



Carnegie Mellon University
Language Technologies Institute



PETER CLARK

ALLEN INSTITUTE FOR AI

Peter Clark is a Senior Research Manager and founding member of AI2, and has led Project Aristo since its inception in 2014. His research focuses upon natural language processing, machine inference, and commonsense reasoning, and the interplay between these three areas.

FROM 'F' TO 'A' ON THE N.Y. REGENTS SCIENCE EXAM: A LOOK INSIDE THE ARISTO SYSTEM

Performing well on standardized exams has been a longstanding challenge for AI. Even in 2016, the best AI system achieved less than 60% on an 8th Grade science exam challenge. Recently, AI2's Aristo system achieved surprising success on the Grade 8 New York Regents Science Exams, scoring over 90% on the exam's non-diagram, multiple choice (NDMC) questions. How was it able to do this? How much reasoning is going on inside? And what mistakes does it still make? In this talk, I will overview Aristo and the impact of its various components, in particular its new language model (LM) solvers. I will present several analyses of what is going on inside Aristo, in particular probing how much the LM solvers go beyond simple pattern matching, and what kinds of errors still occur. Finally I will speculate on the larger quest towards knowledgeable machines that can reason, explain, and interact, and what additional capabilities are needed to reach this broader goal.

FRIDAY, MAR 27, 2:30 - 4 PM

CONDUCTED VIA ZOOM

[HTTPS://CMU.ZOOM.US/J/208848796](https://cmu.zoom.us/j/208848796)