Synopsis

Peter Cheeseman

Title: **Implications of Artificial General Intelligence (AGI)**
Synopsis: It looks highly probable that the first AGI is only a few decades away. When it arrives, the implications for continued human existence are dire! One indicator is that we do not know of any alien intelligences--but why? Given the recent discovery that most stars have planets, and that life on earth got started shortly after it formed, it is likely that there will be billions of planets in our galaxy with life. It is also likely that at least some of these lifeforms will have advanced to human level of intelligence and beyond--yet we see no evidence of this! While there are many possible explanations for the missing aliens, one explanation is that alien civilizations developed AGIs and the AGIs then destroyed their creators. Another explanation is that the aliens merged with their AGIs making them individually super-powerful.
In this case it only takes one rogue super-intelligent alien to destroy them all. Even if we are very cautious in developing a "friendly" AGI, the risk of bad unintended consequences is very high. Trying to out-smart an AGI that is orders of magnitude smarter than us is a losing proposition, so once created, it is unstoppable. While I believe there are possible technical methods to develop an AGI that is benevolent to us, mainly by interacting with us as it develops, this misses the main risk. The main risk is not the AGI, but who owns it! Developing an AGI will take considerable resources, so whoever puts up these resources will want to exploit it to recover their investment, and gain a powerful dominating position. Other organizations will regard this monopoly on power to be intolerable, and will try to develop their own AGI or steal the existing AGI technology. Once we have competing hostile organizations with AGIs, the game is up!

Joscha Bach

Title: **Strong AI: Why we should be concerned about something nobody knows how to build**
Synopsis: At the moment, nobody fully knows how to create an intelligent system that rivals or exceed human capabilities (Strong AI). The impact and possible dangers of Strong AI appear to concern mostly those futurists that are not working in day-to-day AI research. This in turn gives rise to the idea that Strong AI is merely a myth, a sci fi trope and nothing that is ever going to be implemented. The current state of the art in AI is already sufficient to lead to irrevocable changes in labor markets, economy, warfare and governance. The need to deal with these near term changes does not absolve us from considering the implications of being no longer the most intelligent beings on this planet.
Despite the difficulties of developing Strong AI, there is no obvious reason why the principles embedded in biological brains should be outside of the range of what our engineering can achieve in the near future. While it is unlikely that current narrow AI systems will neatly scale towards general modeling and problem solving, many of the significant open questions in developing Strong AI appear to be known and solvable.

Marcus Hutter

Title: **The Future of Artificial Intelligence and Humanity**
Synopsis: Intelligent systems have an increasing impact on our society. We already benefit from it particularly through our smart phones and experiences on the Internet. Self-driving cars and household robots could also be standard technologies very soon. Intelligent Systems approach and exceed human intellectual capabilities in an increasing number of domains (expert systems, board games such as Chess and Go, speech recognition and translation, process optimization, search engines), some can autonomously deal with unknown and unexpected situations. Indeed, Intelligent Systems have the potential to be deployed in almost any facet of our lives. The creation of super-intelligent systems will change our society in the next couple of decades more than the industrial revolution has in the last 200 years. It will have many & immense ethical, political, economical, medical, cultural, humanitarian, religious, in art, warfare, and other implications. For instance, it should change our attitude towards un(der)employment and ultimately the value of life itself.

Bio: Marcus Hutter is Professor in the RSCS at the Australian National University in Canberra, Australia. He received his PhD and BSc in physics from the LMU in Munich and a Habilitation, MSc, and BSc in informatics from the TU Munich. Since 2000, his research at IDSIA and now ANU is

centered around the information-theoretic foundations of inductive reasoning and reinforcement learning, which has resulted in 100+ publications and several awards. His book "Universal Articial Intelligence" (Springer, EATCS, 2005) develops the first sound and complete theory of AI. He also runs the Human Knowledge Compression Contest (50'000€= H-prize).

Ben Goertzel

Title: **AGI Will Obsolete Human Life As We Know It -- Thank Goodness**
Synopsis: A surprise