

- 2015 – **Consulting Research Analyst, *Open Philanthropy Project***  
Preparing a survey report on philanthropic opportunities to improve future impacts of artificial intelligence, including the development of autonomous weapons and long-term development of highly capable autonomous systems (superintelligence).
- 2015 – **Program Officer for AI Grants, *Future of Life Institute***  
Co-authored a technical agenda based on discussion at the 2015 Puerto Rico conference on the future of AI. Organized, promoted, and oversaw selection process of an international grant program for work aimed to make future AI systems more robust and beneficial, funded by Elon Musk and the Open Philanthropy Project. Awarded a total of \$7.2 million divided among 37 projects chosen from about 200 applicants.
- 2012 – 2015 **Research Fellow, *Programme on the Impacts of Future Technology, Future of Humanity Institute & Oxford Martin School***  
Research, grant writing, and event organization on the impacts of long-term development in artificial intelligence, including six publications, background research for Nick Bostrom’s bestselling book *Superintelligence*, and \$2.5 million in grants won. Funded by the Alexander Tamas Fellowship on Machine Superintelligence and the Future of AI.
- 2011 – 2015 **Research Associate, *Machine Intelligence Research Institute***  
Developed a mathematical framework for idealized agents that can learn complex goals, and authored a paper on this work. Provided expertise and contributed technical material to academic papers and outreach materials.
- 2009 – 2010 **Software Engineer, *Google, Inc.***  
Monitored and maintained high-volume real-time statistical analysis software for Google Search. Worked on a team to migrate an active web service to a new storage backend, then a new AJAX frontend, with no downtime.
- 2005 – 2008 **Student Researcher, *Intel Labs Pittsburgh***  
Developed a distributed, asynchronous reconfiguration algorithm for large modular robotic ensembles based on the avoidance of nonholonomic motion constraints. First-authored and presented a paper on this work at the IEEE/RSJ *International Conference on Intelligent Robots and Systems*.

## Publications

“Research Priorities for Robust and Beneficial Artificial Intelligence.” S. Russell, D. Dewey, M. Tegmark. Forthcoming in *AI Magazine*, 2015.

“Long-Term Strategies for Ending Existential Risk from Fast Takeoff.” D. Dewey. Forthcoming in *Risks of Artificial Intelligence*, V. Müller (ed.), 2015.

“Crucial Phenomena.” D. Dewey. Forthcoming in *How Should Humanity Steer the Future*, A. Aguirre et al. (eds.), 2015.

“Additively Efficient Universal Computers.” Accepted to the 14th International Conference on Unconventional Computation and Natural Computation, 2015.

“Reinforcement Learning and the Reward Engineering Principle.” AAAI Spring Symposium Series, 2014.

“A Representation Theorem for Decisions about Causal Models.” D. Dewey. In proceedings of the Fifth Conference on Artificial General Intelligence, 2012.

“Learning What to Value.” D. Dewey. In proceedings of the Fourth Conference on Artificial General Intelligence, 2011.

“Design of a Braille Writing Tutor to Combat Illiteracy.” N. Kalra, T. Lauwers, D. Dewey, T. Stepleton, M.B. Dias. *Information Systems Frontiers*, Volume 11 Issue 2, April 2009.

“Generalizing Metamodules to Simplify Planning in Modular Robotic Systems.” D. Dewey, S. Srinivasa, M. Ashley-Rollman, M. De Rosa, P. Pillai, T. Mowry, J. Campbell, S. Goldstein. In proceedings of the International Conference on Intelligent Robots and Systems, 2008.

“Iterative Design of a Braille Writing Tutor to Combat Illiteracy.” N. Kalra, T. Lauwers, D. Dewey, T. Stepleton, M. B. Dias. In proceedings of the International Conference on Information and Communication Technologies and Development, 2007.

## Presentations

*Superintelligence*, moderated panel including Stuart Russell, Elon Musk, Nate Soares, and Nick Bostrom, at Effective Altruism Global 2015, Mountain View, CA.

*What could we do about intelligence explosion?*, presented at 2014 Paris Workshop on the Future of Artificial Intelligence, 2014 International Conference on Autonomous Agents & Multiagent Systems.

*The long-term future of AI*, presented at TEDxVienna, 2013.

## Education

2005 — 2008: **BSc in Computer Science, with Additional Major in Philosophy**  
*Carnegie Mellon University*

## Honors

**The SIAI-CA 2011 Academic Prize**, *Singularity Institute for Artificial Intelligence Canada Association*. Awarded “in recognition of [Daniel’s] efforts to improve AI theory, and in the hope he will make further contributions in this field of study.”

**Second Prize, FQXi Essay Contest 2014: How Should Humanity Steer the Future?** *Foundational Questions Institute*. Awarded for essay “Crucial Phenomena.”