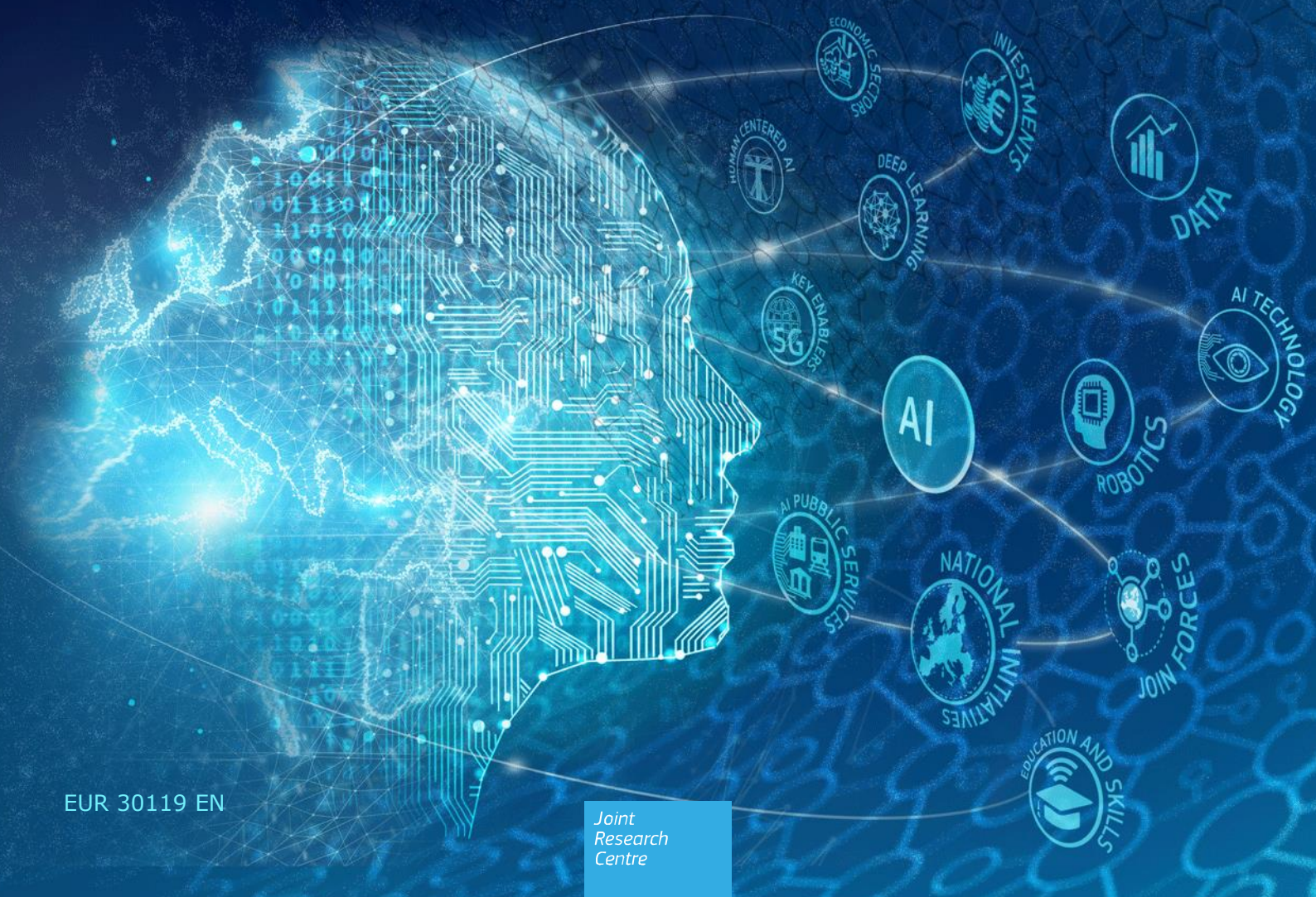
European Commission

JRC TECHNICAL REPORT

AI Watch
Methodology to Monitor the Evolution of AI Technologies



EUR 30119 EN

Joint Research Centre

This publication is a Technical report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication. For information on the methodology and quality underlying the data used in this publication for which the source is neither Eurostat nor other Commission services, users should contact the referenced source. The designations employed and the presentation of material on the maps do not imply the expression of any opinion whatsoever on the part of the European Union concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Contact information
Name: Fernando Martínez Plumed
Address: JRC Seville, Edificio Expo
Inca Garcilaso, 3
41092 Sevilla
Email: Fernando.MARTINEZ-PLUMED@ec.europa.eu

EU Science Hub
https://ec.europa.eu/jrc

JRC120090

EUR 30119 EN

PDF ISBN 978-92-76-17153-9 ISSN 1831-9424 doi:10.2760/643950

Luxembourg: Publications Office of the European Union, 2020

How to cite this report: Martínez-Plumed, F., Hernández-Orallo, J. and Gómez, E., *AI Watch: Methodology to Monitor the Evolution of AI Technologies*, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-17153-9, doi:10.2760/643950, JRC120090.

Contents

Authors

Fernando Martínez-Plumed, Joint Research Centre, European Commission

José Hernández-Orallo, Universitat Politècnica de València

Emilia Gómez, Joint Research Centre, European Commission

Abstract

In this report, we present a methodology to assess the evolution of AI technologies in the context of the AI WATCH initiative. The methodology is centred on building the *AIcollaboratory*, a data-driven framework to collect and explore data about AI results, progress and ultimately capabilities. From the collaboratory framework we later extract qualitative information related to the state of the art, challenges and trends of AI research and development.

This report first describes the administrative context of study, followed by the proposed methodology to build the *AIcollaboratory* framework and exploit it for qualitative assessment. In addition, we present some preliminary results of this monitoring process and some conclusions and suggestions for future work.

Foreword

This report is published in the context of AI WATCH, the European Commission knowledge service to monitor the development, uptake and impact of Artificial Intelligence (AI) for Europe, launched in December 2018.

AI has become an area of strategic importance with potential to be a key driver of economic development. AI also has a wide range of potential social implications. As part of its Digital Single Market Strategy, the European Commission put forward in April 2018 a European strategy on AI in its Communication "Artificial Intelligence for Europe" COM(2018)237. The aims of the European AI strategy announced in the communication are:

- To boost the EU's technological and industrial capacity and AI uptake across the economy, both by the private and public sectors
- To prepare for socio-economic changes brought about by AI
- To ensure an appropriate ethical and legal framework.

Subsequently, in December 2018, the European Commission and the Member States published a "Coordinated Plan on Artificial Intelligence", COM(2018)795, on the development of AI in the EU. The Coordinated Plan mentions the role of AI Watch to monitor its implementation.

AI WATCH monitors European Union's industrial, technological and research capacity in AI; AI-related policy initiatives in the Member States; uptake and technical developments of AI; and AI impact. AI WATCH has a European focus within the global landscape. In the context of AI Watch, the Commission works in coordination with Member States. AI WATCH results and analyses are published on the AI WATCH Portal[1].

From AI Watch in-depth analyses, we will be able to better understand EU's areas of strength and areas where investment is needed. AI Watch will provide an independent assessment of the impacts and benefits of AI on growth, jobs, education, and society.

AI Watch is developed by the Joint Research Centre (JRC) of the European Commission in collaboration with the Directorate-General for Communications Networks, Content and Technology (DG CONNECT).

This report addresses the following objectives of AI WATCH: Monitor the evolution of AI technologies. As part of this objective this report particularly aims to develop a methodology to monitor the state of the art, evolution, benchmarking capabilities and future challenges in AI.

[1] https://ec.europa.eu/knowledge4policy/ai-watch_en

1 Introduction

1.1 Motivation

AI has become an area of strategic importance with potential to be a key driver of economic development and with a wide range of potential social implications. The need for assessing and understanding what AI is capable of is becoming more relevant as academia and technology companies in AI are rushing to achieve breakthroughs for particular benchmarks, specific problems or narrow tasks, usually at the expense of massive data, computation power, embedded heuristics, strong bias, system specialisation, etc. Several national AI strategies are accompanied by catalogues about what their research centres and AI companies claim they are able to do. The study of the impact of AI on the workplace is usually done in terms of some assumptions of what AI or ML is able to do too (Martínez-Plumed et al, 2020, Tolan et al., 2020). Similarly, the analysis of AI risks, both short-term and long-term, depends on the assessment of these capabilities.

But, what is AI capable of? This question is as crucial as elusive, and the answer becomes more difficult as AI progresses in ways that are open-ended about the techniques and resources AI can operate with. The truth is that whenever a computational task is solved, researchers find increasingly challenging to extrapolate whether this task can be reproduced, even when only a few things change: the data, the domain knowledge, the level of noise or uncertainty, the (hyper)parameters, the techniques themselves, the research team, the libraries, the compute, etc.

Benchmarks, competitions and other kinds of challenges are behind much of the recent progress in AI, especially in machine learning (ML) (Hernández-Orallo et al., 2017), but the dynamics of rushing breakthroughs at the expense of massive data, compute, specialisation, etc., has led to a more complex AI landscape, in terms of what can be achieved and how. As a result, policy makers and other stakeholders have no way of assessing what AI systems can do today and in the future.

This does not mean that we must disregard or understate the valuable information that is provided by a plethora of benchmarks. On the contrary, the analysis of the progress of AI must be based on data-grounded evidence, relying on finding and testing hypotheses through the computational analysis of big amounts of shared data (Gewin 2016), using open data science tools (Lowndes et al., 2017). We need to assess whether new AI systems and techniques are simply an incremental improvement for a narrow collection of applications or a real breakthrough representing a more general cognitive ability, which can be established in comparison with the same abilities in humans and other animals. But this analysis must be abstracted from tasks to capabilities, for the purposes of integration[2] and evaluation (Hernández-Orallo, 2017).

In this report, we present a methodology to assess the evolution of AI technologies in the context of AI WATCH, based on the *AIcollaboratory*, a data-driven framework to collect and explore data about AI results, progress and ultimately capabilities. The report first describes the administrative context of study, followed by the proposed methodology to monitor AI evaluation initiatives and use them to provide qualitative assessment. In addition, we present some preliminary results of this monitoring process and some conclusions and suggestions for future work.

1.2 Definitions

We introduce here some relevant definitions in the context of this report.

- *Artificial intelligence (AI)*: we consider the definition proposed in AI WATCH as the basis for our work and Machine Learning (ML) as a particularly relevant subset of AI. For further details on the scope of AI considered in AI WATCH we refer to (Samoili et al., 2019). "Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans[3] that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions".
- *AI system*: a computer system (e.g., algorithm, application, computational model, etc.) capable of performing tasks that involve perception, communication, reasoning, learning, planning and other

[2] CCC AI roadmap: https://cra.org/ccc/ai-roadmap-integrated-intelligence/

[3] Humans design AI systems directly, but they may also use AI techniques to optimise their design. Footnote from original document.

perceptual or cognitive capabilities. AI systems are usually powered by machine learning, natural language processing, machine vision, deep learning, rule inference, etc. and applied to a wide range of applications such as personal assistants, self-driving cars, robotic cleaner or automated translators.

- *AI task*: in computing, a task is a unit of execution or a unit of work[4]. AI tasks are the computing units used for training and evaluating AI systems, and AI systems are built to perform and automate them. Illustrative examples of types of AI tasks involve, among others, facial recognition in images, speech synthesis, relation extraction, question answering or machine translation.
- *AI evaluation*: AI evaluation means making observations of all aspects of AI, for the purpose of comparison. The assessment of AI systems, algorithms and applications may allow us to find their strengths and weaknesses and improve them where necessary. However, AI evaluation should go beyond measuring the performance of an AI system (e.g. the final result that is attained on a task at a single point in time) and embrace other aspects such as the (internal and external) costs of development and deployment (such as data, computations, software, hardware, time) or aspects such as fairness or transparency of the system. These are usually neglected dimensions in the AI literature. In addition, it would be misleading to consider that the progress in AI is represented by the progress of specific AI systems solving specific tasks, benchmarks or challenges. We need to understand the relationships between different tasks and integrate results from multiple systems addressing related AI tasks where actual measurements can be aggregated and combined (e.g., pareto-optimal analysis for assessing performance and costs) into abilities rather than performance. Finally, it is crucial to analyse the state and progress of the area based on data-grounded evidence through the computational analysis of large amounts of AI system results using appropriate tools. There is also a need to assess whether new AI systems and techniques are simply an incremental improvement for a narrow collection of applications or a real breakthrough representing a more general cognitive ability.
- *Benchmarking initiative*: There is an increasing need for a better assessment of the capacities and limitations of AI systems. Theoretical approaches might provide important insights, but only through experimentation and evaluation tools will we achieve a more accurate assessment of how an actual system operates over a series of tasks or environments. This is why several AI experimentation and benchmarking initiatives are now ubiquitous for the evaluation of AI systems. In short, an AI benchmark may be defined as an (open source) platform for evaluating AI performance over a set of (representative) AI tasks from a particular domain (e.g., computer vision, speech recognition or question answering).

[4] https://en.wikipedia.org/wiki/Task_(computing)

2 Proposed methodology

2.1 Monitoring of AI evaluation initiatives

The first stage of AI WATCH proposed methodology consists in the identification and gathering of public information related to AI evaluation. This information is included in scientific publications (e.g. through thematic/academic search engines such as Scinapse[5], AItopics[6] or arXiv[7]), open code and data, and international evaluation initiatives in major AI domains such as computer vision, deep learning, audio and speech recognition, planning and scheduling, evolutionary methods, natural language processing, reinforcement learning and cognitive robotics.

In order to do so, we will periodically monitor benchmark and competition repositories. For each selected repository, we will monitor new tasks, and for each particular task we will study participants, algorithms and changes in metrics. From this monitoring, we will compute statistics per year (e.g. area-specific performance measures/indicators as well as external/internal costs whenever provided) and quantify the activity on the different areas and tasks of AI research (e.g. number of publications and results). Initially, we consider a comprehensive set of open repositories which draw data from multiple (well-established) sources, including academic literature, review articles and code platforms focused on machine learning and AI. More specifically, the main sources of information for AI WATCH may include:

- Papers with Code is the largest, up to date, free and open repository of machine learning papers, which includes data from several sources:
    - NLP-Progress: A hand-annotated repository to track the progress in Natural Language Processing (NLP), including the datasets and the current state-of-the-art for the most common NLP tasks.
    - SQuAD: The Stanford Question Answering Dataset (SQuAD) is an active project for evaluating reading comprehension and question answering which provide up-to-date leader boards.
    - RedditSota: This repository provides state-of-the-art (SoTA) results for a variety of tasks across machine learning problems.
    - SNLI: The Stanford Natural Language Inference (SNLI) Corpus is an active project for Natural Language Inference which reports and catalogues 3-class classification results on the SNLI test set.
    - Cityscapes: This benchmark for semantic segmentation also provides an evaluation server such that authors can upload their results and get a ranking regarding the different tasks.
- EFF: a repository from the Electronic Frontier Foundation that collects problems and metrics from a subset of representative tasks from the AI and machine learning research literature, and aims at tracking progress on them.
- OpenML is an online machine learning platform for sharing and organising data, machine learning algorithms and experiments. Its goal is to make machine learning and data analysis simple, accessible, collaborative and open where people can upload and share data sets and prediction tasks that they then collaboratively solve using machine learning algorithms.

We may also gather data from further repositories to have a more comprehensive picture of the AI landscape. Some examples may include:

- Swedish Chess Computer Association
- Are we there yet?
- Algorithmic Progress in Six Domains
- wer_are_we
- Animal-AI olympics
- ACL

[5] https://scinapse.io/
[6] https://aitopics.org/
[7] https://arxiv.org/

We rely on interfaces and APIs (when available) or web-scraping techniques in order to gather the data from the above repositories. In this regard, Papers with code provides their data (e.g., papers with abstracts, links to code repositories and evaluation tables) under a *Creative Commons* (CC-BY-SA-4[8]) licence, and therefore enabled us to copy, redistribute, remix, transform, and build upon it. To access the data, we rely on the raw files (json) the platform provides[9], which are updated once a week. For its part, EFF provides a Jupyter/IPython notebook[10] hosted in Github where the community can both contribute and gather the already collected data. Finally, OpenML offers interfaces in multiple programming languages which allow scientists to interact with the server using language-specific functions. Note that, once retrieved, all the data from the selected repositories should be then cleaned, transformed and structured to fit *AIcollaboratory* multidimensional data model, which will be explained in subsequent sections.

Therefore, from the aforementioned sources we can track the reported evaluation results (when available or sufficient data is provided) on different metrics of AI performance across separate AI and machine learning benchmarks (e.g., datasets, competitions, awards, etc.). This is possible from a number of AI domains, including (among others) computer vision, speech recognition, music analysis, machine translation, text summarisation, information retrieval, robotic navigation and interaction, automated vehicles, game playing, prediction, estimation, planning, automated deduction, etc. This ensures a broad coverage of AI tasks.

Expected results: *AIcollaboratory* repository (annual update)

2.2 Qualitative assessment

From the information gathered as explained in section 2.1, we will provide an expert assessment to:

- Analyse AI trends and challenges. We will provide a critical analysis of the current state of the art, trends and challenges based on new paradigms, techniques, tasks, and achievements.
- Measure technological readiness and disruption of AI by mapping previous data to technology readiness level (TRL).
- Detect breakthroughs and paradigms change.

Expected results: AI evolution report (annual update)

2.3 Community building

In order to integrate multiple views into the qualitative assessment exercise, we propose to engage with different stakeholders, including the AI research community.

This initiative includes a strategic collaboration with Universidad Politécnica de Valencia, in particular with the research group led by Prof. José Hernández-Orallo. There is a need to interact more broadly with the EU AI community researching dealing with AI progress and evaluation, e.g. CLAIRE[11] and ELLIS[12] networks.

More precisely we propose the organisation of 2 annual events in the context of this work:

1. **Research workshop**: Location: satellite event to major AI conference. AI WATCH plans to organise a yearly workshop on AI evaluation where we will gather insight from researchers on the evolution and future of AI technologies.
    - June, 2020. Presentation of highlight paper (Martínez-Plumed, Gómez and Hernández-Orallo, 2019) and workshop on Evaluating AI Progress (EPAI 2020[13]), 24th European Conference on AI (ECAI 2020[14]) what will take place in Santiago de Compostela, Spain in June 8-12 2020. Our aim is to bring together players from different fields (inside and outside IA) to discuss, in a workshop atmosphere, the different views on:
        - How evaluation in artificial intelligence is done.

[8] https://creativecommons.org/licenses/by-sa/4.0/
[9] https://github.com/paperswithcode/paperswithcode-data
[10] https://github.com/AI-metrics/AI-metrics
[11] https://claire-ai.org/
[12] https://ellis.eu/
[13] http://dmip.webs.upv.es/EPAI2020/
[14] http://ecai2020.eu/

- How it relates to issues such as generality, replicability, reusability, and ethical issues such as privacy, safety, security, fairness, discrimination, accountability, etc.
- How all these parameters can help assess (and ultimately drive) progress in AI.

In this regard, we have proposed a fully multi-disciplinary, transversal, workshop, sustained by a diverse conjunction of organising and programme committees that cover a wide range of backgrounds and expertise in these issues, and an illustrious line-up of invited speakers, who will surely attract a good number of participants from the main communities of the joint conference.

2. **AIcollaboratory workshop**. Location: JRC, autumn 2020. AI WATCH plans to organise a yearly workshop to engage with researchers from different AI domains and discuss about the quantitative and qualitative assessment of AI progress via the *AIcollaboratory*. The goals of this yearly workshop are then to draw qualitative conclusions from the *AIcollaboratory* data and complementary information sources and define future developments in the methodology. It will be followed by an update of the *AIcollaboratory* and the publication of an annual AI progress report.

   This workshop will include technical sessions and will devote an important share of time to discussion and interaction with the *AIcollaboratory*, with a more open discussion about the research challenges around the topic, improvements, future events and initiatives, etc. In this regard, the technical sessions will be devoted to answer questions such as:

   - How should the *AIcollaboratory* be used to properly evaluate, compare and classify AI systems, and thus guide the future of the field?
   - How should it contribute to the analysis of AI progress?
   - How should it track, collate and visualise data relating to AI?
   - How should it help policy makers to assess what AI systems can do, what they will be able to do in the near future, and how the field may get there?
   - How can the interface be improved, or predefined queries and visualisations be created so that everyone can use it and contribute?
   - How can we accurately assess the technological feasibility of automation through the deployment of AI?

   As a result of this workshop, (1) the *AIcollaboratory* would be significantly improved by enhanced feedback from the participants, making it more appropriate for general use, and (2) outreach work plans would be developed to facilitate cooperation and communication between all interested parties.

In addition, the AI WATCH team will attend major AI conferences to share research results and follow the state of the art on current AI trends and challenges.

Expected results: annual *AIcollaboratory* workshop at JRC and annual workshop on AI evaluation at a top AI scientific conference.

3 Preliminary results

3.1 AIcollaboratory

The *AICollaboratory*[15] aims to develop a collaborative initiative for the analysis, evaluation, comparison and classification of AI systems. It is conceived as an open data repository that incorporates different sources of information on methodologies used to characterise all kinds of AI systems, hybrids and collectives thereof. The first prototype of the collaboratory[16] incorporates approaches, data, measurements and evaluations. It provides a way to study, analyse and evaluate, in a complete and unified way, a representative selection of different AI systems, covering, in the long term, the current and future artificial intelligence landscape.

The A*Icollaboratory* aims thus at integrating open data and knowledge in three domains: an inventory of intelligent systems, a behavioural test catalogue and the measurement experiments obtained when systems take tests:

- *Inventory of intelligent systems* (agents): which incorporates information about current, past and future AI systems, including their intrinsic characteristics whenever possible and the hierarchies these agents may be classified in. In this regard, they will be aggregated (e.g., classified into hierarchies) from individualsgroups or organisations, with populations and distributions over them.
- *Testing Catalogue* (instruments): which will integrate a series of behavioural tests such as benchmarks, competitions and related tasks, the dimensions they measure and for which kinds of systems, the possible interfaces and testing apparatus.
- *Repository of experimentation* (results): which will record the results (measurements) of a wide range of AI systems for several benchmarks and tasks, as the main data source of the AI Collaboratory? Data is contributed from scientific papers, experiments, code repositories, AI/robotic competitions, etc., as pointed in section 2.1.

These above three components—roughly corresponding to subjects, instruments and results— have an open and collaborative character, encouraging researchers to be users but also contributors. The current data repository currently covers (as per November 25th 2019) 1800 agents, 581 tasks and 6720 results.

In addition to the infrastructure, there is a series of exploitation tools where actual data science will take place, following state-of-the-art data exploitation tools but customised to the potential users of the *AIcollaboratory* (see Figures 1 and 2):

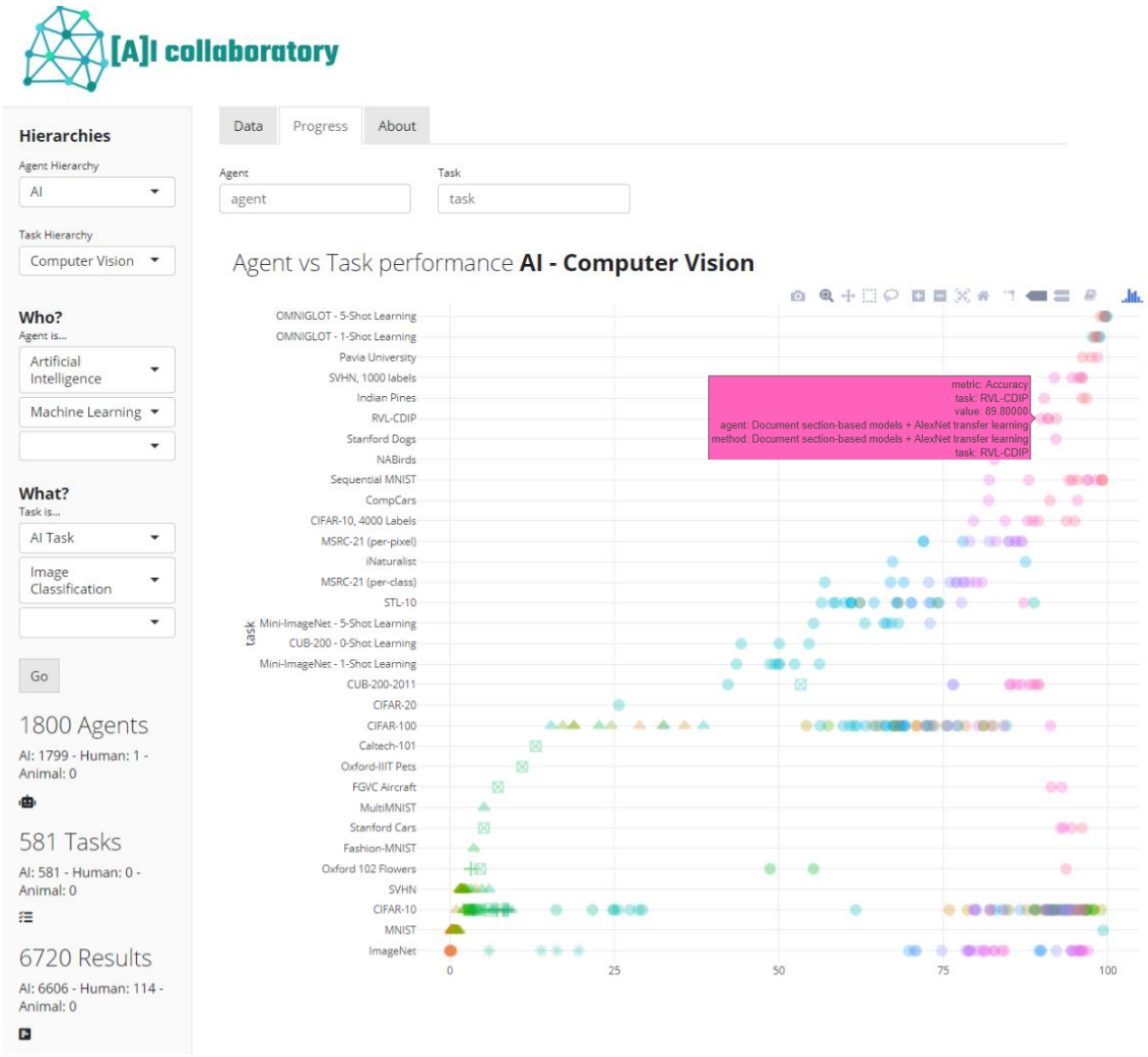
- *Querying tools*: interfaces for powerful (dis)aggregation and cross-comparison, reuse of predefined multidimensional filters, trend analysis, heuristic search, etc.
- *Visualisation tools*: a set of interactive interfaces to perform projections, visual categorisations, progress tracking plots, etc.

The final objective of the *AICollaboratory* is, therefore, to produce innovative basic research at the core of the science of intelligence that will contribute to a richer understanding of (artificial) intelligence, and a better steering of AI progress.

[15] http://www.aicollaboratory.org

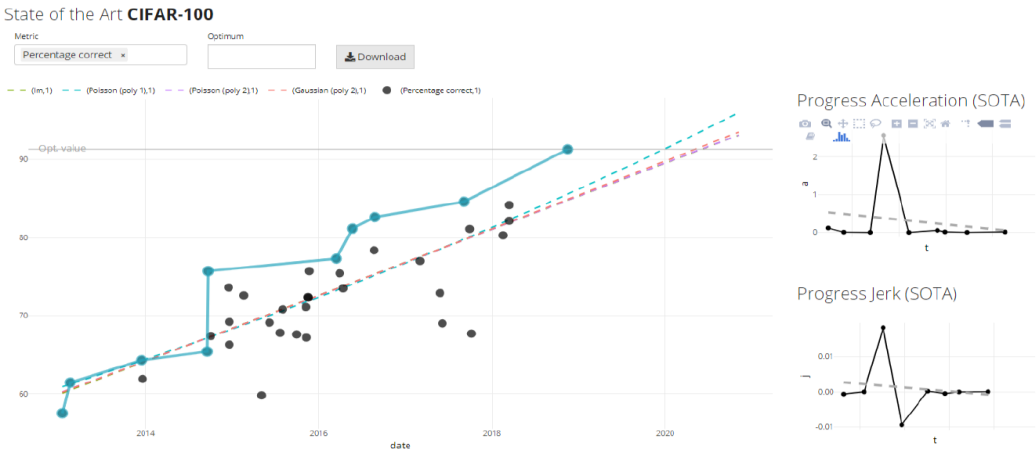
[16] We borrow the term *collaboratory* from the concept of an integrated framework to support scientific collaboration https://en.wikipedia.org/wiki/Collaboratory

Figure 1. Screenshot of the AICollaboratory: AI systems vs task performance. On the left part, the users filter the different dimensions and hierarchies. The plot shows the results (using different metrics) obtained by different AI systems addressing different benchmarks (e.g., CIFAR, OMNIGLOT, MNIST, etc.) involving the selected task (i.e., image classification).



Source: AIcollaboratory, 2020.

Figure 2. Screenshot of the *AICollaboratory*: Task progress. The user can select any particular benchmark (for a specific task) to analyse their progress over time with an overview of the best systems and the pareto frontier.



Results **CIFAR-100**

	agent	method	value	metric	paper
1	GPIPE	GPIPE	91.3	Percentage correct	GPipe: Efficient Training of Giant Neural Networks using Pipeline Parallelism
2	SENet + ShakeEven + Cutout	SENet + ShakeEven + Cutout	84.59	Percentage correct	Squeeze-and-Excitation Networks
3	PyramidNet-272 + SWA	PyramidNet-272 + SWA	84.16	Percentage correct	Averaging Weights Leads to Wider Optima and Better Generalization
4	DenseNet	DenseNet	82.62	Percentage correct	Densely Connected Convolutional Networks
5	WRN+SWA	WRN+SWA	82.15	Percentage correct	Averaging Weights Leads to Wider Optima and Better Generalization
6	Wide ResNet	Wide ResNet	81.15	Percentage correct	Wide Residual Networks
7	CoPaNet-R-164	CoPaNet-R-164	81.1	Percentage correct	Deep Competitive Pathway Networks
8	SimpleNetv2	SimpleNetv2	80.29	Percentage correct	Towards Principled Design of Deep Convolutional Networks: Introducing SimpNet
9	SimpleNetv1	SimpleNetv1	79.37	Percentage correct	Lets keep it simple, Using simple architectures to outperform deeper and more complex architectures
10	ResNet-1001	ResNet-1001	77.3	Percentage correct	Identity Mappings in Deep Residual Networks

Source: AIcollaboratory, 2020.

4 Conclusions

In this report we present the methodology followed to build the *AIcollaboratory*, a data-driven framework developed in the context of the AI WATCH initiative to collect and explore data about AI results, progress and ultimately capabilities. The main goal of the *AIcollaboratory* is to extract quantitative information related to the state of the art, challenges and trends of AI research and development and facilitate further qualitative analysis.

One of the main challenges of the *AICollaboratory* is to maximize engagement by the AI community. Many initiatives do not get enough inertia, funding or popularity and are soon discontinued. We plan to address this in different ways: by co-operating with some other similar initiatives, having strategic collaborations with AI research centres, and establishing interactions with CLAIRE and ELLIS networks of EU researchers on AI by means of annual scientific events.

Ultimately, the *AICollaboratory* aims to be a networked open ecosystem allowing people all over the world to collaborate and build directly on each other's latest data, knowledge and results, also providing provide important benefits for the scientific community and policymakers, as well as produce innovative basic research at the core of the science of intelligence.

5 References

Gewin, V. 'Data sharing: An open mind on open data', Nature, 529 (7584), 117–119, (2016).

Hernández-Orallo, J., 'Evaluation in AI: from task-oriented to ability-oriented measurement',AI Review,48(3), 397–447, (2017).

Hernández-Orallo, J. Baroni, M., Bieger, J., Chmait, N., Dowe, D., Hofmann, K., Martínez-Plumed, F., Strannegård, C., Thórisson, K., "A New AI Evaluation Cosmos: Ready to Play the Game?", AI Magazine, Association for the Advancement of Artificial Intelligence, Vol. 38, No 3, Fall 2017.

Lowndes, J. S., et al., 'Our path to better science in less time using open data science tools', Nature ecology & evolution, 1(6), 0160, (2017).

Martínez-Plumed, F., Hernández-Orallo, J., Gómez, E. 'Tracking AI: The Capability is (Not) Near'. 24th European Conference on Artificial Intelligence (ECAI 2020), Santiago de Compostela, 8-12 June, 2020

Martínez-Plumed, F., Tolan, S., Pesole, A., Hernández-Orallo, J., Fernández-Macias, E., Gómez, E.: "Does AI Qualify for the Job? A Biderectional Model Mapping Labour and AI intensities", Proceedings of the AAAI/ACM conference on Artificial Intelligence, Ethics and Society (AIES 2020), New York, USA, February 7-8, 2020.

Samoili, S., López-Cobo, M., Gómez, E., De Prato, G., Martínez-Plumed, F. and Delipetrev, B., AI WATCH. 'Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence', EUR, Publications Office of the European Union, Luxembourg, 2019, JRC118163.

Tolan, S, Pesole, A., Martínez-Plumed, F., Hernández-Orallo, J., Fernández-Macias, E., Gómez, E. 'Measuring the occupational impact of AI beyond automation: tasks, cognitive abilities and AI benchmarks', Submitted, (2020).

List of figures

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications from EU Bookshop at: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

KJ-NA-30119-EN-N

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub

ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

EU Science, Research and Innovation

EU Science Hub

Publications Office of the European Union

doi:10.2760/643950

ISBN 978-92-76-17153-9