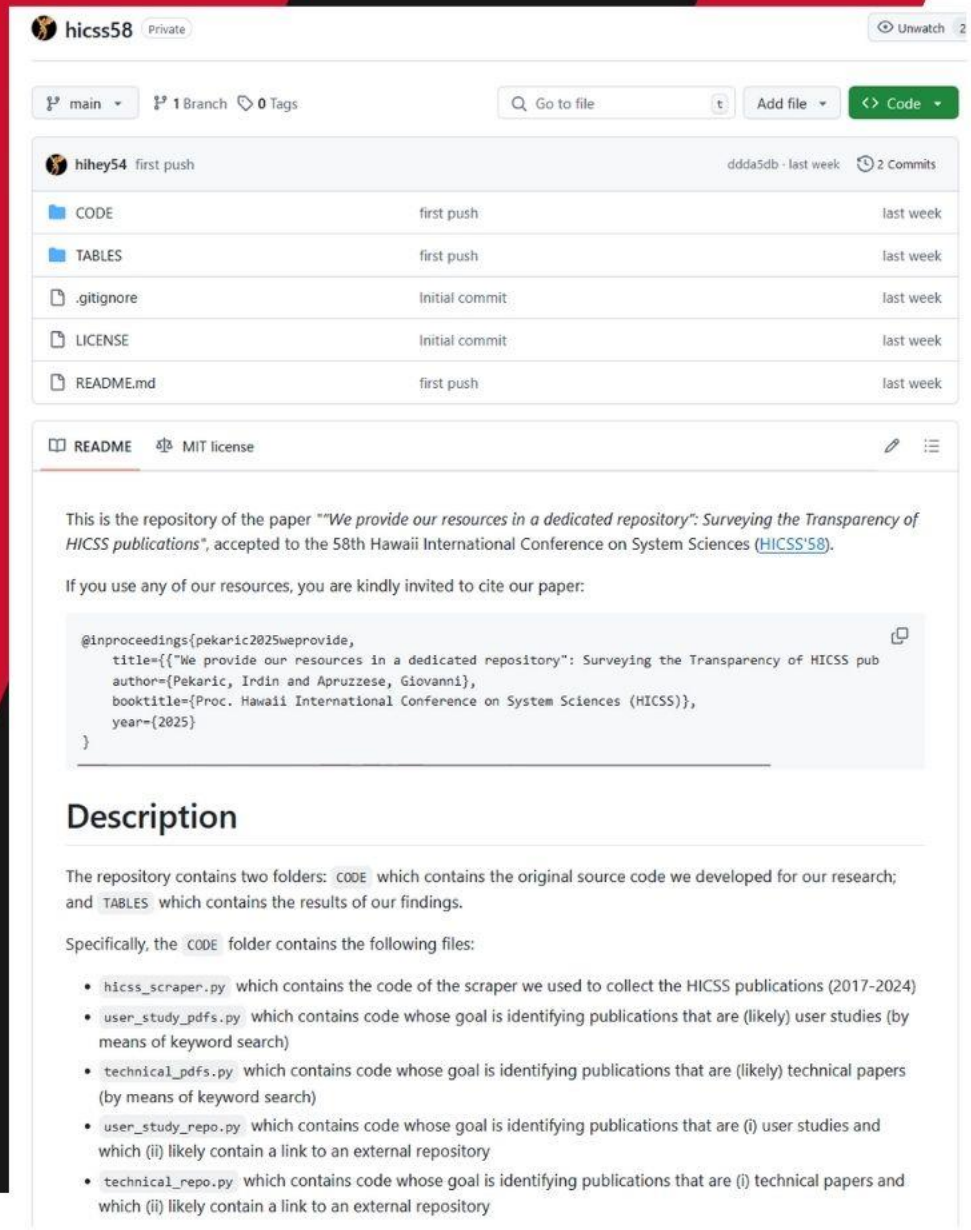
- There is a pressing need for a cultural shift within the research community towards greater openness and sharing of research artifacts to enhance collaboration and reproducibility.

Encouragement for Future Research:

- The findings serve as a call to action for researchers and conference organizers to prioritize and promote the sharing of resources, thereby improving the overall quality and reliability of research outputs.

Inspiration for Methodological Improvements:

- The study highlights the potential for developing more efficient methods, such as automated tools, to assess and encourage resource sharing in future research.



hicss58 Private

Unwatch 2

main · 1 Branch · 0 Tags · Go to file t · Add file · Code

hihey54 first push — ddda5db · last week · 2 Commits

CODE	first push	last week
TABLES	first push	last week
.gitignore	Initial commit	last week
LICENSE	Initial commit	last week
README.md	first push	last week

README · MIT license

This is the repository of the paper *"'We provide our resources in a dedicated repository": Surveying the Transparency of HICSS publications"*, accepted to the 58th Hawaii International Conference on System Sciences (HICSS'58).

If you use any of our resources, you are kindly invited to cite our paper:

```
@inproceedings{pekaric2025weprovide,
    title={{"We provide our resources in a dedicated repository": Surveying the Transparency of HICSS pub
    author={Pekaric, Irdin and Apruzzese, Giovanni},
    booktitle={Proc. Hawaii International Conference on System Sciences (HICSS)},
    year={2025}
}
```

Description

The repository contains two folders: `CODE` which contains the original source code we developed for our research; and `TABLES` which contains the results of our findings.

Specifically, the `CODE` folder contains the following files:

- `hicss_scraper.py` which contains the code of the scraper we used to collect the HICSS publications (2017-2024)
- `user_study_pdfs.py` which contains code whose goal is identifying publications that are (likely) user studies (by means of keyword search)
- `technical_pdfs.py` which contains code whose goal is identifying publications that are (likely) technical papers (by means of keyword search)
- `user_study_repo.py` which contains code whose goal is identifying publications that are (i) user studies and which (ii) likely contain a link to an external repository
- `technical_repo.py` which contains code whose goal is identifying publications that are (i) technical papers and which (ii) likely contain a link to an external repository

REPOSITORY

https://github.com/hihey54/hicss58

THANK YOU

FOR YOUR ATTENTION