SEMINAR ANNOUNCEMENT

## The Institute for Assured Autonomy and the Department of Computer Science

*Present*

# Trustworthy AI

September 22, 2020 | 11:00 am
Click **here** to access this virtual event
http://bit.ly/Jeannette-Wing
Password: 366442

Jeannette Wing
Professor, Columbia University

**ABSTRACT**

Recent years have seen an astounding growth in deployment of AI systems in critical domains such as autonomous vehicles, criminal justice, healthcare, hiring, housing, human resource management, law enforcement, and public safety, where decisions taken by AI agents directly impact human lives. Consequently, there is an increasing concern if these decisions can be trusted to be correct, reliable, fair, and safe, especially under adversarial attacks. How then can we deliver on the promise of the benefits of AI but address these scenarios that have life-critical consequences for people and society? In short, how can we achieve trustworthy AI? Under the umbrella of trustworthy computing, there is a long-established framework employing formal methods and verification techniques for ensuring trust properties like reliability, security, and privacy of traditional software and hardware systems. Just as for trustworthy computing, formal verification could be an effective approach for building trust in AI-based systems. However, the set of properties needs to be extended beyond reliability, security, and privacy to include fairness, robustness, probabilistic accuracy under uncertainty, and other properties yet to be identified and defined. Further, there is a need for new property specifications and verification techniques to handle new kinds of artifacts, e.g., data distributions, probabilistic programs, and machine learning based models that may learn and adapt automatically over time. This talk will pose a new research agenda, from a formal methods perspective, for us to increase trust in AI systems.

**BIO**

Jeannette M. Wing is Avanessians Director of the Data Science Institute and Professor of Computer Science at Columbia University. From 2013 to 2017, she was a Corporate Vice President of Microsoft Research. She is Adjunct Professor of Computer Science at Carnegie Mellon where she twice served as the Head of the Computer Science Department and had been on the faculty since 1985. From 2007-2010 she was the Assistant Director of the Computer and Information Science and Engineering Directorate at the National Science Foundation. She received her S.B., S.M., and Ph.D. degrees in Computer Science, all from the Massachusetts Institute of Technology. Professor Wing's general research interests are in the areas of trustworthy computing, specification and verification, concurrent and distributed systems, programming languages, and software engineering. Her current interests are in the foundations of security and privacy, with a new focus on trustworthy AI. She was or is on the editorial board of twelve journals, including the Journal of the ACM and Communications of the ACM. Professor Wing is known for her work on linearizability, behavioral subtyping, attack graphs, and privacy-compliance checkers. Her 2006 seminal essay, titled Computational Thinking, is credited with helping to establish the centrality of computer science to problem-solving in fields where previously it had not been embraced. She is currently a member of: the National Library of Medicine Blue Ribbon Panel; the Science, Engineering, and Technology Advisory Committee for the American Academy for Arts and Sciences; the Board of Trustees for the Institute of Pure and Applied Mathematics; the Advisory Board for the Association for Women in Mathematics; and the Alibaba DAMO Technical Advisory Board. She has been chair and/or a member of many other academic, government, and industry advisory boards. She received the CRA Distinguished Service Award in 2011 and the ACM Distinguished Service Award in 2014. She is a Fellow of the American Academy of Arts and Sciences, American Association for the Advancement of Science, the Association for Computing Machinery (ACM), and the Institute of Electrical and Electronic Engineers (IEEE).

**Johns Hopkins University**

3400 N. Charles Street
Baltimore, MD 21218

**HOW TO REACH US**

IAA Email: IAAinfo@jhu.edu Website: iaa.jhu.edu

CS Email: contactus@cs.jhu.edu Website: cs.jhu.edu

JOHNS HOPKINS UNIVERSITY

COMM-20-03170