SOUNDSOFTWARE.AC.UK PRIZES FOR REPRODUCIBILITY IN AUDIO AND MUSIC RESEARCH

Chris Cannam, Luís A. Figueira and Mark D. Plumbley

Queen Mary University of London
Centre for Digital Music
{chris.cannam, luis.figueira, mark.plumbley}@eecs.qmul.ac.uk

1. INTRODUCTION AND MOTIVATION

SoundSoftware is an EPSRC-funded Software Sustainability project, which aims to support the sustainable development and use of software and data to enable high quality research within the audio and music research community. One of our objectives is to encourage researchers to carry out reproducible and repeatable experiments and to publish code and datasets along with their papers.

In a survey of UK we conducted in 2010, we found that even though most audio and music researchers develop software during their research, most do not release or share it. In our 2012 paper [1] we identified a set of barriers to open publication: lack of education and confidence with code; lack of access to facilities and tools; platform incompatibilities; and, significantly, lack of incentive.

In earlier work ([1, 2]) we examined various approaches to the first two barriers. It has been suggested [3] that publications are cited more if they come with software and data. But this long-term advantage does not always provide the immediate incentive researchers need at publication time.

In order to motivate researchers, raise awareness of reproducible research, and offer an immediate incentive, we decided to sponsor Reproducible Research Prizes for Audio and Music Researchers. These would be awarded for publications whose work could be repeated using associated software and datasets, particularly containing work deemed influential to the UK audio and music research community.

Here we describe the process of organizing the first edition of this award. We then describe the entries and the evaluation process for each entry, describing general issues common to many entries. We close with ideas for future work and conclusions.

2. PROCESS

We made the initial call for participation on the 24 April 2013, asking participants to submit entries within 3 weeks.

Entries were to be judged in one of four categories:

- Journal submission, unpublished

- Conference submission, unpublished

- Journal paper, published

- Conference paper, published

The categories were chosen in order to be able to offer prizes relevant to the type of work: payment of open-access Article Processing Charges for published work, or travel bursaries to present the work at a conference or visit a UK institution.

Entrants were asked to provide links and and instructions on how to reproduce the work, to explain how the work would enable high quality research in the UK audio and music research community, and to describe steps taken to ensure sustainability of the software for future researchers.

3. PARTICIPATION

We received a total of 13 entries, from 12 different researchers. Of these, half were based in the UK (3 being from our own institution, Queen Mary University of London). The entries came from 10 different institutions, with only 4 being in the UK.

4. EVALUATION

Submissions were assessed against three criteria, each evaluated by a separate panel:

Ease of reproducibility of the results We attempted to obtain and run the software associated with the paper and regenerate the results shown. We only applied this criterion where software was included (some submissions used datasets but not software).

Quality of sustainability planning A team from the Software Sustainability Institute[1] examined the submissions, taking into account whether the code and data were stored in a suitable repository, whether version

---

[1]http://www.software.ac.uk/

control was used, which tools for community involvement were available, and whether the work was properly licenced.

Potential to enable high quality research in the UK audio and music research community Each submission was sent to two external reviewers recognised in the field and their scores were averaged.

We eliminated those submissions which received the lowest grade of score in any single criterion, and shortlisted the rest. Shortlisted submissions were assigned to categories according to the type of publication they contained, as listed in the call, and the winner in each category was that with the best average score across all criteria. More information about evaluation is available from our website[2].

5. OUTCOMES

The winning entries were announced at the SoundSoftware Workshop 2013[3] and on our website[4].

The winning journal submission, submitted by Bob Sturm, was a report on the reproducibility and quality of previously-published work from other authors; the winning conference submission, from Dimitrios Giannoulis, was for a challenge in acoustic scene classification with reference implementations; and the winning published conference paper, from Piotr Majdak, was for a transfer format for audio data. There were no entries for published journal papers, so we awarded a special prize for a Technical Report from Colin Raffel reproducing earlier work in large-scale pitch analysis. We also awarded two Honourable Mentions, to Dan Stowell and to Polina Proutskova.

Of note is that three of the four prizes went to work produced outside the UK; and all of the awards went to work explicitly intended to facilitate or report on sustainability and reproducibility in audio and music research, rather than to research work that just happened to be reproducible. While the awarded papers were all excellent work, this suggests there is room for improvement across the whole body of work in this field in the UK.

6. LESSONS LEARNED

Our prize categories were not a good match for the submissions actually received. For instance, we received 6 published conference papers, but no published journal papers. In a future edition of this award we will choose a different set of categories, perhaps separating works that provide infrastructure for reproducible and sustainable research from works that are reproducible examples of normal research publications.

Evaluating replicability of the submissions was hard, because of the wide range of technologies involved: this is one reason why sustainability is difficult in general. We don't wish to mandate a particular platform, but we will make more explicit the necessity to repeat the experiments in the paper, ideally from a single command, and suggest that submitters test this on a second computer or virtual machine before entering. We are also planning to publish a list of the good practices we observed in the best submissions received.

In this first edition of the prizes, we did not make detailed information about the evaluation criteria available until we published the results. In any future editions, we will publish evaluation criteria along with the initial call.

7. REFERENCES

[1] C. Cannam, L. A. Figueira and M. D. Plumbley, "Sound Software: Towards Software Reuse in Audio and Music Research", in *Proceedings of the IEEE 2012 International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012)*, 2012, 2745-2748.

[2] L. A. Figueira, C. Cannam and M. D. Plumbley, "Software Techniques for Good Practice in Audio and Music Research", in *Proceedings of the 134th International Audio Engineering Society Convention (AES 134)*, 2013.

[3] Vandewalle, J. Kovacevic, and M. Vetterli, Reproducible research in signal processing – what, why, and how, in *IEEE Signal Processing Magazine, vol. 26, no. 3, pp. 3747*, 2009

8. ACKNOWLEDGMENTS

---

[2] http://soundsoftware.ac.uk/rr-prize-how

[3] http://soundsoftware.ac.uk/soundsoftware2013

[4] http://soundsoftware.ac.uk/rr-prize-winner-announcement