23https://par.nsf.gov/servlets/purl/10443571

Abstract
Learning analytics has increased in popularity. Open science - which describes a set of practices to make research more open, transparent, and reproducible – has exploded in recent years, resulting in more open data, code and materials for researchers to use. Without prop knowledge, many researchers do not make their datasets, code and materials openly available. Those that are available are difficult, if not impossible to reproduce.
Purpose of current study is to take a closer look at field by examining previous papers withing proceedings of International Conference on Learning Analytics and Knowledge (LAK 23). Aswell as document the rate of open science adoption (e.g preregistration, open data), also how well available data and code could be reproduced.
Examined 133 research papers, 15minutes for each paper to identify open science practices and attempt to reproduce the results.
Less than half of the research adopted standard open sciecne principles, with approximately 5% fully meeting some of the defined principles.
Unable to successfully reproduce any of the papers in given time period.
We conclude by providing recommendations
on how to improve the reproducibility of our research as a field
moving forward.

1 Introduction
Research using learning analytics and open science practices has
both increased in popularity over the past decade. Learning analytics aimed at improving our understand of learning and education and open science focused on improving scientific practices. Learning analytics has developed the field of education technology through applications and integration in Learning Management
Systems (LMS), collecting student data to produce better models
in a circular development cycle [3, 10].
Open science, which describes practices to make research more open, transparent, and reproducible, has been receiving widespread attention and increased
adoption in felds such as Psychology and Statistics [28, 35].
Learnign analytics and open science tend to run parallel, occasionally overlapping. Conferences such as the International Conference on Learning Analytics and Knowledge do not actively promote open science within the field, which is disappointing, as open science practices are designed to improve the rigor and reproducibility of research. But even when researchers are aware of open science, they may not have enough knowledge or experience to properly make their work meet open science principles.
… limited time to only 15 minutes per paper, logistical contraints and the idea that openness and reproducibility should, when done well, require little time to identify and reproduce.
By pointing out the
shortcomings in open science and reproducibility of our work and others, the authors hope to stimulate the education technology community to develop and share their practices.
Specifcally, this work aims to accomplish the following tasks:
(1) Document and analyze which papers within the proceedings of the International Conference on Learning Analytics and Knowledge (LAK) meet the open science principles and subfields defined by this work.
(2) Determine and document whether the results within the papers are reproducible within a 15-minute timeframe

2 Background

2.1 Open Science

Open science is an umbreall term used to describe that the methodologies, datasets, analysis, and results of any piece of research are openly accessbile [15,35].

While each subcategory has existed for decades, open science itself became prevalent near the beginning of the 2010s with an explosion in popularity near the middle of the decade[28]. Early 2010s many researchers were finding issues when peer reviewing others work: replication failures, unclear practices, overused materials, etc. The open science movement was starting to gain traction as a way to ameliorate these issues by combining mitigations for each part of the problem into a collection of practices, standards, and recommendations.

Largescale studies in the mid-2010's cast doubt on the reproducibility and replicability of research, including replication efforts in psychology (Open Science Collaboration, 2015)[4] and other felds (Baker,2016)[2]. As a result many began to actively use and improve science standards. This work will make use of some of the principles of open science in addition to its reproducibility catalyst to evaluate the selected papers.

Open science practices are designed to improve the rigor and reproducibility. Beyond serving a larger goal of cumulative science, where access and openness are no longer a barrier, and materials and data are widely shared, open science aims to makes our work better.

2.2 Education Technology and Learning Analytics

Education Technology is a field of study investigating the development and evaluation of a learner, their process, and their materials[6, 9]. Progress within the field is cyclical: Researchers develop new technology or processes for teachers and learners to use, teachers and learners provide data and feedback, and researchers use the data to better improve their technology or process to begin the cycle over again.

With a broad defnition, education technology
covers a wide variety of topics, most of which create or integrate with Learning Management Systems (LMS)[5]. One such subfield that focuses more on the data collection, modeing and analysis side is learning analytics.

Learning Analytics is the process by which data collected from learners is investigated to improve learning [3,10,34]. With the creation of the International Conference on Learning Analytics and Knowledge (LAK) in 2011, a common definition was adapted along with a platform for this subfield of research to become independent [29].

As the subfeld focuses more on topics that are in line with the open science principles, papers relating to learning analytics submitted to the LAK conference will be used as the dataset for this work.

3 Methodology: Open Science peer review

Evaluated every research acrticle, short paper, and poster recorded in LAK proceedings of the last two years: 12th and 11th LAK22/21 [31,32]. Only these two because open science has slowly expanded since the early 2010s. Impractical to assume that papers submitted to earlier LAK conferences would meet all defined open science principles, when they were not originally present. Second the older a paper the greater likelihood that the work would face challenges to reproducibility: old packages and libraries, lost data etc.

3.1 Open Methodology

Open Methodology is an uncommonly used 'umbrella' term under open science, which represents the details of the methods and evaluation used by the authors, including but not limited to the setup, logic fow, aggregations of results, etc. [15]. The methodologies

provide additional information that goes beyond the original paper, such that reviewing researchers are able to replicate or reproduce the study themselves.
Papers submitted to the LAK conference tend to follow a similar structure and almost always contain a section dedicated to the methodology–data collection and analysis performed–within the paper.
Documenation will mark a paper as open methodology if available on ACMDL. Naive approach because papers typically do not contain all information needed to reproduce a given result [13,16].
Each paper in ACMDL can be marked with additional metadata tags. Focusing on “Open Access” and “Public Access”. ‘Open Access’ is used to defne a paper that is made freely available online to anyone [1].‘Public Access’ meanwhile is a requirement assigned to a paper that must be made freely available within one year of its publication (e.g. publications from the United States National Institutes of Health (NIH).
In both cases the paper is accessible by anyone on ACMDL.

3.2 Open Data

Open Data is used to defne a single or multiple datasets which can be used and made openly accessible without restriction, or with restrictions that do not prevent its use by the recipient [17, 18]. Mark a paper as containing open data if the paper contains a link, or a link to another paper with a link, to the dataset. If a paper does not use or collect data, then this field will be marked as non-applicable.
Not all data can be openly released. May contain personally identifiable information, not anonymized or openly available. In these cases additional option as added, data can be obtained by request from the author. But this must be explicitly stated within the paper; otherwise marked as containting no open data.
In addition to the open data field, a seperate field was added to determine whether there was data documentation available for the dataset. Data documentation contains a one-to-one mapping of the field name in the dataset to a description of what the field is and what values it may contain (e.g a ‘data dictionary’, table, specification, etc.). The documentation will mark a paper as having data documentation if the paper or a separate document contains all felds in the dataset. If a paper does not use or collect data, then field marked as non-applicable. May be marked as partial, indicating that there exists some documentation on at least one field.

3.3 Open Materials

Open Materials term used to define technologies that can be used and made openly accessible. Within the literature open materials tend to be synonymous with Open Source as it is defined within the context of software development [11, 24].
Open materials defne a slightly broader scope: not only containing open source software, but also technologies that are free-to-use (e.g. Google Sheets).
This documentation will mark a paper as containing open materials if the paper contains a link, or a link to another paper with a link, all of the materials and source used. If a paper does not use a material (e.g. theoretical or argumentative papers), then this field will be marked as non-applicable.
In addition three subfields are added. First, a materials documentation field, which determines wehter the materials and source provided contains documentation for their usage.
Second, a README field is added, to determine wether or not a README is present within the source. READMEs usually contain an introduction to the source in addition to information on the documentation or setup necessary to use the provided source [14].
Finally, a license field is added, which determines whether the source has a permissible software license, such that the source can be used openly by any user [8, 24, 27]. These

fields will be marked as non-applicable only if the open materials field is marked as non-applicable.
The open materials and materials documentation fields can be marked as partial, indicating that at least one of the materials or materials documentation provided is openly accessible.

3.4 Preregistration

Preregistration is a term used to define the steps conducted for the paper without knowledge of the research outcomes [20, 21, 33]. These are typically made prior to the initial starting point of an experiment, such that any prediction or result obtained is unbiased by any observations made by the researcher. Additionally, if the experiment needs to be altered, an update or additional preregistration can be made to preserve the initial methodologies. This registration is made openly accessible and is held on some independent registry (e.g., the Open Science Foundation). Standard preregistration practices involve citing the preregistration within the body of the paper and identifying which analyses/aspects of the research were preregistered. This documentation will mark a paper as containing a preregistration if the paper contains a link to the preregistration or a project containing the preregistration. If a paper does not need to use a preregistration (e.g. theoretical or argumentative papers), then this field will be marked as non-applicable.

4 Methodology: Reproducibility

Reproducibility is a term used to defne performing the exact same experiment with the exact same method to see if the reviewer can achieve the exact results reported in the paper[19, 22, 23]. However it can have different levels of completeness. This work will define reproducibility as the ability to take the dataset collected and the analysis method used and return the exact same results and figures given in the paper. It is impractical to collect the data from scratch or write a similar implementation based on the pseudocode in the paper, so if the dataset or the source is not provided, then the paper will be marked as non-reproducible. The reproducibility field will only be marked as non-applicable if either the dataset or materials fields is also marked as non-applicable.
While the requirement could contain additional fields, such as whether a README is present for necessary setup of if a licens is provided to legally use the source, this work assusmes that the minimum required to reproduce a paper is the dataset and the source which consumes it. Per reproducibility test only 15 minutes, arbitarily chosen as a reasonable length to run a provided command or script as this work assumes that, at most, only a few setups, preprocessings, and analyses need to be run.
The 15-minute timeframe only applies to a sequence of non-author actions to perform. This means that if the author has laid out the necessary actions to perform, then even if the 15-minute timeframe is exceeded, if the exact results are produced, then it will be marked as reproducible. Additionally, the 15-minute timeframe will only be exceeded once the last running action of a sequence has fnished executing.
Reproducibility of a work should not be tied to any specific platform.

As there are particular nuances and problems that are specifc to a piece of software, this work will also defne some common problems in popular software and the steps taken towards providing a reproducibility test.

4.1 Python

Python is a programming language used commonly by researchers for integration to read data, conduct analyses, and generate results in a consumable format. The language itself is under an open source license and contains thousands of libraries, known as packages, tailored to most needs. Although a highly useful tool, there are a number of reproducibility

issues if no steps are taken to freeze the author's current environment.

[Python version, package issues here]

4.2 R
[reproducibility issues for R]

4.3 JavaScript: Node.js
[Port issues]

5 Results

Figure 1: A breakdown of the open methodologies of papers submitted to the 11th and 12th conference. Shows the total (top) and breakdown per conference (bottom)..

133 papers within the LAK conference proceedings, 71 submitted to the 11th LAK conference and 62 to the 12th, which are classified as "Research Article", "Short Paper", and "Poster".

Figure 2: A breakdown of the open methodologies of papers submitted to the 11th and 12th conference. Each paper is broken down into its type (top) and whether it was misclassifed by ACMDL (bottom).

In Figure 2, out of the submitted papers, 9 are publicly accessible, 38 are openly accessible, and the remaining 86 are available but not open. None were classifed incorrectly by ACMDL.

Figure 3: A breakdown of the open data of papers submitted to the 11th and 12th conference by open methodology. Breaks down the requirements into having the data publicly available (top) and the documentation for the dataset (bottom).

In Figure 3, we can see that only 5% of all papers make their raw dataset available with 2% explicitly mentioning that the data can be requested. For documentation, 2% of all papers fully document their work, while 38% mention some form of dataset documentation at least once in their paper. The 2% of works that fully document their dataset are using publicly available data which already has the available documentation. Open and Public Access papers adhere more to open data principles compared to the other papers, with Public Access papers adhering slightly better.

Figure 4: A breakdown of the open materials of papers submitted to the 11th and 12th conference by open methodology. Breaks down the requirements into having the materials publicly available (top) and the documentation for the materials (bottom).

For open materials in Figure 4, we can see only 5% of papers provide the full materials needed to reproduce the paper, while 29% mention using at least one open material. In combination with the data felds, only 2% of the papers have the full materials available along with the raw dataset needed to produce a reproduction. No papers fully documented their materials. The materials documentation either could not be located within the 15-minute timeframe or was not available for every public method in the provided source.

Figure 5: A breakdown of the open materials subfelds and preregistration of papers submitted to the 11th and 12th conference by open methodology. Breaks down the requirements into having a README (left), a permissible license (middle), and a preregistration (right).

Looking at the open materials subfelds and preregistration in Figure 5, 7% of papers contained a README with 3% of those containing a permissible license to use their source. Only one paper provided a preregistration, though its materials could not be accessed. A paper is only marked as non-applicable for these felds if the materials felds are also non-applicable, so the graphs are confated to make a point. If only looking at papers with open materials, 75% of the papers had a README while 45% contained a permissible license. Open access papers provided more papers with READMEs and permissible licenses, followed by Available papers, and fnally none by Public Access papers. This correctly correlates with the breakdown of the open materials feld. The only preregistration is provided by an Available paper.

Figure 6: A breakdown of the Reproducibility Test of papers submitted to the 11th and 12th conference.

As shown in Figure 6, not a single paper was reproducible within a 15-minute timeframe. The main cause for nonreproducibility was du e to a lack of libraries provided or that the libraries themselves did not have a version number attached to them. The secondary cause was that some of the sources used some form of randomness but did not make the randomness deterministic to accurately reproduce (e.g. fxed seed). One paper additionally provided a Bash script, which would only work on Linux and older versions (newer with some work) of MacOS17.
Although none of the work was reproducible in a 15-minute timeframe, if given a longer time period, we estimated that the 2% of papers that contain both the raw dataset and source were likely to be reproducible with enough configuration.

7 Future Work

Another direction for future work will be to improve upon our reproduction setup. This will consist of documenting the required materials and versions used to test the software in addition to providing simple setups to run the reproduction. Each of these steps will then be sourced to the required repository or hosted ourselves if not possible. In addition, a DOI will be assigned to any dynamic resources to provide a permanent identifer in case of reference degradation.

8 Conclusion

Although about 30% of the papers meet some of the very broad partial defnitions for the principles of open science, approximately 5% of the papers meet most of the full defnitions.Most of the provided subcategories are simply neglected, which could be due to negligence, but most likely is due to inexperience and unclear resources. An average researcher cannot be expected to know anything about the open science principles, much less understand the subcategories required within them. Additionally, most reproducibility metrics tend to occur within networks of people who are familiar with each other, so as long as the researcher understands how to reproduce their own work, they will almost never run into this issue.
In general, a paper does not need to meet all the open science requirements or the reproducibility metric, depending on circumstances. For example, datasets with personally identifable information may not be released openly.