

Peter Clark, PhD

peterc@allenai.org, <https://pclark425.github.io/>

AREAS OF EXPERTISE

- Automated Scientific Discovery
- Natural Language Processing
- Machine Reasoning
- Commonsense Reasoning
- Knowledge Acquisition
- Large Language Models
- Machine Learning

CAREER PROGRESSION

Interim CEO 2022-2023

Allen Institute for AI (AI2)

Senior Research Director 2013-now

Allen Institute for AI (AI2)

Senior Research Scientist 2010-2013
Vulcan Inc.

Associate Technical Fellow 2004-2010
Boeing Research & Technology

Research Scientist 1996-2004
Boeing Research & Technology

Research Fellow 1994-1996
Computer Science, Univ Texas at Austin

Research Associate 1992-1994
National Research Council, Canada

Postdoctoral Fellow 1991-1992
University of Ottawa, Canada

Research Scientist 1985-1991
Turing Institute, Glasgow, UK

EDUCATION

PhD Computer Science 1991
Strathclyde Univ., Glasgow, UK

**Masters in Information Technology:
Knowledge-Based Systems** 1985
University of Edinburgh, UK

BA in Physics 1984
First Class Honors (top 8% of graduates)
Oxford University, UK

PROFILE

Internationally recognized expert in artificial intelligence, in particular in the areas of automated scientific discovery, knowledge representation and reasoning, commonsense reasoning, and natural language understanding, with 35 years of experience and over 250 refereed publications in the field (28k citations, h-index 68).

SELECTED ACHIEVEMENTS

- **Significant, impactful research** in the field of artificial intelligence as reflected in: over 250 refereed publications in the area, including 5 Best Paper awards (3 x EMNLP (2024, 2023, 2014), AKBC (2014), AAAI (1997), multiple conference and workshop keynotes (e.g., ACL, NeurIPS, AIED, ESWC, CLEF, EMNLP), co-chair for the International Conference on Knowledge Capture (2005); tutorial co-chair for the NAACL Computational Linguistics Conference (2010); Three Boeing Breakthrough Recognition awards (2006,7,9); appointment as Boeing Associate Technical Fellow (2004); AAAI Senior Member (2014).
- **Co-founder of the Allen Institute of AI** (2012-2013). With others at Vulcan, planned, hired for, and launched the Allen Institute for AI.
- **Interim CEO for AI2 (2022-23)**, successfully steering the 200+ organization through change in leadership and the start of the ChatGPT revolution, including initiating AI2's work on the OLMo Open Language Model (now one of the foremost open models).
- Co-lead of [project Asta](#), a 40 person team developing an agentic ecosystem for **advancing scientific discovery** through all stages of research (literature analysis, hypothesis generation, planning experimentation, data analysis, and reporting), with particular focus on automated scientific discovery (e.g., [CodeScientist](#)).
- Led **Project Aristo**, the flagship project of the Allen Institute of AI, resulting in the [first system to score of 90%](#) on the New York Regents Science Exam (8th Grade multiple choice) in 2018. This milestone gained significant attention in the community and made the headlines of the [New York Times business section](#).
- Developed a new, influential line of research in training language models for formal reasoning, developing systems [RuleTaker](#) and [ProofWriter](#) (700+ citations). RuleTaker was the first system to show that language models were capable of general logical reasoning.
- Extensive contributions on **understanding and tracking belief** in modern neural systems, including [identifying and resolving internal inconsistencies](#), and enabling continual learning through a dynamic external memory of updates (both from the [environment](#) and [user](#)).
- Created a new research line on **representing and reasoning about processes**, including an [EMNP'14 best paper](#).
- Directly involved with, and major contributions to, **Vulcan's Project Halo** since its inception in 2002. From 2002-2010, major contributions to the AURA knowledge acquisition and reasoning system, in particular co-creating the original vision of a system able to acquire knowledge via graphical tools.
- Created vision for and developed Boeing's **Technical Expert Locator**, an AI application for finding expertise within the company. This tool was deployed and in full commercial use from 2001-2010.
- Developed a [theory of concept composition](#) with Prof. Bruce Porter. This work won a AAAI'97 Best Paper award, and was later a cornerstone of the Shaken knowledge acquisition system.
- Co-developed a **new machine learning algorithm**, [CN2](#) (1987). CN2 is still extensively used in the ML community (3500+ citations)

WORK HISTORY

Interim CEO, Allen Institute for AI

2022-2023

- Successfully steering the 200+ organization through change in leadership and the start of the ChatGPT revolution, including initiating AI2's work on the OLMo Open Language Model (now one of the foremost open models).

Senior Research Director, Allen Institute for AI

2013-present

- Research Lead and Project Co-Director, Project Asta (2024-present)**

Co-leading a 40 person team developing [Asta](#), an **agentic ecosystem for advancing scientific discovery** through all stages of research (literature analysis, hypothesis generation, planning experimentation, data analysis, and reporting). Significant research artifacts include:

- [The Asta system](#), an agentic research assistant for scientists
- [CodeScientist](#), for autonomous scientific discovery
- [AstaBench](#), the first comprehensive benchmark testing all stages of the discovery process

- Research Lead and Project Director, Project Aristo (2013-present)**

Lead of a 15 person research team developing solutions for natural language processing, machine reasoning, and commonsense knowledge. Highlights include:

- Developing the first AI system to [ace an 8th grade science exam](#) (scoring over 90%, 2018), using a combination of neural and structured reasoning approaches.
- Developed a new line, influential of research on **using language models for formal reasoning**, including the systems [RuleTaker](#) and [ProofWriter](#) (700+ citations). RuleTaker was the first system to show that language models were capable of general logical reasoning when trained appropriately.
- Created a new research line on **representing and reasoning about processes**, including an EMNLP'14 best paper ("[Modeling biological processes for reading comprehension](#)") and a sequence of follow-up works developing datasets and systems (e.g., "[Tracking State Changes in Procedural Text](#)", "[Reasoning about actions and state changes by injecting commonsense knowledge](#)")
- Novel work on explicit articulation and representation of **the underlying beliefs of language models**, enabling both more consistent reasoning and explanation by modern neural systems. ([Language Models with Rationality](#), [BeliefBank](#)).
- Developed modern approaches to **reasoning with text**, including release of [EntailmentBank](#), a large-scale dataset of multistep entailment trees. EntailmentBank has been used extensively to train new entailment reasoning systems.
- Development of new methods for continual learning by AI agents ([CLIN](#), [Skill Set Optimization](#))
- Impactful contributions of evaluation datasets, including ARC, QuaRTz, SciTail, QasC, EntailmentBank, OpenbookQA, and AstaBench. The [ARC Reasoning Challenge](#) (3000+ citations), now used as a standard evaluation component in many modern evaluation suites.
- Novel methods allowing neural systems to learn at inference time by managing a continuous, dynamic external memory, e.g., [MemPrompt](#), and [Towards Teachable Reasoning Systems](#) where users can correct system errors via a persistent memory.

Senior Research Scientist, Vulcan Inc

2010-2013

- Co-creator of the **Allen Institute of AI** (2012-2013). With others at Vulcan, planned, hired for, and launched the Allen Institute for AI.
- AURA project (2010-2013)**
Designed and developed the **Suggested Question mechanism** for AURA, based on an early prototype from SRI, allowing AURA to propose semantically similar, answerable questions when it was unable to handle a user's question directly. This mechanism was critical for AURA's robustness and usability in trials with students in 2011.
- General (2010-present)**
With other team members, continuous involvement in defining, critiquing, and planning the evolution of the Halo Program. Substantial technical involvement in other aspects of Halo, including the "hybrid" prototype.

Peter Clark, PhD

peterc@allenai.org, <https://pclark425.github.io/>

Associate Technical Fellow (2004-2010), Research Scientist (1996-2004), Boeing Research

- Principal Investigator for Boeing's work in the Vulcan Halo Project (2002-2010)
Involvement in all aspects of the SRI/Boeing/UT team's work on the Halo AURA system for acquiring and reasoning with scientific knowledge bases since the project's inception. In particular:
 - Developed AURA's **natural language interface for asking questions**, allowing users to ask questions in a simplified form of English (CPL/BLUE) rather than a formal query language.
 - Developed AURA's **knowledge representation and reasoning engine**, KM (The Knowledge Machine), allowing AURA to represent knowledge and perform reasoning to answer users' questions.
- Principal Investigator for IARPA AQUAINT (Advanced QA for Intelligence) Project (2006-8)
- Principal Investigator for Boeing's work in DARPA's Reading to Learn Project (2005-6)
- Principal Investigator for Boeing's work in DARPA's MOBIUS project (2005-8)
- The Technical Expert Locator (Boeing Internal) (1999-2001)
 - Created vision for and developed Boeing's **Technical Expert Locator** (TEL), an AI tool for finding experts that exploited Boeing's Aerospace Thesaurus as a massive source of semantic knowledge. The TEL was deployed and in full commercial use from 2001-10, and was a primary vehicle for locating expertise within the company.

Neutral Representation Project (Boeing Internal) (1996-2000, Principal Investigator 1999-2000)

- Development of techniques, and identification of the limits, for representing aircraft design knowledge in a portable ("neutral") format, to improve reuse and reduce vendor dependence.

Research Fellow - University of Texas at Austin

1994-1996

- Developed a **theory of concept composition** with Prof. Bruce Porter, by which representations of complex, concepts can be assembled from modular components. This work won a AAAI'97 Best Paper award.
- Developed and implemented the **DCE Help-Desk Assistant**, a prototype knowledge-based system that inferred answers to certain types of customer queries using machine reasoning.

Research Associate – Knowledge Systems, Canadian National Research Council 1992-4

- Created and developed **Electronic Trader (ET)**, a prototype for automatically detecting arbitrage opportunities using AI techniques. ET was licensed and used in field trials by Canada's Export Development Corporation.

Postdoctoral Fellow – Dept Computer Science, The University of Ottawa

1991-2

- Developed a technique for **guiding inductive learning using domain knowledge**, with Prof Stan Matwin, biasing the machine away from "clearly nonsensical" hypotheses. This work has been cited over 70 times.

Research Scientist – The Turing Institute, Glasgow, UK

1985-1991

- Developed a **computational model of argumentation**, based on Toulmin's work, by which an expert system and user could interact ("cooperatively argue") to jointly solve a problem. This work was the basis of my PhD.
- Principal Investigator and developer of Optimist, a **full-scale, commercial AI system** that implements the above theory. Optimist assisted geologists in oil exploration, and was in commercial use from 1989-1994.
- Co-developed a new **rule induction algorithm**, CN2, now well-known and frequently cited in the community.

AWARDS AND RECOGNITION

- EMNLP Best Paper Awards (2014,2023,2024)
- Multiple invited talks, e.g., AIED'24, AKBC'21, ESWC'24, workshops at AAAI'24, NeurIPS'22, ACL'23.
- Ai2 Test of Time Award (2022)
- AAAI Senior Member (2014)
- AKBC Best Paper Award (2014)
- Boeing Associate Technical Fellow (2004)
- NSF Grant Review Committee (2003)
- SRI Letter of Commendation for work on DARPA's RKF project (2000)
- Keynote speaker for the Conference and Labs of the Evaluation Forum (CLEF 2012)
- Invited speaker at the DARPA Machine Reading Program (2011)
- Tutorial co-chair for NAACL Computational Linguistics Conference (2010)
- Organizing Committee for International Recognizing Textual Entailment Competition (2010)
- Three Boeing Breakthrough Technology Awards (2006,7,9)
- Co-chair of International Conference on Knowledge Capture (2005)
- AAAI Best Paper Award (1997)

SELECTED PUBLICATIONS

- D. Agarwal, B. Majumder,..., Peter Clark. **AutoDiscovery: Open-ended Scientific Discovery via Bayesian Surprise.** in NeurIPS'25.
- P. Jansen, O. Tafjord, M. Radensky, ..., P. Clark. **CodeScientist: End-to-End Semi-Automated Scientific Discovery with Code-based Experimentation.** in Findings ACL'25.
- N. Kassner, O. Tafjord, A. Sabharwal, K. Richardson, H. Schütze, P. Clark. **Language Models with Rationality.** in EMNLP'24.
- B. Beglin, K. Yang, S. Gupta, K. Richardson, E. Bransom, P. Clark, A. Sabharwal, T. Khot **SUPER: Evaluating Agents on Setting Up and Executing Tasks from Research Repositories.** In EMNLP'24. (**Best paper award**)
- A. Madaan, N. Tandon, ..., P. Clark. **Self-Refine: Iterative Refinement with Self-Feedback.** In NeurIPS'23 (**2200+ citations**)
- W. Yu, M. Jiang, P. Clark, A. Sabharwal. **IfQA: A Dataset for Open-domain Question Answering under Counterfactual Presuppositions.** In EMNLP'23. (**Outstanding paper award**)
- P. Lu, S. Mishra, T. Xia, L. Qiu, K. Chang, S. Zhu, O. Tafjord, P. Clark, A. Kalyan. **Learn to explain: Multimodal reasoning via thought chains for science question answering.** In NeurIPS'22. (**1800+ citations**)
- O. Tafjord, B. Dalvi, P. Clark. **ProofWriter: Generating Implications, Proofs, and Abductive Statements over Natural Language,** Findings of ACL'21.
- N. Kassner, O. Tafjord, H. Schutze, P. Clark. **BeliefBank: Adding Memory to a Pre-Trained Language Model for a Systematic Notion of Belief.** In EMNLP'21.
- P. Clark, O. Tafjord, K. Richardson. **Transformers as Soft Reasoners over Language,** IJCAI'20.
- P. Clark, O. Etzioni, et al., **From 'F' to 'A' on the N.Y. Regents Science Exams: An Overview of the Aristo Project.** AI Magazine 41 (4) pp 39-53, 2020.
- T. Mihaylov, P. Clark, T. Khot, A. Sabharwal. **Can a suit of armor conduct electricity? a new dataset for open book question answering.** In EMNLP'18. (**2000+ citations**)
- P. Clark, I. Cowhey, O. Etzioni, et al. **Think you have solved question answering? try ARC, the AI2 reasoning challenge.** arXiv:1803:05457. 2018. (**3500+ citations**)
- J. Berant, V. Srikumar, P.-C. Chen, A. Vander Linden, B. Harding, B. Huang, P. Clark, C. D. Manning. **Modeling Biological Processes for Reading Comprehension.** In EMNLP'14. (**Best paper award**)
- P. Clark,, P. Harrison, N. Balasubramanian, O. Etzioni. **Constructing a Textual KB from a Biology TextBook.** In Proc. Workshop on Automatic Knowledge Base Construction and Web-scale Knowledge Extraction (AKBC-WEKEX 2012), 2012. (**Best paper award**)
- P. Clark,, Harrison, P., Yao, X. **An Entailment-Based Approach to the QA4MRE Challenge.** In Proc. CLEF 2012 (Conference and Labs of the Evaluation Forum) - QA4MRE Lab, 2012.
- D. Gunning, V. Chaudhri, P. Clark, K. Barker, et al., **Project Halo Update - Progress Toward Digital Aristotle** In AI Magazine (vol 31 no 3), 2010.
- P. Clark, P. Harrison. **An Inference-Based Approach to Recognizing Entailment.** In Proceedings of 2009 Text Analysis Conference (TAC'09), Gaithsburg, Maryland, 2009.
- P. Clark, P. Harrison. **Large-Scale Extraction and Use of Knowledge From Text.** In Proc Fifth Int Conf on Knowledge Capture (KCap'09), 2009.
- P. Clark, P. Harrison. **Boeing's NLP System and the Challenges of Semantic Representation.** In Proc SIGSEM Symposium on Text Processing (STEP'08), Venice, Italy, 2008.
- P. Clark, P. Harrison, T. Jenkins, J. Thompson, R. Wojcik. **Acquiring and Using World Knowledge using a Restricted Subset of English.** In: The 18th International FLAIRS Conference (FLAIRS'05), 2005.
- N. Friedland, G. Matthews, M. Witbrock, P. Clark, et al., **Project Halo: Towards a Digital Aristotle.** In: AI Magazine 25 (4), 2004, pp. 29-47. AAAI Press.
- P. Clark, V. Chaudhri, S. Mishra, J. Thomere, K. Barker, B. Porter. **Enabling Domain Experts to Convey Questions to a Machine: A Modified, Template-Based Approach.** In 2nd International Conference on Knowledge Capture (KCap'03), 2003.
- P. Clark and B. Porter. **Building Concept Representations from Reusable Components.** In AAAI'97, pages 369-376, CA:AAAI Press, 1997. (**Best Paper Award**).
- P. Clark and T. Niblett. **The CN2 Induction Algorithm.** Machine Learning, 3(4):261-283, 1989. (**3700+ citations**)