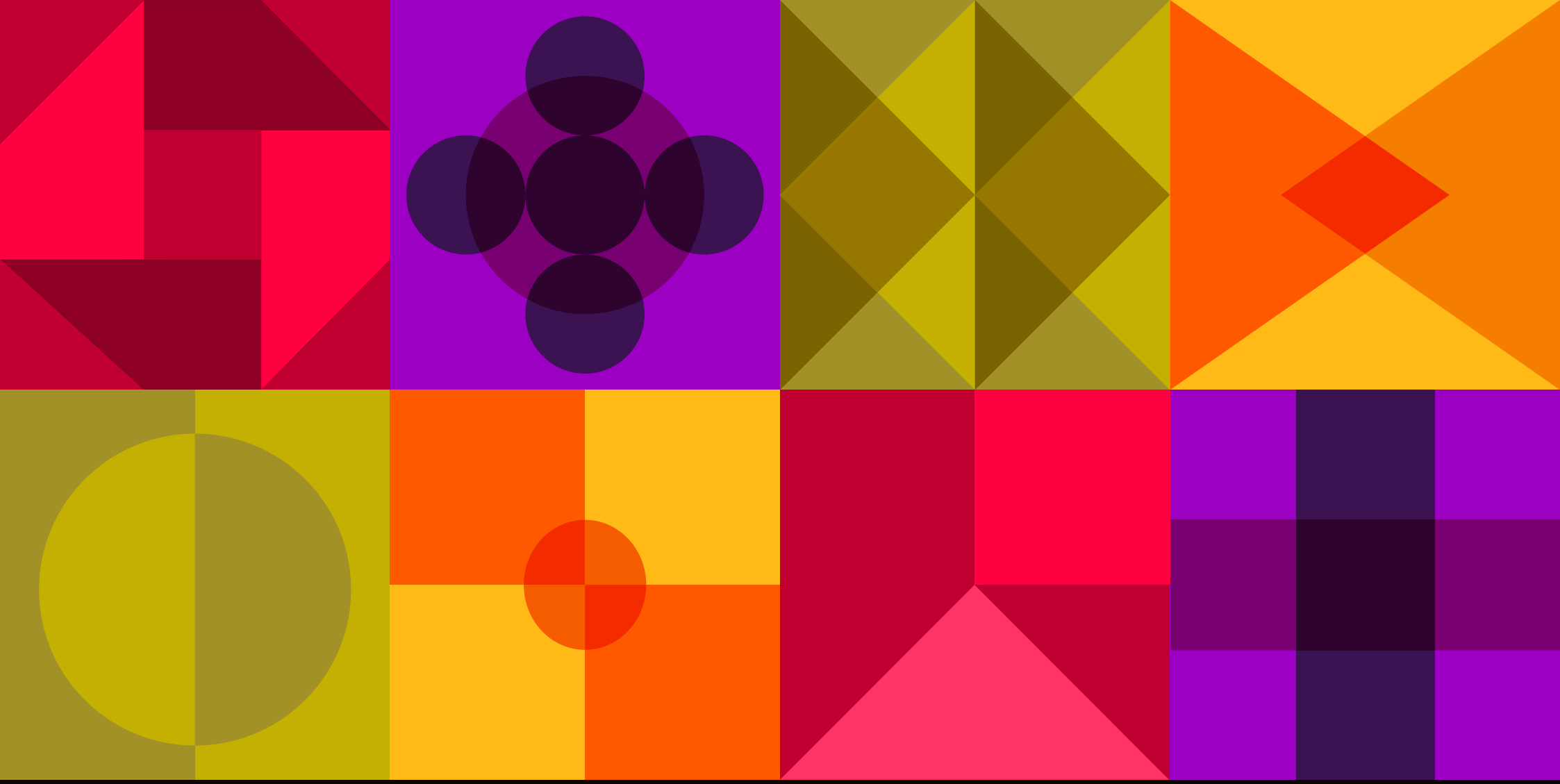

nesta

Patrick J McGovern FOUNDATION

wellcome

OMIDYAR NETWORK

Collective Intelligence Grants Programme

Experiments in collective intelligence design for social impact

Second cohort: 2020-2021

AUTHORS

Issy Gill
Ian Steadman
Kathy Peach

ACKNOWLEDGEMENTS

This report is based on the findings from 15 collective intelligence experiments that were co-funded (but not conducted) by Nesta, the Wellcome Trust, Omidyar Network and the Patrick J. McGovern Foundation.

The experiments were carried out by: Grzegorz Rutkiewicz, Patrycja Tempska, Maria Dowgiałło, Michal Wroczynski, Gniewosz Leliwa and Maciej Brochocki (Samurai Labs), Christian Ochieng, Ruslan Masinjila, Mohamed Alsayed and Cleopa Otieno (Kenya Flying Labs), Ulrike Hahn and Jason Burton (Centre for Cognition, Computation and Modelling, Birkbeck University of London), Maria Camps, Lydia Garcia, Lin Lin and David Bermejo (Spotlab), Johannes Mutter, Jeffrey Pickett, Helen van Baal, Alexander Davies and Mario Lyon (neu by Augmented Thinking), Thomas Chesney (University of Nottingham), Swee-Hoon Chuah (University of Tasmania), Robert Hoffman (RMIT University), Bo Percival, Felix Delattre and Ramya Ragupathy (Humanitarian OpenStreetMap Team), Jacopo Margutti (Netherlands Red Cross), Bastian Greshake Tzovaras and Mad Price Ball (Open Humans), Marc Santolini, Leo Blondel and Thomas Landrain (Just One Giant Lab), Alexa Hagerty, Alexandra Albert, Igor Rubinov and Joe McCraw (Dovetail Labs), Vito Trianni, Ralf Kurvers, Valentina Presutti, Andrea Giovanni Nuzzolese and Alessandro Russo (Institute for Cognitive Sciences and Technologies of the Italian National Research Council), Irving Lin (Human Dx), Robert Trigwell, Prithvi Hirani, Nuno Nunes, Eduardo Zambrano, Beakal Gobena, Elvina Zirima, Pablo Gallego, Tashia Shupert, Debora Gonzalez and Lorenzo Guadagno (International Organization for Migration), Alessia Paglialonga, Marco Zanet, Marco Muselli (Institute of Electronics, Computer and Telecommunication Engineering of the Italian National Research Council), Aziz Guergachi (Ryerson University), David Barber, Rebecca Theal, Rachael Morkem (Queen's University, Canada), Karim Keshavjee (University of Toronto), Robert Kyba (Strategic Global Counsel, Toronto), Ling Tan, Usman Haque and Usamah Khan (Umbrellium), Iram Quarashi (Loop Labs), Franz Miralles (Swisscontact), Rafael Lindemann (RIMISP - Latin American Center for Rural Development), Carmen Tapia and Jesus Chumacero (Banco de Desarrollo Productivo), Martin Murillo (IEEE Humanitarian Activities Committee), Merihan Alhafnawi, Paul O'Dowd, Edmund Hunt, Severin Lemaignan and Sabine Hauert (University of Bristol) and other team members that might not be listed here.

We are very grateful to the following individuals at the organisations who co-funded this grant programme. Haidee Bell and Zaichen Mallace-Lu at the Wellcome Trust, Eshanthi Ranasinghe (formerly) at Omidyar Network, and Claudia Juech and Linda Tsai at the Patrick J. McGovern Foundation.

We would like to thank our colleagues at Nesta's Centre for Collective Intelligence Design for their involvement in supporting and managing the grants programme over this period of time, specifically Joana Murta Rosa, Eva Grobbink, Camilla Bertoncin, Aleksandra Berditchevskaia and Peter Baeck.

We would like to thank the members of our advisory group who help assess applications for the Collective Intelligence Grants. Julien Cornebise, Marc Driscoll, Louise Francis, Erik Johnston, Gitte Kragh, Chris Lintott, Massimo Menichinelli and Lucy Robinson. And the experts from other teams at Nesta who contributed to the assessment process, in particular Rob Fuller, Cath Sleeman, Juan Mateos-Garcia and Samuel Hanes.

We would also like to acknowledge the many hours of work contributed by participants, volunteers and citizen scientists without whom these experiments would not have been possible.

The articles in this report were written by Ian Steadman and the layout for this report was designed by Mike Green.

ABOUT NESTA

We are Nesta, the UK's innovation agency for social good. We design, test and scale solutions to society's biggest problems. Our three missions are to give every child a fair start, help people live healthy lives, and create a sustainable future where the economy works for both people and the planet.

For over 20 years, we have worked to support, encourage and inspire innovation. We work in three roles: as an innovation partner working with frontline organisations to design and test new solutions, as a venture builder supporting new and early stage businesses, and as a system shaper creating the conditions for innovation.

Harnessing the rigour of science and the creativity of design, we work relentlessly to change millions of lives for the better.

Find out more at **nesta.org.uk**

If you'd like this publication in an alternative format, such as Braille or large-print please contact us at:
information@nesta.org.uk

Published October 2021

01

Exploring AI-crowd interaction

We believe that to tackle complex problems we need to mobilise all the resources of intelligence available to us. That's why we funded five experiments that sought to understand how to best combine the complementary strengths of machine intelligence and collective human intelligence.

Using serious games to train AI models for medical diagnosis

Experiment 5: Can citizens playing online games be as effective as physicians in training AI models for image diagnosis?

Who is behind this experiment? Spotlab.

Key finding: AI models that have been trained on medical images annotated by non-expert adults and school children are as accurate at diagnosing diseases as AI models trained on expert-based annotations.

Who is this relevant for?

- National policymakers, practitioners and researchers in health.
- Citizen science projects and crowdsourcing platforms.
- International and humanitarian organisations.

What was the experiment?

This experiment tested whether AI models trained by the general public could outperform medical specialists in analysing medical imagery (digitised blood samples) for the diagnosis of global diseases such as malaria. The AI models were trained by volunteers (adults and school children) performing medical image analysis while playing serious online games (through a website or app) for global diseases diagnosis. The accuracy of models were then tested against the performance of medical specialists.

What did they learn?

The experiment found that AI models trained on images annotated by both adults and school children can obtain similar results to ones trained on expert-based annotations, of around 93 per cent accuracy. The team found that the minimum number of different responses (or annotations) from adults and school children needed to achieve a level of accuracy comparable to experts was 20. In fact, responses from 20 school children alone could achieve a similar level of accuracy to experts.

Millions of people are affected every year by diseases which go undiagnosed, simply because some healthcare systems struggle with limited capacity. Spotlab has been using a game called SpotWarriors to see if the wisdom of a crowd of non-experts is comparable to the expertise of trained doctors – and, in the process, empower people to become part of a global healthcare solution.

"Microscope diagnosis – looking at blood, skin, etc, under a microscope – is the gold standard for several diseases," explains Spotlab's Lydia Garcia. *"But it requires specialist time and resources. It's a very slow process."*

So, what to use instead? *"An AI model,"* says Lin Lin. *"We thought it would be faster. But the problem there is that training AI models requires a lot of data, and getting someone to label all that sample*

data is also difficult because doctors are so busy. So we asked, 'What if we get a group of normal people like us to label the images, and train the model with that data?' – and we prove it's possible."

So how do you get a group of people without expertise to label medical images? The answer is a game called SpotWarriors. After showing players example microscope images of certain diseases, it then asks them to spot the same signs of infection in other images. *"In each image, if the player attempts an identification – identifying a blood cell with a malaria infection, for example – and we know what the correct answer is, they either score or lose points,"* says Lin. *"But if they select a part of the image where we're unsure, it's marked with a question mark for further verification."*

The concept was first developed in 2013 by Spotlab's Miguel Luengo, with its first iteration, Malariaspot. *"He found that 22 non-expert players were equivalent to one specialist,"* says Garcia. *"Every player makes mistakes, but correct guesses tend to be clustered, so you can detect accuracy in the noise."*

The SpotWarriors game overview and mobile app interface

For the SpotWarriors project, the team wanted to test their hypothesis on both kids and adults. They ran 50 workshops in schools, with more than 1,000 children taking part (and this was made much more difficult by the severity of the COVID-19 pandemic in Spain). *"The teens were very competitive,"* says Garcia. *"They were like, 'Oh, now I'm a doctor!' That's really fun. This is a beautiful experiment in some ways, because it empowers people. One of our partners, Elena, is an expert on soil-transmitted parasites – but her mother is so obsessed with playing that she's more of an expert than her daughter now. It's awesome."*

"Every player makes mistakes, but correct guesses tend to be clustered, so you can detect accuracy in the noise."

Since people's guesses tend to cluster around the correct answer in each image in the game, the team could use those averages – from across 700 different 'levels' – to train a neural network in how to spot the same diseases from the same symptoms. They took a different dataset of 7,600 diagnostic images, and set the neural network loose on it to see how accurate it would be. The result? 93 per cent of the time it correctly diagnosed the relevant illness – and not only was that accuracy near-identical regardless of when either guesses from only adults or only kids were used to train the neural network, it was on par with the accuracy displayed by medical professionals in Spain and Kenya, who took part via a tele-microscopy platform. Lay people, even kids, can actually compare to experts when it comes to diagnosing certain diseases.

The fact that only 700 images are used to train the model is a limitation – expanding the dataset would likely improve the model even further. The team has floated the idea of having hospitals upload new diagnostic images to SpotWarriors in the future whenever an in-house expert is unsure of the correct diagnosis – say, they're only 90 per cent certain – and letting players collectively tackle it instead. Those guesses, in turn, would help improve the model further still by becoming part of its training dataset.

"We also learned a lot about the importance of game design here," says Garcia. *"With parasites, we designed the game to be like Tinder – you swipe if you see one. But with malaria, we made it more like scrolling around on Google Maps and placing pins. We got better results with the former."* The game also only works with diseases where the key markers fit in a box on a screen. But it has potential. Wherever people kill time by looking at their phones – on the subway, in bed, on the toilet – they could be analysing pictures on their phone, helping to train models like this.

"Everyone is an agent of change," argues Garcia. *"This game gives us a chance to promote principles like solidarity and empathy. It doesn't always have to be that solutions come from politics – they can be out there in society."*

A SpotWarriors school workshop

58 Victoria Embankment
London EC4Y 0DS

+44 (0)20 7438 2500
information@nesta.org.uk
@nesta_uk
www.facebook.com/nesta.uk
www.nesta.org.uk