

Vision-and-Language Algorithmic Reasoning (VLAR 2023)  
Workshop and Challenge

October 3, 2023, Paris, France

Held in conjunction with ICCV 2023

<https://wvlar.github.io/iccv23>

## CALL FOR CONTRIBUTIONS

The focus of this workshop is to bring together researchers in multimodal reasoning and cognitive models of intelligence towards positioning the current research progress in AI within the overarching goal of achieving machine intelligence. An important aspect is to bring to the forefront problems in perception, language modeling, and cognition that are often overlooked in state-of-the-art research and that are important for making true progress in artificial intelligence. One specific problem that motivated our workshop is the question of how well current deep models learn broad yet simple skills and how well do they generalize their learned models to solve problems that are not part of their learning set; such skills even children learn and use effortlessly (e.g., see the paper “[Are Deep Neural Networks SMARTer than Second Graders?](#)”). In this workshop, we plan to bring together outstanding researchers to showcase their cutting edge research on the above topics that will inspire the audience to bring out the missing pieces in our quest to solve the puzzle of artificial intelligence.

---

### IMPORTANT DATES

#### \* Paper Track

**Submission deadline: \*\*\*July 24, 2023\*\*\*** (11:59PM EDT) (extended from July 20)

Paper decisions to authors: August 7, 2023

Camera-ready deadline: August 18, 2023 (11:59PM EDT)

#### \* SMART-101 Challenge Track

Challenge open: June 15, 2023.

Submission deadline: **\*\*\*September 15, 2023\*\*\*** (11:59PM EDT) (extended from September 1).

Arxiv paper deadline to be considered for awards: September 1, 2023 (11:59PM EDT).

Public winner announcement: October 3, 2023 (11:59PM EDT).

---

## TOPICS FOR PAPER TRACK

We invite submissions of original and high-quality research papers in the topics related to vision-and-language algorithmic reasoning. The topics for VLAR 2023 include, but are not limited to:

- \* Large language models, vision, and cognition including children's cognition
- \* Foundation models of intelligence, including vision, language, and other modalities
- \* Artificial general intelligence / general-purpose problem solving architectures
- \* Neural architectures for solving vision & language or language-based IQ puzzles
- \* Embodiment and AI
- \* Large language models, neuroscience, and vision
- \* Functional and algorithmic / procedural learning in vision
- \* Abstract visual-language reasoning, e.g., using sketches, diagrams, etc.
- \* Perceptual reasoning and decision making
- \* Multimodal cognition and learning
- \* New vision-and-language abstract reasoning tasks and datasets
- \* Vision-and-language applications

---

## SUBMISSION INSTRUCTIONS FOR PAPER TRACK

- \* We are inviting only original and previously unpublished work. Dual submissions are not allowed.
- \* All submissions are handled via the workshop's CMT Website:  
<https://cmt3.research.microsoft.com/VLAR2023>.
- \* Submissions should not exceed **four (4) pages** in length (excluding references).
- \* Submissions should be made in PDF format and should follow the official ICCV template and guidelines.
- \* All submissions should maintain author anonymity and should abide by the ICCV conference guidelines for double-blind review.
- \* Accepted papers will be presented as either an oral, spotlight, or poster presentation. At least one author of each accepted submission must present the paper at the workshop.
- \* Presentation of accepted papers at our workshop will follow the same policy as that for accepted papers at the ICCV main conference
- \* Accepted papers will also be part of the ICCV 2023 workshop proceedings.
- \* Authors may optionally upload supplementary materials, the deadline for which is the same as that of the main paper and should be submitted separately.

---

## INSTRUCTIONS FOR PARTICIPATING IN THE SMART-101 CHALLENGE TRACK

As part of VLAR 2023, we are hosting a challenge based on the Simple Multimodal Algorithmic Reasoning Task – SMART-101 – dataset, which is available for download here:

<https://smartdataset.github.io/smart101/>. The accompanying CVPR 2023 paper “Are Deep

Neural Networks SMARTer than Second Graders” is available here:

<https://arxiv.org/abs/2212.09993>.

\* The challenge is hosted on Eval AI and is open to submissions:

<https://eval.ai/web/challenges/challenge-page/2088/overview>

\* The challenge participants are required to make arXiv submissions detailing their approach.

These are only used to judge the competition, and **will not be reviewed** and will not be part of workshop proceedings.

\* Winners of the challenge are determined both by performance on the leaderboard over a private test set as well as the novelty of the proposed method (as detailed in the arXiv submission). Details are made available on the challenge website.

\* Prizes will be awarded on the day of the workshop.

---

## KEYNOTE SPEAKERS

[Prof. Anima Anandkumar](#), NVIDIA & Caltech

[Dr. François Chollet](#), Google

[Prof. Jitendra Malik](#), Meta & UC Berkeley

[Prof. Elizabeth Spelke](#), Harvard University

[Prof. Jiajun Wu](#), Stanford University

---

## WORKSHOP ORGANIZERS

[Anoop Cherian](#), Mitsubishi Electric Research Laboratories

[Kuan-Chuan Peng](#), Mitsubishi Electric Research Laboratories

[Suhas Lohit](#), Mitsubishi Electric Research Laboratories

[Kevin A. Smith](#), Massachusetts Institute of Technology

[Ram Ramrakhya](#), Georgia Institute of Technology

[Honglu Zhou](#), NEC Laboratories America, Inc.

[Tim K. Marks](#), Mitsubishi Electric Research Laboratories

[Joanna Matthiesen](#), Math Kangaroo USA

[Joshua B. Tenenbaum](#), Massachusetts Institute of Technology

---

## CONTACT

Email: [vlariccv23@googlegroups.com](mailto:vlariccv23@googlegroups.com)

SMART-101 project: <https://smartdataset.github.io/smart101/>

Website: <https://wvlar.github.io/iccv23>