TRUSTABLE EMBEDDED ARTIFICIAL INTELLIGENCE

« From a system and software engineering point of view »

François TERRIER - 2019/03/07

ECOSYSTEM STRUCTURATION on AI SAFETY

Why do we need a AI Safety Landscape?

- AI Safety stretches the limits of the traditional discipline of safety engineering.

- More consensus in terminology and meaning is key towards aligning the understanding of engineering and socio-technical concepts

- Focus on generally accepted knowledge so that the knowledge described is applicable to most AI Safety problems.

What concrete aspects do we target?

AISafety 2019

- **Bring together the most relevant initiatives and leaders** interested on developing a map of AI Safety knowledge to seek consensus.

- Expected outcome is a series of workshop **reports summarizing discussions about a landscape of AI Safety**, including needs, challenges, practices and gaps.

- **Align and synchronize with other Related initiatives** towards a body of knowledge.

- Way of Working : Open **face-to-face meetings** that will take place together with the international workshops of AISafety (at IJCAI) and SafeAI (at AAAI).

Organisation Committee

AISafety 2019

Huáscar Espinoza, Commissariat à l´Energie Atomique, France

Han Yu, Nanyang Technological Univ. Singapore

Xiaowei Huang, University of Liverpool, UK
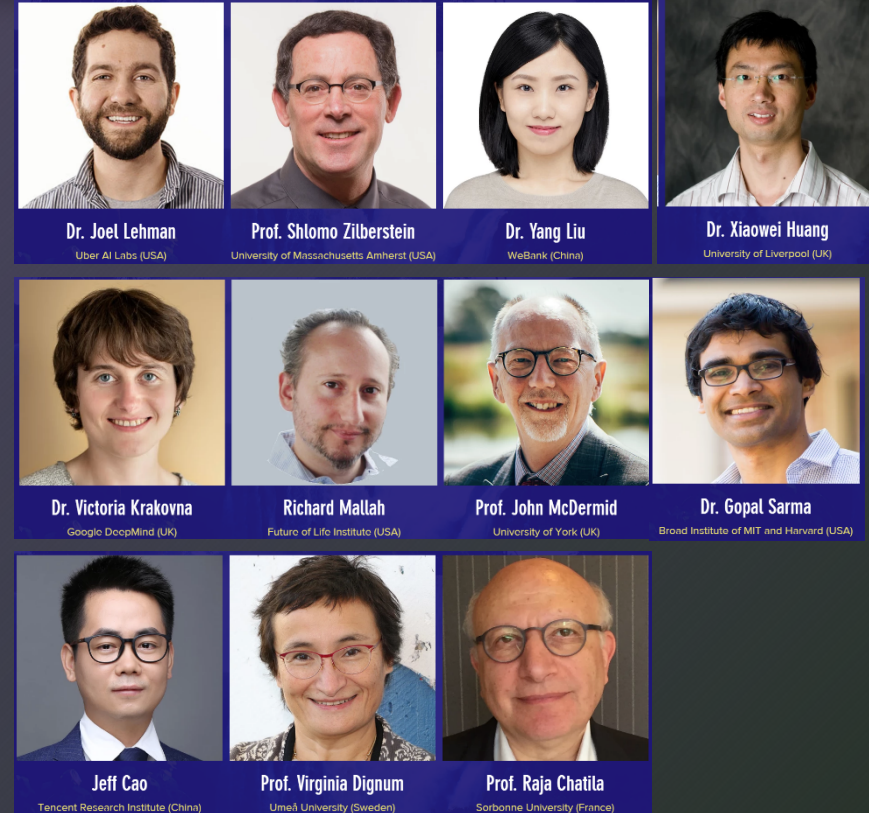
Freddy Lecue, Thales, Canada

Cynthia Chen, University of Hong Kong, China

José Hernández-Orallo, Universitat Politècnica de València, Spain

Seán Ó hÉigeartaigh, Univ. of Cambridge, UK

Richard Mallah, Future of Life Institute, USA

Dr. Joel Lehman
Uber AI Labs (USA)

Prof. Shlomo Zilberstein
University of Massachusetts Amherst (USA)

Dr. Yang Liu
WeBank (China)

Dr. Xiaowei Huang
University of Liverpool (UK)

Dr. Victoria Krakovna
Google DeepMind (UK)

Richard Mallah
Future of Life Institute (USA)

Prof. John McDermid
University of York (UK)

Dr. Gopal Sarma
Broad Institute of MIT and Harvard (USA)

Jeff Cao
Tencent Research Institute (China)

Prof. Virginia Dignum
Umeå University (Sweden)

Prof. Raja Chatila
Sorbonne University (France)

More information at www.ai-safety.org/ai-safety-landscape