

Panel on AI Ethics Education
at the
Symposium on Educational
Advances in Artificial
Intelligence 2017
sponsored by the
Future of Life Institute

Responsible Use of AI Technologies

- Recent Headlines
 - Facebook touts AI benefits as job risks loom
 - The pros and cons of using a robot as an investment adviser
 - Robots can kill and deliver beer. Do we need humans?
 - As Artificial Intelligence evolves, so does its criminal potential
 - Should your driverless car hit a pedestrian to save your life?

Responsible Use of AI Technologies

- Issues

- AI systems can process large quantities of data, detect regularities in them, draw inferences from them and determine effective courses of action -- sometimes faster and better than humans and sometimes as part of hardware that is able to perform many different, versatile and potentially dangerous actions.
- The behavior of AI systems can be difficult to validate, predict or explain since they are complex, reason in ways different from humans and can change their behavior via learning.
- Their behavior can also be difficult to monitor by humans in case of fast decisions, such as buy and sell decisions on stock markets.

Responsible Use of AI Technologies

- Questions

- Do we need to worry about the reliability, robustness, and safety of AI systems?
- Do we need to provide oversight of their operation?
- How do we guarantee that their behavior is consistent with social norms and human values?
- Who is liable for incorrect AI decisions?
- How will AI technology impact standard of living, distribution and quality of work, and other social and economic aspects?

Responsible Use of AI Technologies

- Activities at AAAI 2017
 - Workshop: AI and OR for Social Good
 - Workshop: AI, Ethics and Society
 - Invited Panel: AI for Social Good
 - ACM SIGAI Student Essay Contest on the Responsible Use of AI Technology
 - JAIR Special Track on AI and Society
- Activities elsewhere
 - OSTP Workshops on Preparing for the Future of AI (on benefits and risks of AI)
 - Partnership on AI (to benefit people and society)
 - IEEE Standards Association's Global Initiative for Ethical Considerations in the Design of Autonomous Systems
 - ...

Panel on AI Ethics Education

- The panel will discuss how, as educators, we can incorporate ethical issues into undergraduate or graduate AI classes via class assignments or projects, a single lecture, a larger segment of lectures or a whole class
- Format
 - 13 minute statement by each panelist
 - 15 minute discussion

Panel on AI Ethics Education

- Ben Kuipers, Professor at the University of Michigan
 - Fellow of AAI
 - Fellow of IEEE
 - Fellow of AAAS
 - Ethics officer of ACM SIGAI
 - Currently spends his sabbatical on researching AI ethics



Panel on AI Ethics Education

- Illah Nourbakhsh, Professor at Carnegie Mellon University (CMU)
 - Member of the Global Future Council on the Future of AI and Robotics for the World Economic Forum
 - Senior Advisor to the Future Society, Harvard Kennedy School
 - Executive Board of the Society of Responsible Robotics
 - Book author of “Robot Futures” and “Parenting for Technology Futures”
 - Developed “Ethics and Robotics” class (resources available at ethicsandrobotics.org)



Panel on AI Ethics Education

- Judy Goldsmith, Professor at the University of Kentucky
 - Senior member of AAAI
 - Mentoring awards from CRA and AAAS
 - Teaching awards on the department, college and university levels
 - Developed “Science Fiction and Computer Ethics” class
 - Co-author (with Emanuelle Burton, Nicholas Mattei, and others) of a textbook on "Science Fiction and Computer Ethics"



Panel on AI Ethics Education

- World Economic Forum
 - Unemployment: What happens after the end of jobs?
 - Inequality: How do we distribute the wealth created by machines?
 - Humanity: How do machines affect our behavior and interaction?
 - Artificial stupidity: How can we guard against mistakes?
 - Racist robots: How do we eliminate AI bias?
 - Security: How do we keep AI safe from adversaries?
 - Evil genies: How do we protect against unintended consequences?
 - Singularity: How do we stay in control of a complex intelligent system?
 - Robot rights: How do we define the humane treatment of AI?