

Special Issue Editorial: Crowd AI for Good

FRANK O. OSTERMANN, University of Twente

LAURE KLOETZER, University of Neuchatel

MARISA PONTI, University of Gothenburg

SVEN SCHADE, European Commission – Joint Research Centre (JRC)

ABSTRACT

This special issue editorial of the *Human Computation* journal on "Crowd AI for Good" motivates explorations at the intersection of artificial intelligence and citizen science, and introduces a set of papers that exemplify related community activities and new directions in the field.

EDITORIAL

This special issue explores the cross-section of two topics that have seen an concomitant rise in the research communities of many disciplines: Machine learning (ML) as a subset of artificial intelligence (AI; compare Battista and Roli, 2018), and Citizen Science (CS; compare Vohland et al, 2021).

Over the past 20 years, CS (including online CS) and AI have been seen rapid growth and a rise in popularity. This double success seems to occur on parallel tracks. Machine learning technology has advanced thanks to the cross-fertilization of computer and cognitive science, and sophisticated models have been proposed in computer vision, speech analysis, music processing and bioinformatics, to name a few areas. Citizen science has expanded among scientific disciplines, from ecology, biology and astronomy, to medicine, geography, mathematics, history and climate change, for example.

The many advances in scientific insight we have witnessed in the past years should not cover up some of the challenges that these approaches face. Namely, ML/AI has come under scrutiny and received criticism because of black box models, and biases that have been caused by unbalanced

training data (see, for example, Annoni et al. 2018). For example, If the training data contains racist and sexist content, or omits important data and is insufficiently diverse and not representative this will be reflected in the model outputs as well. However, labeling training data, and validating model outputs, still require substantial human supervision and work, which is why many applications rely on the same training data sets, and not often enough closely inspect the outcomes. Concerning CS, it has been initially viewed with skepticism because of the unknown quality of the crowd's contributions. Although this has been demonstrated to be a largely unfounded concern, CS still faces the issue of long-term community building (retaining contributors for longitudinal studies), and scaling up to larger geographic areas. While some spectacular insights have been generated through CS research spanning decades¹² (Schade et al. 2020), the big CS platforms are also “graveyards” of many projects that were abandoned too quickly, mostly because of insufficient participation. Further, the syntactic and semantic integration of the various smaller data sets also is an unsolved challenge.

Combining both approaches, citizen science at its best is an inclusive and participatory approach to science that goes beyond crowdsourcing of tasks. It therefore offers great opportunities to address some of the issues from ML/AI. The potential to combine the complementary strengths of ML/AI and CS (e.g., ML/AI ability to scale up easily for large data sets, and CS manual validation to ensure meaningful results) to address their respective weaknesses (e.g., spurious patterns in big data sets, and sustainable organization of large volunteer groups), has been pointed out before (Ostermann et al., 2018). More recently, Rafner et al. (2021) developed a framework for Hybrid Intelligence that aims to facilitate systematic research into combining CS and ML/AI. Indeed, new “Crowd AI” systems, which combine ML/AI and CS, have led to unprecedented capabilities, and for certain problem classes may constitute the only viable solutions today.

At the same time, growing awareness of the broad applicability of AI to societal problems including humanitarian issues has led to recent initiatives in the public and private spheres, including Microsoft's AI for Good, Google's AI for Social Good, as well as the UN's AI for Good Global Summit, which seeks to apply machine intelligence to 17 Sustainable Development Goals (SDGs) set by the UN General Assembly.

Hence this special issue, for which the first plan originated from the workshop *Exploring the interplay between Human Learning and Machine Learning* - The Citizen Science Perspective, organised by the Joint Research Centre (JRC) in collaboration with the COST Action “[Citizen Science to promote creativity, scientific literacy, and innovation throughout Europe](#)”. The workshop was the first event within this COST Action devoted to the interaction between machine

¹ <https://www.theguardian.com/environment/2021/may/18/world-is-home-to-50bn-birds-breakthrough-citizen-science-research-estimates>

² <https://www.theguardian.com/environment/2017/oct/18/warning-of-ecological-armageddon-after-dramatic-plunge-in-insect-numbers>

learning and human learning in citizen science, and aimed to raise awareness about opportunities and issues emerging from this interrelation (Kloetzer, Schade, & Ponti, 2019).

Some important questions that this special issue aims to address include: What are current examples of such human-machine collaboration in CS? Can we create a typology of current various uses of ML/AI in CS, and can we identify the conditions that make such projects successful? What are their current and remaining problems and limits? And last but not least, what can we learn from the successful case of Citizen Science which could be useful for understanding and designing other situations of human – machine collaboration?

This special issue also intends to report on initiatives at the intersection of these two movements by examining the opportunities, methods, and impacts of Crowd AI as applied to societal problems.

The journey was delayed by Covid-19 pandemic, but now the special issue is out, featuring diverse accounts of the successful combination of AI with CS and participatory science, namely:

- Ponti et al. report on a recent symposium held at the ECSA 2020 international CS conference, capturing some vibes from this community.
- Chamberlain et al. examine how to use lightweight, decentralized engagement forms with social networks for connecting formal or authoritative agencies and institutions with "organic" (spontaneous, unorganized) collaboration, without demotivating or scaring off participants.
- Portela tests conversational agents (aka chatbots) as interface agents for communication with citizen science project participants, and suggests design elements to facilitate using conversational agents in new CS projects.
- Palmer et al. show that the combination of humans and machine as "Crowd AI" allows to unlock the potential of ecological Big Data, in this case more than 1,500 camera traps in southern and eastern African countries, that generate millions of images each year.
- Anjum et al. investigate hybrid workflows for three popular image analysis tasks, where deep learning is combined with crowdsourcing to verify and correct predicted image labels, and generate new annotations for images with low prediction confidence.

Together, these papers offer together a glimpse of the richness for potential human-machine computations, while providing state-of-the-art ways forward. We hope that this special issue will

be one of the stepstones to advance the research in the field, and to enable collective intelligence for real social benefit.

1. REFERENCES

- Annoni, A., Benczur, P., Bertoldi, P., Delipetrev, B., De Prato, G., Feijoo, C., Fernandez Macias, E., Gomez Gutierrez, E., Iglesias Portela, M., Junklewitz, H., Lopez Cobo, M., Martens, B., Figueiredo Do Nascimento, S., Nativi, S., Polvora, A., Sanchez Martin, J., Tolan, S., Tuomi, I., & Vesnic Alujevic, L. (2018). Artificial Intelligence: A European Perspective. *EUR 29425 EN, Publications Office of the European Union, Luxembourg, 2018*, ISBN 978-92-79-97219-5, doi:[10.2760/91283](https://doi.org/10.2760/91283), JRC113826.
- Biggio, B. & Roli, F. (2018). Wild Patterns: Ten Years after the Rise of Adversarial Machine Learning. *Pattern Recognition* 84 : 317–31. <https://doi.org/10.1016/j.patcog.2018.07.023>.
- Kloetzer, L., Schade, S., & Ponti, M. (2019). Exploring the interplay between Human Learning and Machine Learning. *Report Citizen Science COST Action 15212*. <https://cs-eu.net/news/report-wg-2-wg-4-exploring-interplay-between-human-learning-and-machine-learning>
- Ostermann, F.O., Garcia-Chapeton, G., Kraak, M-J., % Zurita-Milla, R. (2018). Towards A Crowdsourced Supervision Of The Analysis Of User-Generated Geographic Content: Engaging Citizens In Discovering Urban Places. *Zenodo Pre-Print*. doi: [10.5281/zenodo.1215681](https://doi.org/10.5281/zenodo.1215681).
- Rafner, J., Gajdacz, M., Kragh, G., Hjorth, A., Gander, A., Palfi, B., Berditchevskaia, A., et al. (2021). Revisiting Citizen Science Through the Lens of Hybrid Intelligence. *ArXiv:2104.14961*. <http://arxiv.org/abs/2104.14961>.
- Schade, S., Tsinaraki, C., Manzoni, M., Berti Suman, A., Spinelli, F.A., Mitton, I., Kotsev, A., Delipetrev B. & Fullerton. K.T. (2020). Activity Report on Citizen Science – discoveries from a five year journey. *EUR 30551 EN, Publications Office of the European Union, Luxembourg, 2020*, ISBN 978-92-76-28369-0, doi:[10.2760/172609](https://doi.org/10.2760/172609), JRC123500.
- Vohland, K., Zandstra, A. L., Ceccaroni, L., Lemmens, R., Perelló, J. Ponti, M., Samson, R. & Wagenknecht, K. (eds.) (2021). The Science of Citizen Science. *Cham: Springer International Publishing*. <https://doi.org/10.1007/978-3-030-58278-4>.