AIMS & SCOPE

The 1st international workshop on Evaluating Progress in Artificial Intelligence (EPAI 2020, http://dmip.webs.upv.es/EPAI2020/) would be held in Santiago, Spain (August 30th, 2020)

Artificial intelligence (AI) and machine learning (ML) capabilities are growing at an unprecedented rate. Countless AI applications are being developed and can be expected over the long term. In hindsight, one would say that progress certainly has taken place just looking at the range of tasks that AI and ML are able to solve autonomously today (according to the benchmarks) and were not solvable a few years ago, from machine translation to medical image analysis or self-driving vehicles. Moreover, progress in AI is widely believed to have substantial social and economic benefits, and possibly to create unprecedented challenges. In order to properly prepare policy initiatives for the arrival of such technologies, accurate forecasts and timelines are necessary to enable timely action among policymakers and other stakeholders.

However, there is still much uncertainty over how to assess and monitor the state, development, uptake and impact of AI as a whole, including its future evolution, progress and benchmarking capabilities. While measuring the performance of state-of-the-art AI systems on narrow tasks is useful and fairly easy to do, where the assessment really becomes difficult, though, is in trying to map these narrow-task performances onto more general AI and how it can have an impact on society in terms of benefits, risks, interactions, values, ethics, oversight into these systems, etc.

This workshop will welcome formalisations, methodologies and testbenches for the evaluation of AI systems. The goal is also to measure the field's progress. More specifically, we are interested in theoretical or experimental research focused on the development of concepts, tools and clear metrics and indicators to characterise and measure AI/ML systems and how this relates to, among others, metrics of intelligence (and other cognitive abilities), and rates of development, progress and impact.

We want to bring together as wide a range of people as possible from most areas in AI and ML but we will also target people from cognitive science, comparative psychology, neuroscience, psychometrics, philosophy of science and technology, measurement theory, etc., with the aim of encouraging cross-disciplinary approaches and theoretical and experimental analysis of how AI evaluation should be done at present and in the future.

We welcome regular papers, demo papers about benchmarks or tools, and position papers, and encourage discussions over a broad list of topics (not exhaustive):

- Analysis of progress scenarios (simulations), AI progress forecasting, and associated issues and risks: privacy, safety and security, surveillance, inequality, bias, discrimination, transparency, regulations, accountability, sanctions, and workforce/management displacement.
- Proposals for new general tasks, benchmarks, competitions, evaluation environments, workbenches and general AI development platforms.
• Analysis and comparisons of AI/ML benchmarks and competitions. Lessons learnt.
• Theoretical or experimental accounts of the space of tasks, abilities and their dependencies.
• Methods for AI evaluation, including measures and indicators for their progress and impact.
• Analysis of disruptive AI technologies, AI readiness, and other indexes.
• Evaluation of the technical capabilities and performances of the major AI-based systems.
• Analysis of AI and its ethical, legal (law, regulation and governance), social and economic impact.
• Evaluation of the uptake of AI across different industries and sectors in the economy.
• Analysis of the impact of AI on employment: insights on the role of workplace organisation in shaping the effect of new technologies on labour markets (opportunities, challenges, etc.).
• Better understanding of the characterisation of task requirements, costs and difficulty (energy, time, trials needed...) beyond algorithmic complexity.
• Evaluation of social, verbal, reasoning and other general cognitive abilities abilities in multi-agent systems, video games, artificial social ecosystems, conversational bots, dialogue systems and personal assistants.
• Evaluation of multi-agent systems in competitive and cooperative scenarios, evaluation of teams, approaches from game theory.
• Assessment of replicability, reproducibility and openness in AI / ML systems.
• Dominant and neglected AI paradigms, limitations and possibilities.

INVITED SPEAKERS

Invited speakers to be announced.

SUBMISSION

We solicit submissions (full or short papers) including:

• Original research contributions.
• Applications and experiences.
• Surveys, comparisons, and state-of-the-art reports.
• Tool or demo papers.
• Position papers related to the topics mentioned above.
• Work in progress papers.

Submitted papers must be formatted according to the camera-ready style for ECAI 2020. Manuscripts must be submitted electronically via the easychair conference management system using the following link: https://easychair.org/conferences/?conf=epai2020

Authors of accepted papers will be asked to present the paper during the workshop. Online pre-proceedings containing all accepted papers will be prepared before the date of the conference. Depending on the number and quality of submissions, we will examine the possibility of targeting a volume or a journal special issue.

Papers are allowed a maximum of seven (7) pages, references excluded. References can take up to one additional page. Formatting Guidelines, LaTeX Styles and Word Template can be downloaded from http://ecai2020.eu/wp-content/uploads/AuthorsPack-ECAI2020.zip

IMPORTANT DATES

• Submission deadline: May 15th, 2020
• Notification of acceptance: June 12th, 2020
• Camera-ready: June 15th, 2020
• Workshop: August 30th, 2020
PROGRAM CHAIRS

- Emilia Gomez-Gutierrez - Joint Research Centre, European Commission
- Seán Ó hÉigeartaigh - Cambridge’s Centre for the Study of Existential Risk (CSER)
- Jose Hernandez-Orallo - Technical University of Valencia
- Fernando Martínez-Plumed - Technical University of Valencia
- Giuditta De Prato - Joint Research Centre, European Commission

PROGRAM COMMITTEE (confirmed)

- Jordi Bieger (CADIA, Reykjavik University)
- Botty Dimanov (University of Cambridge)
- Enrique Fernández Macías (European Commission’s Joint Research Centre)
- Peter Flach (University of Bristol - Alan Turing Institute)
- Ross Gruetzemacher (Auburn University)
- Aaron Li-Feng Han (Dublin City University)
- Annarosa Pesole (European Commission’s Joint Research Centre)
- Ricardo Prudêncio (Universidade Federal de Pernambuco)
- Ute Schmid (University of Bamberg)
- Charlotte Stix (Eindhoven University of Technology)
- Songül Tolan (European Commission's Joint Research Centre)
- Ricardo Vinuesa (KTH Royal Institute of Technology)
- Jess Whittlestone (University of Cambridge)

CONTACT PERSON

Fernando Martínez-Plumed (fmartinez@dsic.upv.es)