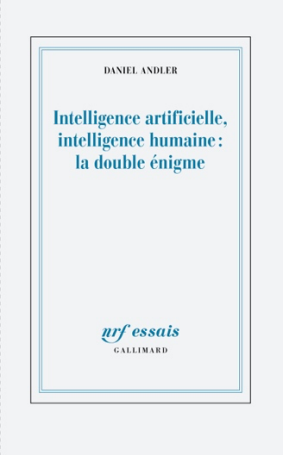
Original title: *Intelligence artificielle, intelligence humaine : la double énigme*
Paris, Gallimard, 432 pp.; published May 4, 2023. Publisher's page here
Author: Daniel Andler, emeritus professor of philosophy at Sorbonne Université, senior researcher in the Department of Cognitive Studies of École normale supérieure (Paris), member of Académie des sciences morales et politiques. Author's page here.



Artificial intelligence, human intelligence: the dual enigmas

What is the relation between artificial intelligence and human intelligence? What should it be, what will it likely be as it evolves? Those are questions as old as AI itself, and there is no shortage of answers to be found. The recent spectacular arrival in the public arena of ChatGPT and with it generative AI has brought the debate to a new level of intensity. Any answer depends on what we mean by artificial intelligence and what we mean by human intelligence. And neither of these notions is self-evident.

The first objective of the book is to present artificial intelligence, which is both a project and a series of technical achievements. Far from dating from yesterday, the project was born after the Second World War, and is rooted in a conception of the human mind which inspired both AI itself and cognitive science. It is impossible to understand today's AI and seriously speculate what its next steps might be without having an idea of its trajectory. Early AI, born in the 1950s, took conscious reasoning as its paradigm. From the 1980s onwards, AI a sharp turn away from this so-called "symbolic" AI, and espoused "connectionism", rebranded as "deep learning", which places perception at the heart of intelligence. Although AI owes deep learning its present triumph, the two approaches co-exist, as becomes evident by taking a closer look at the systems produced by AI. Although pitched at different levels of description, both symbolic and connectionist systems belong to the same special realm, that of input-ouput algorithms that produce solutions to certain problems encountered by human cognition.

The book then turns to human intelligence. While most psychologists understand it as the ability to solve problems, the author proposes to view it as the ability to deal with situations: an individual's intelligence consists in the way in which she acts in the concrete situations that she encounters. Nor does it reside in any specific mental process or function. Rather, it is a norm that applies to behavior: it qualifies the fit between an individual and his world. But there is no objective standard against which this fit can be measured, in the way a piece of string Is measured in inches or the power developed by a cyclist when climbing a pass is measured in kilowatts. Yet intelligence is not a purely subjective matter: like aesthetic or ethical judgments, it gives rise to reasoned debates which often, but not necessarily, lead to consensus.

The distinction between situation and problem (as made precise for the purpose of the analysis) is crucial: a situation presents itself to an individual in the here-and-now in the form of a call to act (including mentally) in a particular context. Human intelligence manifests itself in the way the human agent organizes context — an operation that for an intelligent artificial system makes no sense. It does not encounter situations. It only solves problems that human agents put to it.

Now is often happens that when facing a given situation, a human agent will formulate a problem, one whose solution, she believes, will indicate how best to act. It is only at this point that artificial intelligence can help humans, and in fact it can solve an ever-increasing variety of problems.

This is how the work sets out to resolve two closely related enigmas. The first is that human intelligence, although it undoubtedly results from our mental capacities, and manifests itself in every situation thanks to the particular deployment of these capacities, does not reside in any of the components, or combination thereof, uncovered by cognitive science. (This is in line with the common intuition that it is elusive.) The second enigma is that despite its spectacular progress, artificial intelligence, while constantly aiming for human intelligence, does not come significantly close to it.

What place, then, should be assigned to such a powerful intellectual tool, which is devoid of intelligence in the human sense of the term? This is the central question in the ethics of artificial intelligence. It is not a matter of deploying artificial intelligence in an ethical way - how can we ensure that the introduction of intelligent artificial systems takes place under the right ethical conditions? The more basic question is whether, when and under what conditions it is ethical to deploy AI, in view of its expected and probable net contribution and the concrete conditions under which the deployment would be carried out.

Yet the profession, encouraged by its recent successes, forges ahead, and aims, at least officially, to catch up with and even surpass human intelligence. The race to give its systems ever-greater autonomy, which in the eyes of their designers is an essential part of true intelligence, is in fact leading us towards an unknown fraught with threats. What artificial intelligence should bring us is not colleagues or replacements, but docile assistants of a new kind, with whose help we may be able to solve some pressing problems, both small and large.

TABLE OF CONTENTS

PART ONE. THE IDEA, THE UNDERTAKING, THE TRAJECTORY OF AI

CHAPTER 1. *Artificial intelligence at a glance: from vision to strategic withdrawal*

1.1 What does "artificial intelligence" mean?

1.2 AI's Promethean moment

1.3 Strategic withdrawal

1.4 The five stages of strategic withdrawal

1.4.a *Renouncing "thinking" thought*

1.4.b *Renouncing equivalence with human intelligence*

1.4.c *Renouncing general intelligence*

1.4.d *Renouncing independence from human intelligence*

1.4.e *Renouncing reflection*

1.5 The three faces of contemporary AI

1.6 An uncertain horizon

CHAPTER 2. *The sources of artificial intelligence*

2.1 Formal epistemology and logic

2.2 Computation, algorithm

2.4 The computer

2.4 System and function

2.5 Behavior and autonomy

2.6 Information and communication

2.7 The neo-mechanist unification

2.8 Cybernetics: crucible and matrix

CHAPTER 3. *The classical era*

3.1 How AI looks

3.2 The foundations of classical or symbolic AI

3.3 In search of algorithms

3.4 The seasons of classical symbolic AI

3.5 The troubles of symbolic AI: from microworlds to expert systems

3.6 Does symbolic AI suffer from a fundamental deficiency?

CHAPTER 4: *The age of connectionism*

4.1 The conceptual framework of neurocomputing and connectionist models

4.2 From perception to prediction

4.3 Two conceptions of what human intelligence contributes to artificial intelligence

4.4 The difficult rise of neurocomputing and the triumph of deep learning (DL)

4.4.a Initial coexistence: from McCulloch and Pitts (1943) to the perceptron (late 1960s)

4.4.b *The Thirty Years' War (1970-2000)*

4.4.c *The road to victory (2000-2012)*

4.4.d *The reign of deep learning (2012 to 2017 and beyond)*

4.5 The weaknesses of deep learning

4.6 Its limited skills

CHAPTER 5. *The beginnings of a third era?*

5.1 Developments in deep learning

5.2 The transformer transforms language processing

5.3 The sudden emergence of large language models (LLMs)

5.4 Strenghts and weaknesses of generative AI

5.5 Robotics and AI getting closer again

5.6 Augmented intelligence

PART TWO. THE QUESTION OF INTELLIGENCE AND THE FUTURE OF AI

CHAPTER 6. *Can one compare artificial intelligence and human intelligence?*

6.1 First-step rhetoric and the no-miracle argument

6.2 Two disagreements

6.3 Three escapes from the comparison issue

6.4 The fourth box in the table: accept the terms of the comparison

6.5 Searching for the missing ingredient
6.5.a Consciousness
6.5.b Meaning
6.5.c Common sense

6.6 The limits of additivity and the end of the road

CHAPTER 7. *Intelligence and animality*

7.1 The return of an old question

7.2 In search of a definition of intelligence

7.3 Organism and environment: the ecological perspective

7.4 Problem and situation

7.5 Intelligence: how to deal with Situations

CHAPTER 8. *Intelligence and humanity*

8.1 From Situation to Problem

8.2. Two types of Situation?

8.3 The breakdown of the analytical method

8.4 A trick of human intelligence: cooperation

8.5 Autonomy

8.6 The double face of the concept of intelligence, and the enigma dispelled

CHAPTER 9. *The artificial intelligence of dreams*

9.1 The project of a powerful artificial intelligence

9.2 The goal of artificial general intelligence

9.3 The two paths: algorithmic, biological

9.4 Is powerful artificial intelligence possible?
(1) The IGA's arguments

9.5 Is powerful artificial intelligence possible?
(2) The case of super-intelligence

9.6 The double limit that super-intelligence, and therefore *AI*,will probably never cross

CHAPTER 10. *Artificial Intelligence and the Good*

10.1 The ethical awakening of AI

10.2 What does the ethics of artificial intelligence aim for?

10.3 Charters and principles

10.4 Ethical reflection is before us

10.5 The “control problem” and values alignment

10.6 So that the cows are well looked after (old French maxim– looking for English equivalent)

Acknowledgements

References

Notes

Subject index

Author index