



Laying foundations to quantify the "Effort of Reproducibility"

Akhil Pandey Akella

Northern Illinois University

Dekalb, IL, USA

aakella@niu.edu

David Koop

Northern Illinois University

Dekalb, IL, USA

dakoop@niu.edu

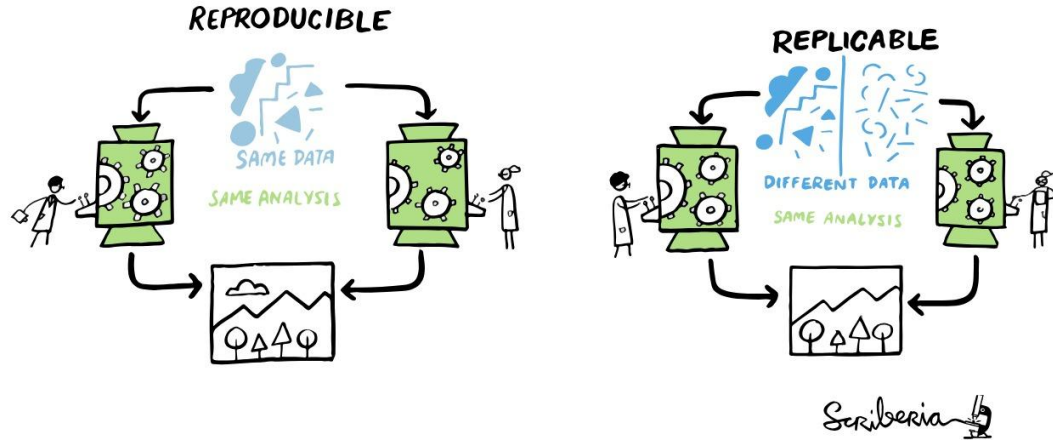
Hamed Alhoori

Northern Illinois University

Dekalb, IL, USA

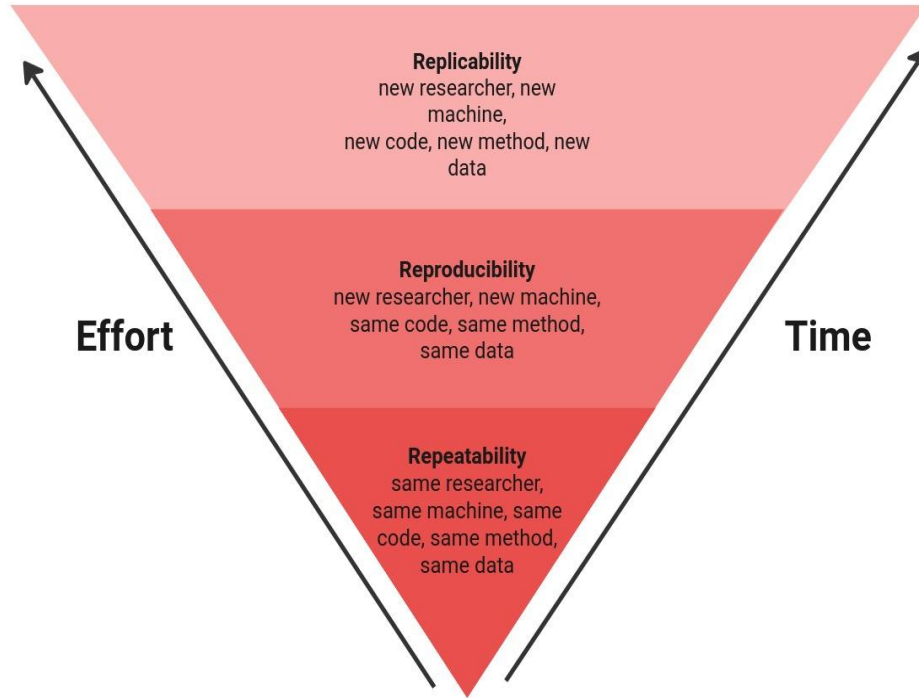
alhoori@niu.edu

What is Reproducibility ?



miro

Effort of Reproducibility



miro

Quantifying the 'Effort of Reproducibility'

OBTAIN RAW DATA FROM
ML REPRODUCIBILITY
CHALLENGE



STEP 1

EXTRACT META
INFORMATION FROM
REPRODUCIBILITY REPORT



STEP 2

DISCOVER FACTORS THAT
CONTRIBUTE TO
EASINESS & DIFFICULTY



STEP 3

Data Collection

OBTAIN RAW DATA
FROM ML
REPRODUCIBILITY
CHALLENGE



YEARS: 2020, 2021

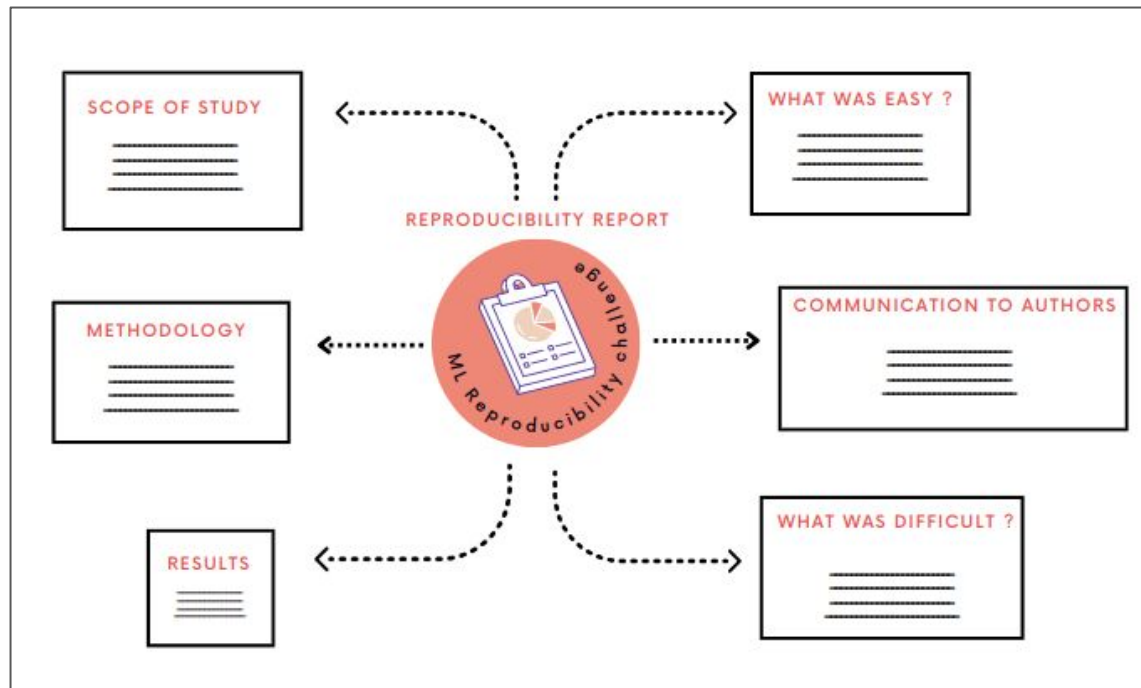
AUTHOR'S NAME, TITLE, DOI, CITATIONS
META INFORMATION ABOUT
ORIGINAL WORK

AUTHOR'S NAME, TITLE, DOI, VOL, ISS
META INFORMATION ABOUT
REPRODUCIBILITY REPORT

OPENREVIEW, PDF'S, CODEBASE
DIGITAL ARTIFACTS

SAMPLE SIZE: 87

Information Extraction

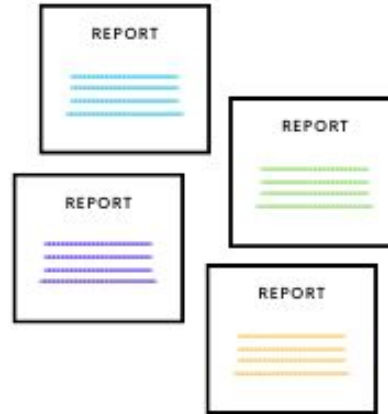


Inductive Analysis

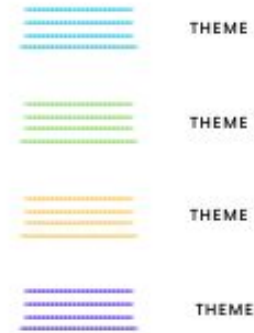
SYSTEMATICALLY READING REPORTS



IDENTIFY DOMINANT THEMES



DEVELOP ENCODING STRATEGY



Scope of reproducibility

[RE] ORIGINAL WORK

RESEARCHERS
REPRODUCING
JUST THE
ORIGINAL
WORK

[RE] ORIGINAL WORK +
EXTENSION

RESEARCHERS
REPRODUCING THE
ORIGINAL WORK AND
ADDING A MINOR
EXTENSION

[RE] MULTIPLE WORKS TO
PRODUCE NEW WORK

RESEARCHERS
REPRODUCING
MULTIPLE ORIGINAL
WORKS TO CREATE
NEW WORK



ORIGINAL ARTICLE

SEMI-SUPERVISED CLASSIFICATION WITH
GRAPH CONVOLUTIONAL NETWORKS

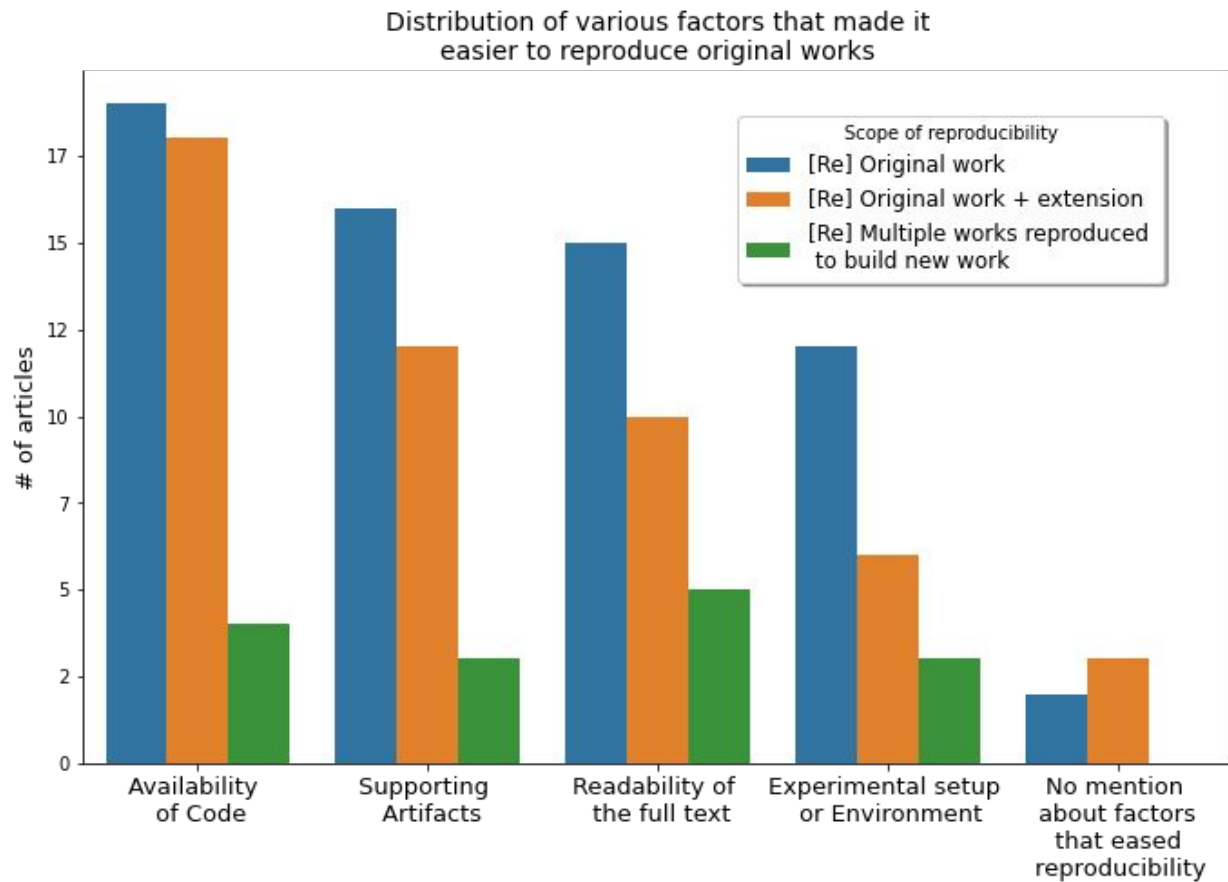
[RE] SEMI-SUPERVISED CLASSIFICATION WITH
GRAPH CONVOLUTIONAL NETWORKS

[RE] SEMI-SUPERVISED CLASSIFICATION WITH
GRAPH CONVOLUTIONAL NETWORKS

+ HYPERPARAMETER
TUNING

[RE] SEMI-SUPERVISED... + [RE] OTHER WORK

Reasons that made it easy



Reasons that made it easy

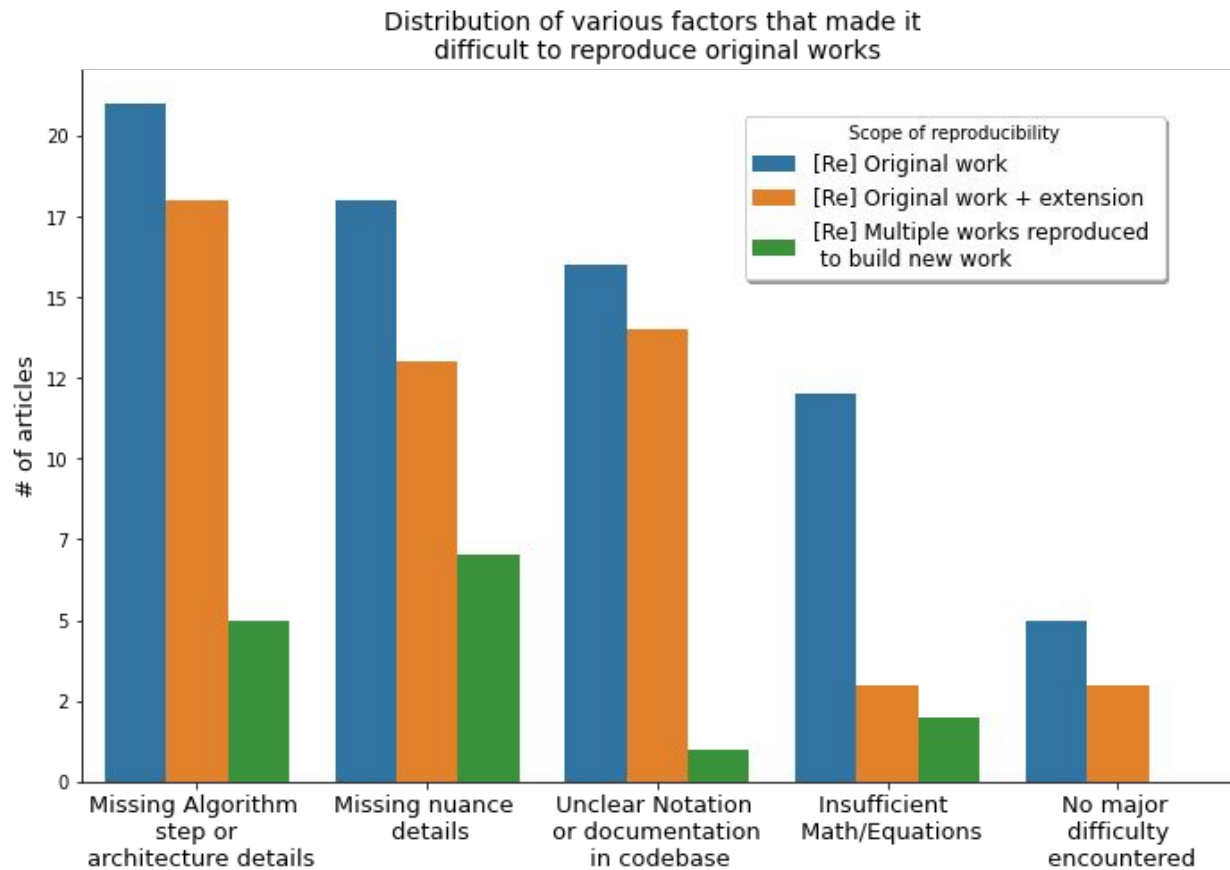
WHAT WAS EASY

IMPLEMENTING MOST OF THE CODE WAS STRAIGHTFORWARD AS AUTHORS OF BOTH PAPERS **PROVIDE SOURCE CODE**. GITHUB ISSUES WERE ANOTHER SOURCE OF RETRIEVING INFORMATION, CLARIFYING PARTS OF THE PAPERS WHEN NEEDED. ADDITIONALLY, BOTH OF THE ORIGINAL PAPERS ARE QUITE **COMPLETE, WELL-WRITTEN** MAKING IT EASY TO FOLLOW.

REASONS IDENTIFIED BY INDUCTIVE ANALYSIS

AVAILABILITY OF CODE
READABILITY OF THE FULL TEXT

Reasons that made it difficult



Reasons that made it difficult

WHAT WAS DIFFICULT

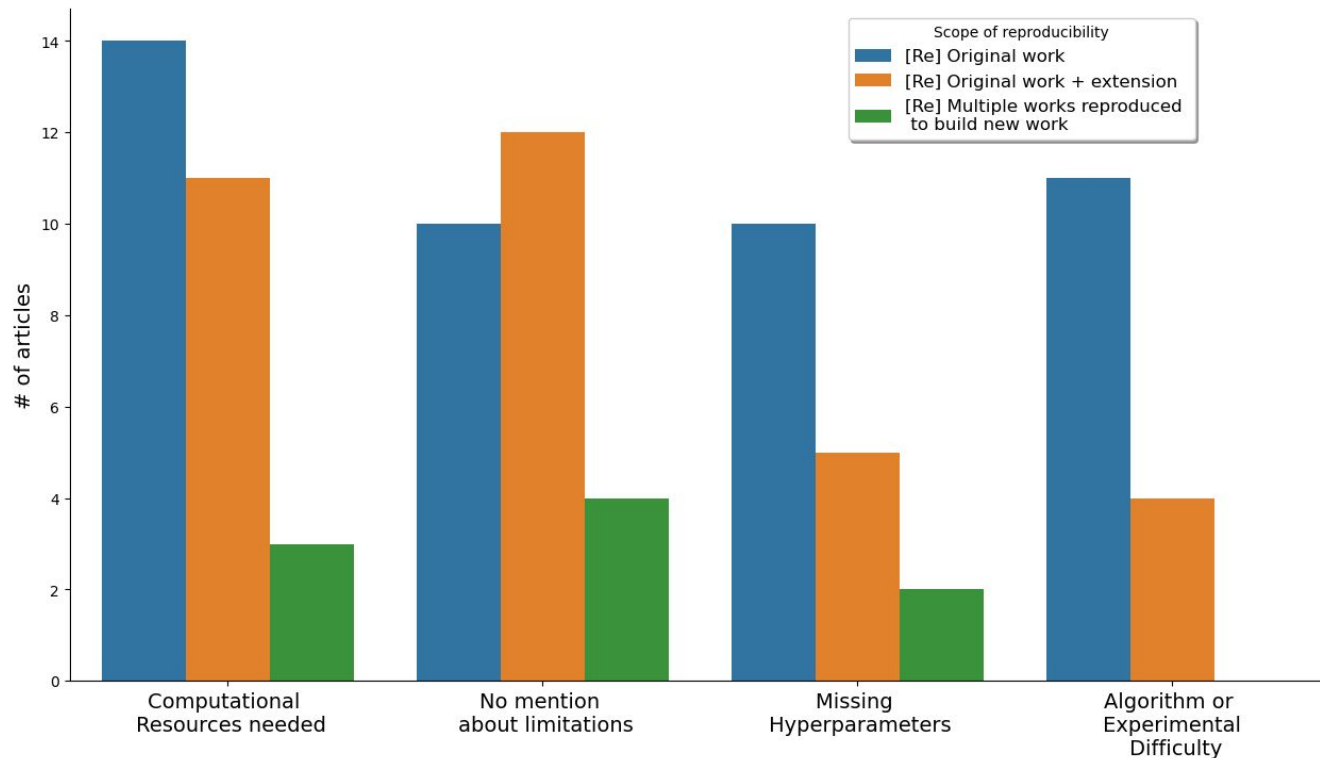
THE ORIGINAL CODE DID NOT CONTAIN ANY DOCUMENTATION, WHICH MADE IT DIFFICULT TO NAVIGATE. NO CODE FOR CALCULATING THE METRICS WAS PROVIDED AND THIS HAD TO BE IMPLEMENTED FROM SCRATCH. DURING THE TRAINING OF THE MODELS, MEMORY ALLOCATION ISSUES OCCURRED. TRAINING AND EVALUATING ON A LARGE DATASET TOOK A CONSIDERABLE AMOUNT OF TIME.

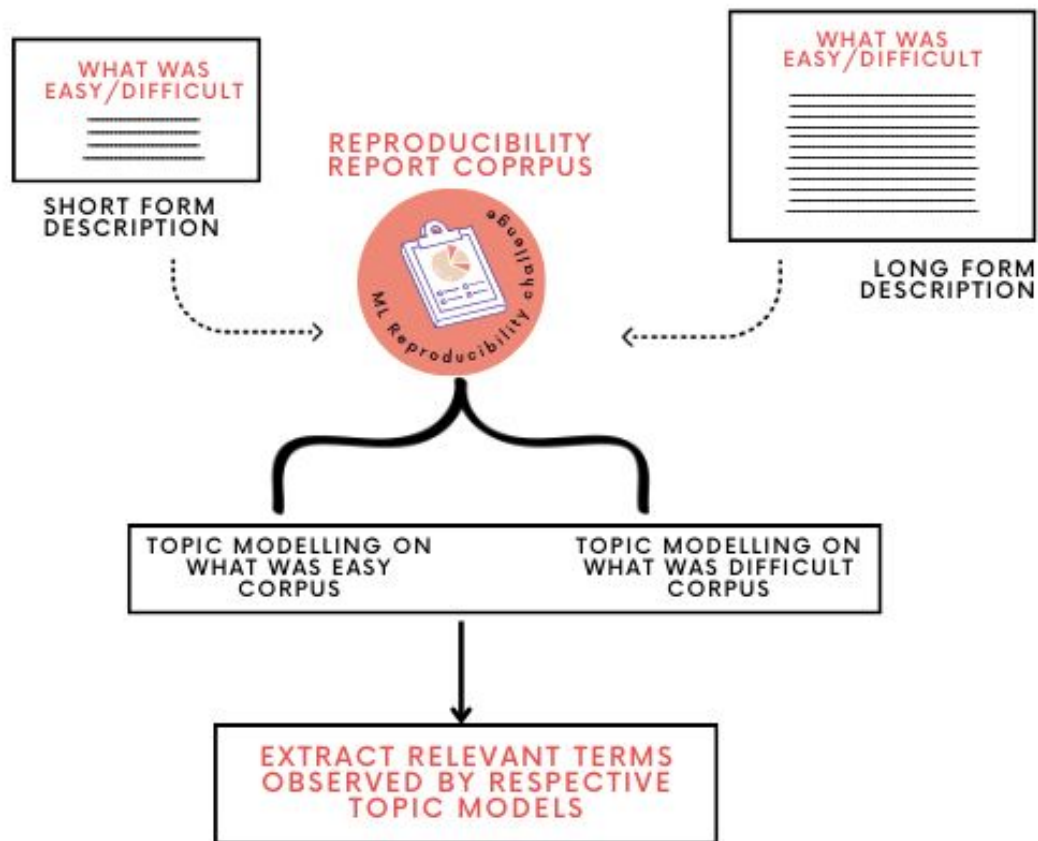
REASONS IDENTIFIED BY INDUCTIVE ANALYSIS

UNCLEAR NOTATION OR
DOCUMENTATION IN
CODEBASE

MISSING ALGORITHM STEP OR
ARCHITECTURE DETAILS

Limitations





Quantitative Analysis

Relevant terms

- 'What was easy' corpus

Topic	Most relevant terms
1	<i>describe, straightforward, understand, documented</i>
2	<i>codebase, repository, source, instructions</i>
3	<i>datasets, training, scripts, experiments</i>
4	<i>results, ideas, evaluation, architecture</i>
5	<i>correspondence, addressed, peer-review, copyright</i>

Relevant terms

- 'What was difficult' corpus

Topic	Most relevant terms
1	<i>dataset, algorithm, implementation, method</i>
2	<i>training, loss, accuracy, learning</i>
3	<i>models, network, training, time</i>
4	<i>difficult, challenges, evaluation, claim</i>
5	<i>methods, features, performance, parameters</i>

Conclusion

- We have laid the groundwork to analyze and identify factors that encapsulate the effort required to reproduce scientific articles.
- We found “**Availability of Code**” as one of the most important factors that made it easy to reproduce the original work.
- We found “**missing algorithm steps or architectural details**” as one of the most important factors that made it difficult while reproducing the original work.
- Studies under the scope “*[Re] Multiple works reproduced to build new work*” never mention “**Algorithm or Experimental difficulty**” as a subjective factor for limitation.
- Our Topic model found “**Communication with the original authors**” as an additional factor that can ease the effort of reproducibility.
- Quantifying the effort required to reproduce a study provides a valuable framework for future research in this area.

Artifacts



<https://reproducibilityproject.github.io/effortly>



Acknowledgement



National Science Foundation

SIGIR
Special Interest Group
on Information Retrieval

ATA Lab



Northern Illinois
University

End | fin