Dr. Kristinn R. Þórisson.

MASTER AND SERVANT

PHOTOS BY **PÁLL STEFÁNSSON.**

***Mica Allan** met to talk morals, machines, intelligence and ideas with **Dr. Kristinn R. Þórisson**, the managing director of the Icelandic Institute for Intelligent Machines (IIIM) and Associate Professor of Computer Science at Reykjavík University, learning what the IIIM is doing to further robotics and support the police in Iceland.*

To a 25-year-old, life before the internet was an ancient time, right back there with the dinosaurs. But in a few decades, the internet could seem ancient to us all with the development of artificial intelligence (AI) and the machines it will spawn. Along with the Googles, the IBMs and the DARPAs (the United States Defense Advanced Research Projects Agency) of this world, AI research is being pursued by academics who spend their days working on algorithms and robots to shape and reach the as-yet-unknown destination that AI will take us to. One such research lab on the AI journey is the non-profit Icelandic Institute for Intelligent Machines (IIIM), based at Reykjavík University (RU).

BRIDGE TO THE FUTURE

IIIM is Iceland's first independent research institute for AI, robotics and simulation, and brings together leading forces in academia and industry to foster the exchange of ideas, inventions and brainpower. Its principal partners are the School of Computer Science and the Center for Analysis and Design of Intelligent Agents; both based at RU, and funded through competitive grants and projects with industry partners, such as the Icelandic prosthetics and orthopedic company Össur.

Dr. Kristinn R. Þórisson is the IIIM's managing director and his background reads like something from a Dan Brown novel. He's pioneered new ideas in communicative multimodal agents at MIT and has taught advanced AI courses at Columbia, RU and the Royal Institute of Technology in Sweden. He's also the co-founder of RU's Centre for Analysis and Design of Intelligent Agents. His experience working in industry has seen him consult for NASA and British Telecom and he has worked with Honda on a humanoid robot, as well as creating the world's first cognitive architecture that learns complex skills by observation and programming itself. In addition, he sits on the editorial board of the *Journal of Artificial General Intelligence* and has twice won the Kurzweil Award for his work on AI (Kurzweil is a futurist, inventor and Director of Engineering at Google, hailed by *Inc.* magazine as the "rightful heir to Thomas Edison.")

Kristinn started IIIM to help bridge research in two sectors that operate very differently in terms of short and long term research focus. "In academia, a year is the blink of an eye, it often takes a year or more to get your work published and academics think ten years ahead; whereas industry depends on getting operational results within two to three years at most. With an academic timeframe a company would go under in a year. So there's this chasm between academia and industry." Having identified this 'chasm,' Kristinn is optimistic that the role IIIM plays, and the work it produces, will be of valuable service to the research community in academia and industry alike.

He explains one of the central goals of IIIM: "What does that chasm mean for society? It means that return on investment in academic research takes longer than it should because the knowledge takes too long to find its way into applications. A key operational goal of IIIM is to bring academic results more quickly to industry, and enable market opportunities to be more visible to academic researchers."

Founded in 2009, the IIIM works with a dozen companies on a variety of projects. One such project is with Össur, where machine learning makes prosthetics better adapt to their users. Another project is with the National Commissioner of the Icelandic Police, developing software to run on new mobile devices for patrolling officers. Kristinn's hope is that the success of these projects, which has already been demonstrated, will clearly underline the service to society that the institute is rendering. The ultimate aim? For the institute to become self-sufficient. Some may say this is a tall order, given that it receives only about half the funding of similar initiatives abroad. However, IIIM has received ISK 400 million (USD 3.1 million) from the Icelandic Centre for Research and the area of AI and simulation technologies is one of three Centers of Excellence programs from 2009 to 2015.

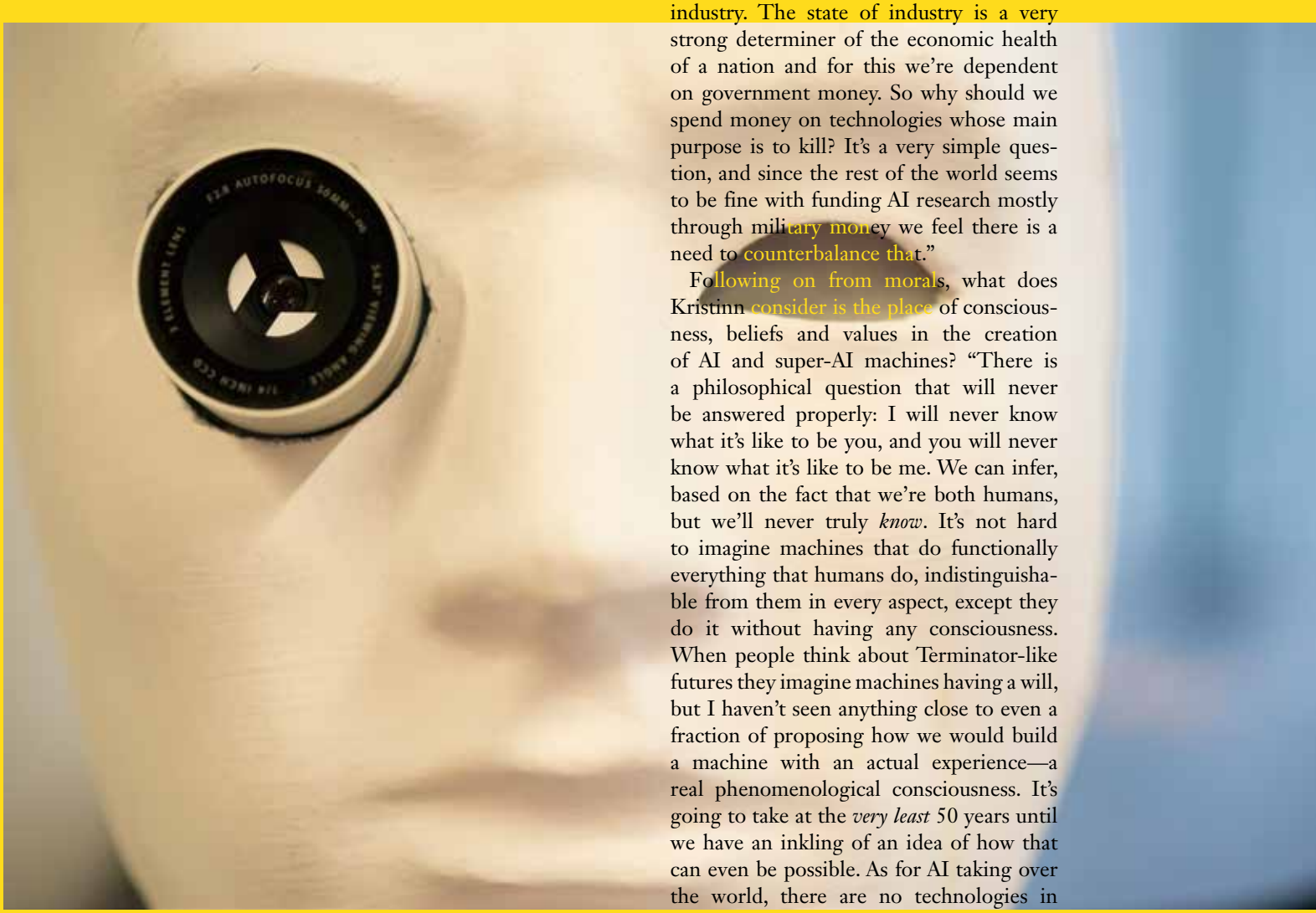
FRIEND OR FRANKENSTEIN?

Given his time in the field, how does Kristinn think the approaches to AI have developed? "By-and-large their nature is fundamentally the same as it was 50 years ago. People are using the same methodologies. AI systems are advanced tools. There's been a lot of talk recently about autonomy, which is really what's at the heart of the dystopian future some have painted, where machines takeover of their own accord. At RU we have some interesting new developments on the subject of truly self-programming systems—something that would be necessary for such a Frankensteinian future. But this work is at a very, very early stage. For the pessimists, in the worst scenarios autonomous robots take over humanity and destroy it. This, however, is still so hypothetical to be entertained as anything but science fiction, and we are not in that camp."

The AI future the IIIM is helping develop does not reflect the images of machine mania that blockbusters often portray in splashy color on the big screen. Kristinn views the advancement of AI in a different light. "In more realistic visions of the future, machines take over our jobs. But that is just the industrial revolution extending its reach into the information age, and should not come as a surprise. Automation takes over because of the structure we've erected for people to make decisions for how they pay their bills and taxes, and the methods used to decide how to apply those: in a market economy automation will always be preferred over manual labor, when applicable and available. AI today is like Henry Ford's conveyor belt. It's a tool. Granted—you can do a hell of a lot more with an 'information conveyor belt' than a physical conveyor belt, but it's still a conveyor belt obeying the same principles for its application—and abuse."

MORALS AND MACHINES

Perhaps to address the concern about a potential Frankensteinian future, and suspicions about how AI will be developed and used, IIIM is the very first research and development group to reject the development of technologies intended for military operations and to issue an ethics policy, which it released in late August 2015. What's been the response to this? Kristinn explains: "From those who have already taken steps and thought about these issues, it's been overwhelmingly positive. From the rest of the community—not much feedback these first few weeks. We are getting some press, but no deluge of requests for commentary or interviews. Which is a bit surprising, given how novel it is." The IIIM's Ethics Policy for Peaceful R&D (research



and development) includes upholding the law and honoring the UN's Humans Rights Declaration, as well as taking one step further. "We decided against using military funding because the IIIM's specific focus is to speed up innovation and to make basic academic research more relevant to industry. The state of industry is a very strong determiner of the economic health of a nation and for this we're dependent on government money. So why should we spend money on technologies whose main purpose is to kill? It's a very simple question, and since the rest of the world seems to be fine with funding AI research mostly through military money we feel there is a need to counterbalance that."

Following on from morals, what does Kristinn consider is the place of consciousness, beliefs and values in the creation of AI and super-AI machines? "There is a philosophical question that will never be answered properly: I will never know what it's like to be you, and you will never know what it's like to be me. We can infer, based on the fact that we're both humans, but we'll never truly *know*. It's not hard to imagine machines that do functionally everything that humans do, indistinguishable from them in every aspect, except they do it without having any consciousness. When people think about Terminator-like futures they imagine machines having a will, but I haven't seen anything close to even a fraction of proposing how we would build a machine with an actual experience—a real phenomenological consciousness. It's going to take at the *very least* 50 years until we have an inkling of an idea of how that can even be possible. As for AI taking over the world, there are no technologies in any mainstream AI lab today that suggest how to make machines that are sufficiently autonomous to even be capable of deciding one day that they like to play with yoyos. Today's AI is an advanced power tool. And last time I checked, no one is afraid of power tools taking over the world." ⋆