

CONTENTS

ACKNOWLEDGMENTS ix

INTRODUCTION: ON THE PROMISE AND PERIL OF AI BY JOHN BROCKMAN xv

CHAPTER 1. Seth Lloyd: Wrong, but More Relevant Than Ever 1

It is exactly in the extension of the cybernetic idea to human beings that Wiener's conceptions missed their target.

CHAPTER 2. Judea Pearl: The Limitations of Opaque Learning Machines 13

Deep learning has its own dynamics, it does its own repair and its own optimization, and it gives you the right results most of the time. But when it doesn't, you don't have a clue about what went wrong and what should be fixed.

CHAPTER 3. Stuart Russell: The Purpose Put into the Machine 20

We may face the prospect of superintelligent machines—their actions by definition unpredictable by us and their imperfectly specified objectives conflicting with our own—whose motivations to preserve their existence in order to achieve those objectives may be insuperable.

CHAPTER 4. George Dyson: The Third Law 33

Any system simple enough to be understandable will not be complicated enough to behave intelligently, while any system complicated enough to behave intelligently will be too complicated to understand.

CHAPTER 5. Daniel C. Dennett: What Can We Do? 41

We don't need artificial conscious agents. We need intelligent tools.

CHAPTER 6. Rodney Brooks: The Inhuman Mess Our Machines Have Gotten Us Into 54

We are in a much more complex situation today than Wiener foresaw, and I am worried that it is much more pernicious than even his worst imagined fears.

CHAPTER 7. Frank Wilczek: The Unity of Intelligence 64

The advantages of artificial over natural intelligence appear permanent, while the advantages of natural over artificial intelligence, though substantial at present, appear transient.

CHAPTER 8. Max Tegmark: Let's Aspire to More Than Making Ourselves Obsolete 76

We should analyze what could go wrong with AI to ensure that it goes right.

CHAPTER 9. Jaan Tallinn: Dissident Messages 88

Continued progress in AI can precipitate a change of cosmic proportions—a runaway process that will likely kill everyone.

CHAPTER 10. Steven Pinker: Tech Prophecy and the Underappreciated Causal Power of Ideas 100

There is no law of complex systems that says that intelligent agents must turn into ruthless megalomaniacs.

CHAPTER 11. David Deutsch: Beyond Reward and Punishment 113

Misconceptions about human thinking and human origins are causing corresponding misconceptions about AGI and how it might be created.

CHAPTER 12. Tom Griffiths: The Artificial Use of Human Beings 125

Automated intelligent systems that will make good inferences about what people want must have good generative models for human behavior.

CHAPTER 13. Anca Dragan: Putting the Human into the AI Equation 134

In the real world, an AI must interact with people and reason about them. “People” will have to formally enter the AI problem definition somewhere.

CHAPTER 14. Chris Anderson: Gradient Descent 143

Just because AI systems sometimes end up in local minima, don’t conclude that this makes them any less like life. Humans—indeed, probably all life-forms—are often stuck in local minima.

CHAPTER 15. David Kaiser: “Information” for Wiener, for Shannon, and for Us 151

Many of the central arguments in The Human Use of Human Beings seem closer to the 19th century than the 21st. Wiener seems not to have fully embraced Shannon’s notion of information as consisting of irreducible, meaning-free bits.

CHAPTER 16. Neil Gershenfeld: Scaling 160

Although machine making and machine thinking might appear to be unrelated trends, they lie in each other’s futures.

CHAPTER 17. W. Daniel Hillis: The First Machine Intelligences 170

Hybrid superintelligences such as nation-states and corporations have their own emergent goals and their actions are not always aligned to the interests of the people who created them.

CHAPTER 18. Venki Ramakrishnan: Will Computers Become Our Overlords? 181

Our fears about AI reflect the belief that our intelligence is what makes us special.

CHAPTER 19. Alex “Sandy” Pentland: The Human Strategy 192

How can we make a good human-artificial ecosystem, something that’s not a machine society but a cyberculture in which we can all live as humans—a culture with a human feel to it?

CHAPTER 20. Hans Ulrich Obrist: Making the Invisible Visible:
Art Meets AI 206

Many contemporary artists are articulating various doubts about the promises of AI and reminding us not to associate the term “artificial intelligence” solely with positive outcomes.

CHAPTER 21. Alison Gopnik: Als Versus Four-Year-Olds 219

Looking at what children do may give programmers useful hints about directions for computer learning.

CHAPTER 22. Peter Galison: Algorithmists Dream of Objectivity 231

By now, the legal, ethical, formal, and economic dimensions of algorithms are all quasi-infinite.

CHAPTER 23. George M. Church: The Rights of Machines 240

Probably we should be less concerned about us-versus-them and more concerned about the rights of all sentients in the face of an emerging unprecedented diversity of minds.

CHAPTER 24. Caroline A. Jones: The Artistic Use of Cybernetic Beings 254

The work of cybernetically inclined artists concerns the emergent behaviors of life that elude AI in its current condition.

CHAPTER 25. Stephen Wolfram: Artificial Intelligence and the
Future of Civilization 266

The most dramatic discontinuity will surely be when we achieve effective human immortality. Whether this will be achieved biologically or digitally isn't clear, but inevitably it will be achieved.

Index 285