The Scientometrics of AI Benchmarks:
Unveiling the Underlying Mechanics of AI Research

Pablo Barredo¹ and José Hernández-Orallo²
and Fernando Martínez-Plumed²,³ and Seán Ó hÉigeartaigh⁴

Abstract

The widespread use of experimental benchmarks in AI research has created new competition and collaboration dynamics that are still poorly understood. In this paper we provide an innovative methodology to explore this dynamics and analyse the way different entrants in these competitions, from academia to tech giants, behave and react depending on their own or others’ achievements. We perform an analysis of over twenty popular benchmarks in AI, linking their underlying research papers. We identify links between researchers and institutions (i.e., communities) beyond the standard co-authorship relations, and we explore a series of hypotheses about their behaviour as well as some aggregated results in terms of activity, breakthroughs and efficiency. As a result, we detect and characterise the dynamics of research communities at different levels of abstraction, including organisation, affiliation, trajectories, results and activity.

Submission under review

¹ Universidad de Oviedo, email: UO237136@uniovi.es
² Universitat Politècnica de València, email: jorallo@upv.es, fmartinez@dsic.upv.es
³ JRC, European Commission, email: fernando.martinez-plumed@ec.europa.eu
⁴ University of Cambridge, email: so348@cam.ac.uk
ACKNOWLEDGEMENTS
This material is based upon work supported by the EU (FEDER), and the Spanish MINECO under grant RTI2018-094403-B-C3, the Generalitat Valenciana PROMETEO/2019/098. F. Martínez-Plumed acknowledges funding of the AI-Watch project by DG CONNECT and DG JRC of the European Commission. J. Hernández-Orallo and S. Ó hÉigeartaigh are also funded by an FLI grant RFP2-152. We thank the anonymous comments received by the EPAI committee.

REFERENCES