

FUTURE OF HUMANITY INSTITUTE

Littlegate House, St. Ebbe's Street, Oxford OX1 1PT
Tel: +44(0)1865 286800
www.fhi.ox.ac.uk



The Future of Humanity Institute at the University of Oxford invites applications for four postdoctoral research positions. We seek outstanding applicants with backgrounds that could include computer science, mathematics, economics, technology policy, and/or philosophy.

The Future of Humanity Institute is a leading research centre in the University of Oxford looking at big-picture questions for human civilization. We seek to focus our work where we can make the greatest positive difference. Our researchers regularly collaborate with governments from around the world and key industry groups working on artificial intelligence. To read more about the institute's research activities, please see <http://www.fhi.ox.ac.uk/research/research-areas/>.

1. Research Fellow – AI – Strategic Artificial Intelligence Research Centre, Future of Humanity Institute (Vacancy ID# 121242). We are seeking expertise in the technical aspects of AI safety, including a solid understanding of present-day academic and industrial research frontiers, machine learning development, and knowledge of academic and industry stakeholders and groups. The fellow is expected to have the knowledge and skills to advance the state of the art in proposed solutions to the “control problem.” This person should have a technical background, for example, in computer science, mathematics, or statistics. Candidates with a very strong machine learning or mathematics background are encouraged to apply even if they do not have experience with AI safety topics, assuming they are willing to switch to this subfield. Applications are due by Noon 6 January 2016. You can apply for this position through the Oxford recruitment website at <http://bit.ly/1M11RbY>.

2. Research Fellow – AI Policy – Strategic Artificial Intelligence Research Centre, Future of Humanity Institute (Vacancy ID# 121241). We are looking for someone with expertise relevant to assessing the socio-economic and strategic impacts of future technologies, identifying key issues and potential risks, and rigorously analysing policy options for responding to these challenges. This person might have an economics, political science, social science, or risk analysis background. Applications are due by Noon 6 January 2016. You can apply for this position through the Oxford recruitment website at <http://bit.ly/1OfWd7Q>.

3. Research Fellow – AI Strategy – Strategic Artificial Intelligence Research Centre, Future of Humanity Institute (Vacancy ID# 121168). We are looking for someone with a multidisciplinary science, technology, or philosophy background and with outstanding analytical ability. The post holder will investigate, understand, and analyse the capabilities and plausibility of theoretically feasible but not yet fully developed technologies that could impact AI development, and to relate such analysis to broader strategic and systemic issues. The academic background of the post-holder is unspecified, but could involve, for example, computer science or economics. Applications are due by Noon 6 January 2016. You can apply for this position through the Oxford recruitment website at <http://bit.ly/1jM5Pic>.

4. Research Fellow – ERC UnPrEDICT Programme, Future of Humanity Institute (Vacancy ID# 121313). This Research Fellowship will work on a new European Research Council-funded UnPrEDICT (Uncertainty and Precaution: Ethical Decisions Involving Catastrophic Threats) programme, hosted by the Future of Humanity Institute at the University of Oxford. This is a research position for a strong generalist, and will focus on topics related to existential risk, model uncertainty, the precautionary principle, and other principles for handling technological progress. In particular, this research fellow will help to develop decision procedures for navigating empirical uncertainties related to existential risk, including information hazards and situations where model or structural uncertainty are the dominating form of uncertainty. The research could take a decision-theoretic approach, although this is not strictly necessary. We also expect the candidate to engage with the research on specific existential risks, possibly including developing a framework to evaluate uncertain risks in the context of nuclear weapons, climate risks, dual use biotechnology, and/or the development of future artificial intelligence. The successful candidate must demonstrate evidence of, or the potential for producing, outstanding research in the areas of relevance to the project, the ability to integrate interdisciplinary research in philosophy, mathematics and/or economics, and familiarity with both normative and empirical issues surrounding existential risk. Applications are due by Noon 6 January 2016. You can apply for this position through the Oxford recruitment website at <http://bit.ly/1HSCKgP>.

Alternatively, please visit <http://www.fhi.ox.ac.uk/vacancies/> or <https://www.recruit.ox.ac.uk/> and search using the above vacancy IDs for more details.

