

CHI research material availability study protocol

Motivation

CHI community is proud of its diversity of research methods and contributions. Due to this methodological diversity, research publications generate different types of research artifacts (e.g., paper prototypes, software, hardware, study protocols, interview scripts, affinity diagrams, or data analysis code).

Recently in other scientific disciplines such as psychology, medicine, and economics, there is a push for transparency and openness of research. This is implemented by making research materials available for other scientists to replicate, reproduce, scrutinize, and build upon. This topic was present at CHI 2018 in the form of paper ([Cockburn et al. 2018](#)), SIGs ([Chuang & Pfeil. 2018](#); [Wacharamanotham et al.. 2018](#); [Remy et al., 2018](#)), and alt.chi ([Echtler & Häußler, 2018](#)). In these events, discussants pointed out that this methodological diversity muddled the criteria for transparency and openness. Furthermore, most HCI research involve human users and, thus, collected data could be subject to privacy concerns. Finally, HCI research has the potential for commercialization. Therefore, concerns about intellectual properties could further discourage (or preclude) researchers from sharing their research materials in the public domain.

Nevertheless, it is unclear how diverse the research materials generated from CHI community is, and how much of them are currently available for further research. It should also be noted that the social sciences (e.g., psychology) rely on diverse methods and the challenges faced in HCI are, unlikely, to be entirely unique. To move towards the transparency and openness that is appropriately defined for CHI community, a necessary first step is to chart out the current landscape of generated research materials and their availability. Therefore, in this questionnaire, we aim to determine the current state of generated materials and their present state of availability in CHI community.

Research questions

1. What is the current situation of the availability of each type of research materials in CHI community?
2. What factors prevents the authors from making these research materials public?

Hypotheses and predictions

Operationalization

Participants

Participant criteria and recruitment strategy

We will invite all first authors of CHI 2018 paper to participate in our survey. The invitation will be sent by email based on the email address that the author specified in the paper (retrieved from the confer system).

Participant factors

Some of the first authors may be graduate students who are inexperienced in some of the terminology used in the questionnaire. We mitigate this by providing elaborated description and examples in the survey. Besides, we ask demographic questions to assess the quality of responses.

Compensation

We will raffle ten prizes of 10 EUR-equivalent gift cards.

Risks

Answering questionnaire about specific publication may risk revealing personal attitude to the community and may jeopardize future career. Therefore, we decided to dissociate the questionnaire from individual publication. The respondents can voluntarily associate the responses with specific paper DOI at the end of the survey, but we will not release DOI-associated data public due to the concern above.

Procedure

1. ✓ Draft the [questionnaire](#)
2. ✓ Elicit feedback from people who expressed their interest in Open Science through personal connection and public outlets
 - a. SIG TOP @ CHI 2018
 - b. SIG Evaluation @ CHI 2018
 - c. Transparent statistics mailing list
 - d. RepliCHI mailing list
 - e. SIGCHI and CHI-meta Facebook groups
 - f. Contacts from Open Science Foundation
3. ✓ Conduct online focus group to elicit more in-depth feedback
4. Send the questionnaire to all first authors using the contact data from confer system
 - a. Phases
 - i. Phase 1: send the questionnaire out to 10% of the authors (stratified sampling by session) and wait one week for the responses to check and fix potential problems
 - ii. Phase 2: send the questionnaire out to the rest of the authors
 - b. Allow one month (July 2018) to collect the responses
 - c. If the first author is unreachable (the email bounced back), re-send to the next author as listed on the paper.
 - d. Email a reminder two weeks after the initial invitation.

Data collection & analysis

Data collection

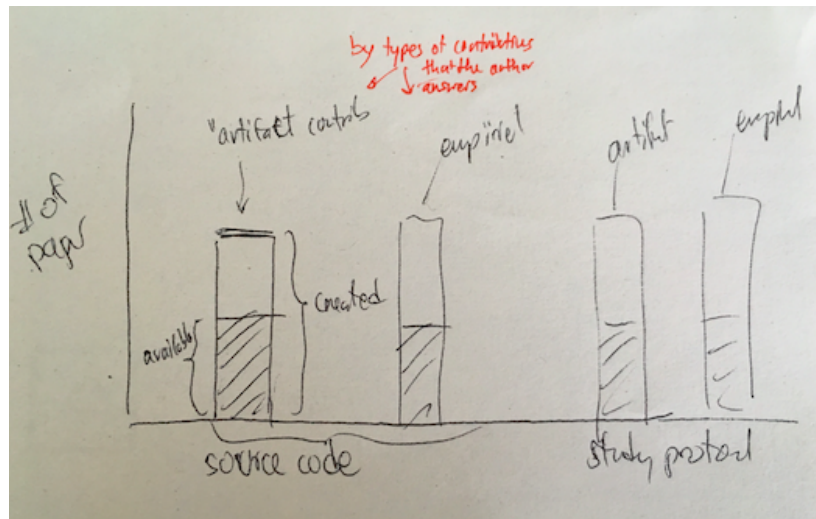
The data is collected on via a LimeSurvey platform installed on zpac.ch/chi_material_survey. The survey will be open for one month.

Error management

Selection bias is a potential threat to validity. (People who care about open science may be more likely to answer this questionnaire.) However, due to our decision avoid associating responses with DOIs (see Risks section), we cannot collect non-participant data. Since there's no population data available, we cannot [statistically correct](#) for such bias. We will acknowledge the potential of this bias as a limitation.

Data analysis plan

- The data analysis will be conducted twice: one week after releasing phase 1 invitation and at the end of the survey period.
 - The phase 1 analysis is intended to check for problems in the survey. We will run the whole analysis script and read responses from open questions.
- Proportion of produced research materials and those that are made available ← describe by research contribution type and research material types with the following visualization



- By type of research material, plot the frequency of responses about the channel that the materials are made public
- By type of research material, plot the frequency of responses about the reasons preventing the materials from being made public
- Proportion of overall replies received to the email sent out over time → responsiveness of the questionnaire is an indirect indicator of responsiveness of the request for research materials
- Plot proportion of material availability by research lab size
- Plot proportion of material availability by country of the lab

Questionnaire Design Rationale

Overall:

- We are interested in the association between the “why” associated with “what”. We ask both in this survey to (1) provide a concrete case for the respondent to think about instead of eliciting their opinion in abstract project and (2) to limit the number of “why” questions we need to ask the respondents

About questions:

- Informed consent may influence how the participant responses. Thus, they are considered as a part of study material
- Debriefing materials may influence participants’ responses in subsequent sessions in a multi-session study, thus they are considered as a part of the study material

Dissemination of our research materials

- We will make the responses from the questionnaire available after anonymizing fields including such as voluntary DOI and email address

Future work

- Apply the same survey for previous CHI conferences to assess how the availability evolves over time.
- Survey attitude of CHI community about which solutions would be preferred for each types of materials (opt-in vs. opt-out)
 - Antti: “It would be interesting to see what they think about different solutions to this (e.g., conferences/journals expecting release of data/code when possible...)”
 - “open science should be OPT-OUT, i.e. it should be the default choice, but exceptions would of course be allowed with good reasons (which will happen, no doubt). Presently openness is OPT-IN and only a small proportion even considers it.”

- If the number of responses is sizable, try using the responses with the volunteered DOI to train an ML algorithm to detect contribution types and research materials from the text. (Apply LDA to extract topic model)

Contributor names for the protocol

If you add comments above, please type your name here, so we can acknowledge you in the [contribution log](#).

Names:

Lewis L. Chuang

Kelly Caine

Pro memoria: [unused stuff from previous versions of the protocol](#)

“replication”: to effectively repeat an experiment.

“reproduce”: to reproduce results from the same dataset.

<http://science.sciencemag.org/content/334/6060/1226>

- Additional measures to avoid or assess self-selection bias ([possible methods](#))

Contributor roles: http://dictionary.casrai.org/Contributor_Roles