

Peer Reviewing data at ACL

Proposal

Iryna Gurevych, Ilia Kuznetsov
UKP Lab, Technical University of Darmstadt

Preliminary Remarks

- Peer review is a promising area for NLP
 - As application
 - As text type and domain
- However, reviewing data is sensitive and there are concerns:
 - Privacy, anonymity, copyright, dissemination of unpublished results...
- We want to
 - Start a constructive discussion
 - Determine core issues
 - Discuss solutions for peer reviewing data collection at *ACL
- Based on the full-text Proposal

Executive Summary

- We want to start a constructive discussion on peer reviewing data in the ACL community
- We propose a numerical repository to enable meta-studies of peer reviewing at ACL, at no privacy risk
- We collect the main issues and risks associated with peer reviewing texts and propose solutions to address the challenges related to Privacy, Anonymity, Consent, Copyright and Workload
- We report on a pilot **peer reviewing data collection campaign** at COLING-2020

Outline

- Motivation
- Numerical data: *ACL meta-science repository
- Textual data: Challenges and Solutions
 - Anonymity
"I want to stay anonymous to the authors and reviewers"
 - Privacy and Security
"I don't want negative feedback to be published, and I worry about my research ideas"
 - Copyright
"Who owns the data, how can this data be used and who gets the credit"
 - Consent
"We don't want to trick the reviewers and authors or get them into trouble"
 - Management overhead
"PCs are very busy"
- Pilot study at COLING-2020

Motivation

Why peer review

- Publication overload in research. Peer reviewing is a bottleneck.
 - Time-consuming and expensive
 - Less experienced researchers involved
 - Bias and miscalibration
 - Full-AI peer review is neither feasible nor desirable
 - *AI-assisted* peer review can make the process faster and robust to bias
-
- Peer reviews are sensitive. The few existing datasets lack clear terms of use and copyright
 - the reviewers are not protected; the content is not protected
 - building upon the data is troublesome
 - We need better workflows based on **explicit, informed** consent, clear licensing and secure data storage strategies

Why *ACL

- Benefits of the home community
 - Control over reviewing setups
 - Access to data for own research
 - Short way to applications, no domain shift
 - Easy to annotate compared to out-of-domain, e.g. biology or sociology
 - Better awareness of NLP-related risks etc. → *informed* consent
- Peer reviews as text genre are interesting from the NLP/CL perspective
- Liberal copyright for papers (ACL Anthology)

- Develop datasets, approaches and tools for own use within the ACL, then export to other communities

Setup

General workflow and parties involved

- Authors submit blind-submission versions
- Reviewers provide the initial reviews (text + scores)
- [optional] *Rebuttal*
 - Authors and reviewers debate in the rebuttal stage
 - Reviewers might update their reviews
- [optional] *Discussion*
 - Reviewers discuss among themselves in the discussion stage
 - Reviewers might update their reviews
- Program chairs aggregate the reviews and make a final decision
 - [optional] Program chairs prepare meta-reviews
- Authors of the submitted papers provide camera-ready
- External public

Data involved

- Blind submissions
- Initial reviews (text and scores)
- Rebuttal texts
- Discussion board logs
- Final reviews (after Rebuttal and Discussion)
- Meta-reviews
- Camera-ready versions
- Metadata

Numerical data

*ACL Meta-Science Repository

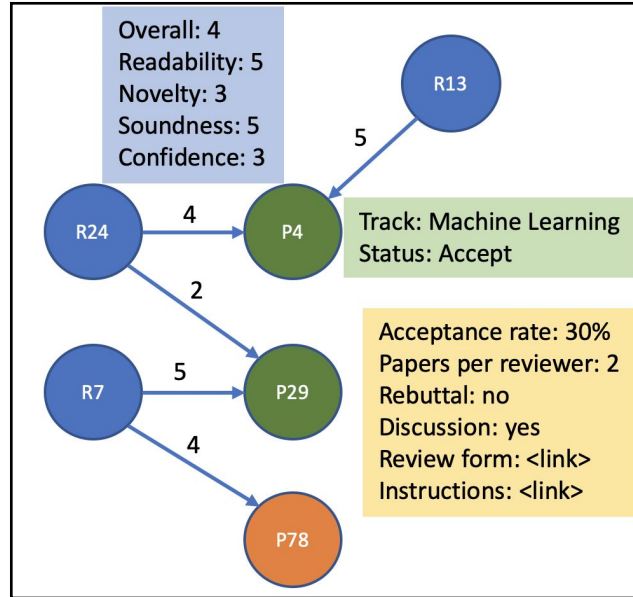
Idea

- Most issues come from peer review texts, unpublished work contents and reviewer identities
- However, there is a lot of **purely numerical data** generated by each *ACL event. It is anonymous and publicly reported by PCs at each conference
 - Acceptance rates
 - Score distributions
 - Workflows (rebuttal, discussion)
 - # submissions per reviewer
 - Etc.
- Collected each time and not publicly available in machine-readable formats
- We propose to **standardize the conference reporting** and make numerical data publicly available on a dedicated website

Data to publish

- Scores (overall, aspect, confidence)
- Anonymous reviewer (R152) and paper (P241) identifiers
- Reviewer-paper graph (anonymous)
- Conference metadata
 - Rebuttal/Discussion yes or no
 - Acceptance rates
 - The reviewing template used
 - # reviews per reviewer, # reviewers per paper
 - etc.
- Principled way to **test peer reviewing workflows** and policies, **monitor consistency**, study score-related biases, etc. Enables meta-studies.
- Useful for NLP community and for **bridging to the meta-science** and general peer reviewing research outside NLP

Data example



Implementation

- Can be **almost entirely automated**
- No anonymity / privacy / copyright issues; just numbers
- **Dedicated website** with numerical data from past events, basic statistics and visualizations.
 - For the community
 - For researchers in meta-science and science-of-science
 - For program chairs to generate reports
 - For program chairs to pick the workflows and policies that work
- **UKP Lab will manage data collection and the website**
 - if proven useful, the ACL community may consider overtaking the service for long-term maintenance.
- Possible extras
 - Citation counts and other scientometrics for accepted papers
 - Monitoring the state of rolling review
 - Auxiliary supervision for score-based NLP models

Textual data Challenges and Solutions

Overview of Challenges

- Anonymity
“I want to stay anonymous to the authors and reviewers”
- Privacy and Security
“I don’t want negative feedback to be published, and I worry about my research ideas”
- Copyright
“Who owns the data, how can this data be used and who gets the credit”
- Consent
“We don’t want to trick the reviewers and authors or get them into trouble”
- Management overhead
“PCs are very busy”

Anonymity (D1)

- In double-blind review authors and reviewers are anonymous to each other
- Challenges:
 - Some reviewers **sign their reviews** (identity open to the authors but not to the public)
 - Reviewer identities might be **disclosed in the discussion** and meta-reviews to the other reviewers, but not to the authors and not to the public
 - Authors of accepted papers are identifiable via ACL Anthology (**anyway**)
 - Authors of rejected papers are identifiable via arXiv or subsequent publication (**anyway**)
- Solutions:
 - **No metadata** linking to the reviewer or author identities is collected.
Identity can be disclosed voluntarily as part of the license attribution (D3).
 - Reviewers are **notified in advance** to make an informed decision about signing the reviews
 - Reviewers and authors are notified about the **potential risks** of de-anonymization via author profiling techniques by malicious parties

Anonymity (D1)

- Option 1: Reviewer identity entirely hidden
- Option 2: Global reviewer ID within venue
 - Anonymize the reviewers using a random, but stable identifier, e.g. [Ilia Kuznetsov](#) → R162
 - Allows grouping reviews from the same reviewer together. Enables:
 - Preference learning
 - Reviewer-level tasks (as opposed to single-review-level tasks)
 - Slightly increases anonymity risk
 - **Challenge:** anonymity leak in one review (e.g. signed despite being informed) exposes the authorship of other reviews
 - **Solutions / Options:**
 - Notify the reviewers in advance
 - Global ID only consistent within one venue / event / time span
 - Only use the Global ID in the numerical dataset

Privacy and Idea Safety (D2)

- Most data in the current peer reviewing workflow is private.
- Two main challenges in making it public:
 - Reviews contain **negative feedback**
 - Leaking **unpublished ideas** (through blind submissions or through reviews)
- Data sensitivity depends on data type and the acceptance status of the paper:

	neg. feedback	idea safety
blind submission / accept	N/A	no risk (public)
reviews and rebuttals / accept	low risk	no risk (public)
discussion boards	medium risk	medium risk
reviews and rebuttals / reject	high risk	high risk
blind submission / reject	N/A	high risk

Privacy and Idea Safety (D2)

- Solutions

- Only collect data for **accepted papers**
- Explicit, **informed consent** from both authors and reviewers.
- Authors agree to publication of the reviews.
- Later option for rejected papers: a **privacy period** of **TWO YEARS**
 - Minimizes the risk of leaking research ideas from rejected papers
 - Reduces the risk of negative feedback impact

- Data access for the community

- **Open datasets** for accepted papers and their reviews (if consented to);
- **Secure storage** and protected experimental environment for cases where reviewers agree but authors disagree; later also usable for rejected papers
 - Data stored on a secured server and accessed in a shared task environment, e.g. TIRA
- Sharing **data derivatives**
 - Models trained on data can be shared, at a slight theoretical privacy risk

Copyright (D3)

- OpenReview terms of service and licensing in progress; ACL data is private (reviews, blind submissions) or CC-BY (published papers).
- Challenges
 - Need a licence to enable research on the datasets
 - CC **requires attribution**, but we want to keep reviews and blind submissions anonymous
 - CC is **irrevocable**
- Solution (see the [doc](#) for details)
 - *Copyright © 2021 administered by the Association for Computational Linguistics (ACL) on behalf of ACL content contributors: Professor John Smith, Dr. Susan Lee, Dr. Michael Jones, and other contributors who wish to remain anonymous. Content is made available under a [Creative Commons Attribution 4.0 International License [creativecommons.org]].*
 - Contributors transfer a license to ACL and have an option to voluntarily disclose their identity. The agreement must make it clear that the CC licenses are **irrevocable** unless there is breach.

Consent (D4)

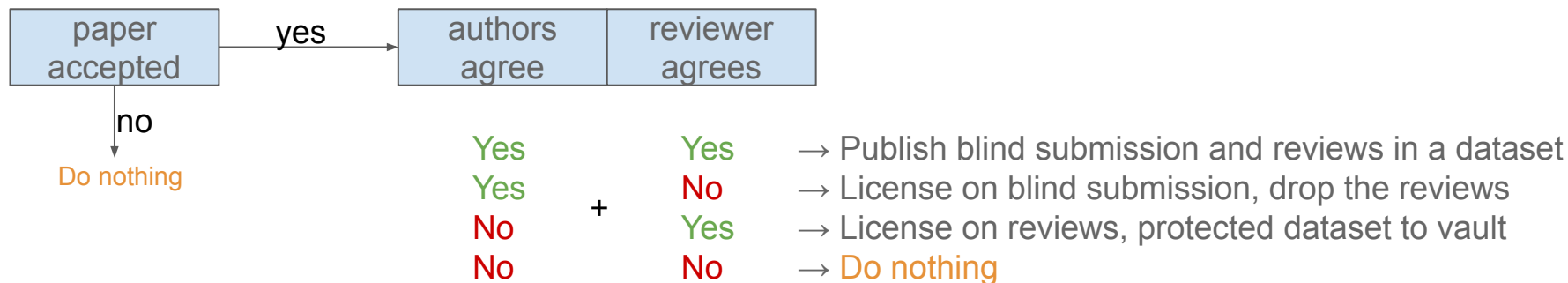
- The contributors must **explicitly consent** to the data collection, be informed about the purposes and the associated risks, and given agency over their data (GDPR).
- Challenges
 - Skewing the distribution towards accepted papers and benevolent reviews
 - Obtaining consent during the event creates a stress situation; obtaining consent after the event reduces the turnaround
 - Consent can be withdrawn. This would affect the derivative datasets and harm reproducibility
- Solutions
 - **Explicit consent** on paper/venue basis, **no** default opt-in.
Make it clear that donating the data is **optional**.
 - **Notification** about future consent collection **before** the reviewing begins.
Consent collection after the acceptance notification.
 - **Conflict with licensing**: License cannot be withdrawn, but consent can.
 - **For GDPR, a License (D3) provides stronger legal basis for data processing than consent.**

Workload (D5)

- Challenge:
 - Getting the license and extracting the data creates **additional work** for the program chairs.
- Solution:
 - **External workflow** managed by the community / moderator
 - Consent collection **integrated into the conference management system** or managed via external provider.
 - Forms and most data manipulations can be easily automated. Automate whenever possible.
 - High automation potential with transition to OpenReview.

Summary Workflow (D6)

- Based on the above, we propose the following **general workflow** as the first iteration.
- All participants are notified about the future data collection in advance.
- Upon review completion, we ask reviewers for their **reviews**.
- After the acceptance decisions, we ask the authors of **accepted** papers for their **blind submission**, and **permission to publish peer reviews** for their paper.
- If reviewers agree to publish their reviews, but the authors don't want that, the reviews are put to a vault (e.g. TIRA) and will only be used for internal and testing purposes
~ shared task evaluation setup



COLING-2020 Pilot

Pilot overview

- Test-drive for textual and numerical data collection at COLING-2020
- Uses an earlier version of the workflow
- Goals
 - Test the workflows and measure the PC workload
 - Get community feedback
 - Collect data with clear consent and copyright status
- Implemented measures
 - **Anonymity:** no personal data, global reviewer identifier, only reviews before discussion
 - **Privacy:** explicit mention of risks in the consent form; only approach authors of accepted papers, but reviewers for all papers; 2-year privacy period
 - **Copyright:** CC-0
 - **Consent:** explicit consent from authors and reviewers, notification after acceptance deadline
 - **Workload** for PC: two e-mail messages; authorize access via SoftConf; negotiating the consent forms (since no standard form is yet available)

Consent forms

Reviewers

https://forms.office.com/Pages/ResponsePage.aspx?id=HBOHA8qCXkqX_VQgmzB6s4Y3gOE-fypItlQ_eix2fkBUN1pJUjIFUTBVNTFTUkpVR05GTk1MNkYyNC4u

Authors

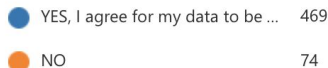
https://forms.office.com/Pages/ResponsePage.aspx?id=HBOHA8qCXkqX_VQgmzB6s4Y3gOE-fypItlQ_eix2fkBUMVIENzNHOUVOT0ZJVIFOWDhJVzNPRUtDUS4u

Results

- Out of 1.5k reviewers, ~500 participated
- Out of ~580 authors of accepted papers, 190 participated. Half of the authors who gave consent also provided TeX sources.
- Collected 1300 review texts, 140 blind submission PDFs and 80 TeX sources
- ~ equivalent to ICLR-2017 but with clear copyright and consent

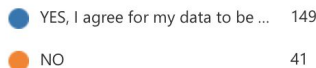
I agree that the text and structured data from my peer review reports and discussion boards can be freely used for the purposes and under conditions described above.

[Weitere Details](#) [Insights](#)



2. I agree that the blind-submission versions of my papers can be freely used for the purposes and under conditions described above.

[Weitere Details](#)



Discussion

- Generally encouraging feedback, useful questions and workflow details from the community
- Heads-up e-mail before consent collection is important, so that people keep it in mind while writing their reviews
- Discussion boards' logs are questionable as they contain informal discussions between reviewers
- Turnaround could be higher (~30% survey participation)

Summary

Recap

- We want to start a constructive discussion on peer reviewing data in the ACL community
- We propose a numerical repository to enable meta-studies of peer reviewing at ACL, at no privacy risk
- We collect the main issues and risks associated with peer reviewing texts and propose solutions to address the challenges related to Privacy, Anonymity, Consent, Copyright and Workload
- We report on a pilot **peer reviewing data collection campaign** at COLING-2020

Questions and discussion points

- *ACL Numerical repository?
- What are the remaining issues and concerns regarding the peer reviewing text collection and how can they be addressed?
- Is the proposed summary workflow good?
- What is the opinion of the community on the issues and solutions?
- How to make the data most useful to the ACL community?