

Global Catastrophic Risk Symposium  
World Congress on Risk 2012, 18-20 July, Sydney  
Hosted by the Society for Risk Analysis

**Symposium title:** Global Catastrophic Risk

**Chair:** Seth D. Baum, Department of Geography, Pennsylvania State University & Executive Director, Global Catastrophic Risk Institute, USA

**Symposium abstract:**

Global catastrophic risks (GCRs) are risks of events that could significantly harm or even destroy civilization at the global scale. GCRs are thus risks of the highest magnitude, regardless of probability. Major GCRs include climate change, pandemics, nuclear warfare, and potential new technologies. This symposium features interdisciplinary and international perspectives on GCR. The symposium emphasizes the challenge and importance of understanding future GCRs based on environmental and technological change.

GCR has emerged in recent years as an important topic in risk analysis. While individual GCRs have long received attention, new efforts are being made to form an integrated study of GCR. The reasons for studying all GCRs together are twofold. First, GCRs have important interaction effects and risk-risk tradeoffs. For example, extreme climate change could destabilize nuclear-armed countries. Also, developing new technologies such as artificial intelligence could help address other GCRs, or it could be a GCR on its own. Second, different GCRs often face the same issues. For example, the high stakes involved in global catastrophe raise unique analytical and ethical issues. Also, the unprecedented nature of global catastrophe poses challenges for handling uncertainty shared by all GCRs. For these two reasons, forming an interdisciplinary GCR research community is of high importance.

The global nature of GCR lends itself especially well for discussion by an international community. Global catastrophes affect all nations and potentially all individual human beings. In the extreme case, global catastrophes can cause human extinction. Thus all countries have a stake in reducing GCR. Indeed, some GCRs, such as the possibility of a large asteroid impact, tend to bring all of humanity together. But other GCRs, such as nuclear warfare, are fundamentally based on conflict between humans. Still other GCRs, such as climate change, require difficult global collective action. By developing an international community of GCR researchers and other professionals, we can help address the inherently global challenges of GCR.

One important challenge for GCR analysis involves understanding how ongoing changes to our world affect GCR. This challenge is especially acute for GCR because it is such a long-term issue. Global catastrophes do not occur frequently, but are nonetheless important because of their high magnitude. Thus, GCR analysis involves considering events that could happen decades, centuries, or even further into the future. Our society will inevitably undergo extensive changes over these timescales, complicating the analysis. Many of these changes are directly tied to specific GCRs. For example, climate change is widely expected to cause humanity considerable harm, but we do not know how severe this harm could be. Also, ongoing advances in computer hardware and

software render it increasingly likely that artificial intelligence could eventually outsmart humanity, becoming a major destabilizing force, for better or for worse. But technological change is always difficult to predict, and artificial intelligence is no exception. It is thus incumbent upon us as risk analysts to embrace the uncertainty surrounding environmental, technological, and other societal changes as they are relevant to GCR.

This symposium synthesizes the breadth of GCR topics in the context of our changing world. Speakers from fields including geography, economics, computer software, climate science will present a range of perspectives on GCR. It is part of ongoing efforts to form interdisciplinary and international GCR community at SRA and beyond.

### **Speaker Information:**

**Title:** Astrobiology and the Risk Landscape

**Presenter:** Milan Ćirković, Senior Research Associate, Astronomical Observatory of Belgrade & Associate Professor, Department of Physics, University of Novi Sad, Serbia

**Abstract:**

We live in the epoch of explosive development of astrobiology, a novel interdisciplinary field dealing with the origin, evolution, and the future of life. While at first glance its relevance for risk analysis is small, there is an increasing number of crossover problems and thematic areas which stem from considerations of observation selection effects and the cosmic future of humanity, as well as better understanding of our astrophysical environment and the open nature of the Earth system. In considering the totality of risks facing any intelligent species in the most general cosmic context (a natural generalization of the concept of global catastrophic risks or GCRs), there is a complex dynamical hierarchy of natural and anthropogenic risks, often tightly interrelated. I shall argue that this landscape-like structure can be defined in the space of astrobiological/SETI parameters and that it is a concept capable of unifying different strands of thought and research, a working concept and not only a metaphor. Fermi's Paradox or the "Great Silence" problem represents the crucial boundary condition on generic evolutionary trajectories of individual intelligent species; I briefly consider the conditions of its applicability as far as quantification of GCRs is concerned. Overall, such a perspective would strengthen foundations upon which various numerical models of the future of humanity can be built; the lack of such quantitative models has often been cited as the chief weakness of the entire GCR enterprise.

**Title:** How Do We Analyze Global Catastrophic Risks Rationally?

**Presenter:** Yew-Kwang Ng, Professor, Department of Economics, Monash University, Australia

**Abstract:**

More than two decades ago, I published a paper (Social Choice and Welfare 1991) entitled "Should we be very cautious or extremely cautious on measures that may involve our destruction?". Recently, I published another paper "Consumption tradeoff vs. catastrophes avoidance" (Climatic Change 2011). The proposed presentation would

combine the insights of these two papers to discuss the question: How do we analyze global catastrophic risks rationally? Catastrophic risks includes in particular cases where the whole of humankind on earth or even the whole life system becomes extinct. Even just confining to our own species, extinction is probably the most disastrous outcome we may face. If our expected welfare is infinite, even a small (but strictly positive) increase in the probability of extinction decrease our expected welfare infinitely. The rational objective (which is defensible) of maximizing our expected welfare will then require us to be extremely cautious in matters that may threaten our survival, being prepared to avoid such risks at enormous costs. However, is our expected welfare infinite? Our likely finite, though hopefully very long, lifespan (for the whole human species) seems to suggest that our expected welfare is finite. Nevertheless, could the likely much higher technological levels in our future that may not only transform the objective environment but also our own selves serve to increase our expected welfare indefinitely? The possible effects of likely technological advances on our current optimization on such large issues like global warming would be examined. In contrast to the focus on the trade-off between current and future consumption emphasized by contemporary analysts of global warming, the key issue is really the avoidance of catastrophes. The relevance of the results on happiness studies on choices involving catastrophic risks would also be discussed.

**Title:** Catastrophic Risk And Climate Change

**Presenter:** Steven Sherwood, Professor & Co-Director, Climate Change Research Centre, University of New South Wales, Australia

**Abstract:**

Global warming is fundamentally different from typical “doomsday” scenarios. A rise in global temperatures unprecedented in human history is no longshot possibility, it is overwhelmingly likely without dramatic policy changes. We do now know how serious this will be, and do not expect changes to be sudden, but the huge inertia involved means mitigation must begin very early to be effective. While some potential catastrophes are like being hit by a bus, climate change is more like a cancer, of uncertain prognosis, that gradually becomes less treatable as the patient procrastinates or remains in denial.

Basic physics implies that the warming expected from likely fossil fuel reserves and climate sensitivity values, even in the absence of “tipping points,” would produce enough warming within a few centuries to bring some portions of the Earth to the brink of inhabitability just from summertime heat stress alone without even considering other impacts. In the worst case, large portions of the tropics would be pushed over the brink. The future of civilisation is threatened unless we can either confirm that climate sensitivity and/or fossil fuel reserves are smaller than currently expected, or somehow ensure that a large portion of our fossil fuel endowment (or at least the carbon it contains) remain permanently in the ground. This appears unlikely without strong policies of self-restraint.

An interesting link between societal threats may occur through the so-called discount rate used by economists when comparing costs and benefits of emissions controls. The universal practice of discounting future economic impacts compared to present ones often causes mitigation to appear uneconomic, even though future damages far exceed mitigation costs. It can be argued that the primary reason that individuals discount is to hedge against the risk of unexpected events spoiling the deal. If we adopt

the same philosophy toward climate, our mitigation decisions hinge largely on how pessimistic we are about other threats to society.

**Title:** The Intelligence Stairway And Global Catastrophic Risk

**Presenter:** Jaan Tallinn, Ambient Sound Investments, Estonia

**Abstract:**

I present a model for the succession of optimisation processes on our planet, culminating in a new process that could cause a global catastrophe. The optimisation arguably began 4 billion years ago with the emergence of life and evolution. Evolution can be viewed as a set of competing optimisation processes that optimise the fitness of species for survival and reproduction. A succession occurred about 100 thousand years ago with the emergence of humans. The optimisation power of humans exceeded that of evolution, and so human-driven technological progress replaced evolution as the dominant force. There seems to be a pattern in which optimisation processes tend to produce agents whose optimisation power exceeds that of the process that produced them. I call this pattern the Intelligence Stairway.

Recent developments in computer technology suggest a possible new step in the Intelligence Stairway dominated by computers who are smarter than their creators. If this step occurs, then according to the model, the emergence of such computers (artificial general intelligence, AGI) marks the end of human-driven technological progress and the beginning of a new phase: AGI-driven "intelligence explosion". Moreover, the relative time scales and environmental impacts of evolution and technological progress suggest the uneasy conclusion that the intelligence explosion could be a global catastrophe. Two qualitatively different approaches could avoid the catastrophe. One approach is to mathematically prove that the intelligence explosion will be favorable to humanity (the "Friendly AI" approach). The other approach is to limit AIs to narrow domains so they could not dominate humans. I argue that the latter approach is more pragmatic. We need a coordinated effort to establish and enforce a safety protocol for AI developers so they do not produce AGIs by accident.

**Title:** Regulating Global Catastrophic Risk

**Presenter:** Arden Rowell, Assistant Professor and Richard & Marie Corman Scholar at the University of Illinois College of Law, USA

**Abstract:**

Global catastrophic risks—like climate disruption, global warfare, ecological collapse, pandemic and disruptive emerging technologies—pose particular challenges for regulatory design, because they pair high stakes, social dissensus, and limited information. What role does law have to play in prioritizing, managing and responding to these risks? Which legal institutions should be responsible for managing global catastrophic risks, and should those institutions be local, national, or global? Should laws focus on global catastrophic risks as a category, or should regulatory strategies be tailored to the specifics of each risk? If tailoring is appropriate, what legal strategies are available for managing key global catastrophic risks, and what are the characteristics of circumstances where law is no longer helpful?