# THE MANY MEANINGS OF AI

Artificial intelligence is often seen as the pinnacle of the drive for digitalisation, but there are pitfalls along the way. **Erik Hollnagel** takes us on a potted tour of AI, its many meanings, and its proper use.

## KEY POINTS:

- **We have developed technology throughout history to amplify our abilities, but amplification has worked less well for cognitive abilities than for physical abilities.**
- **Technology is often used to replace or substitute human functions rather than to amplify or support them.**
- **Automation to replace human functions with technology has often failed to support people to remain in control of what happens.**
- **Digitalisation could usefully amplify what we do well, to amplify intelligence, while we stay in control.**
- **Piecemeal development of advanced technical solutions makes it impossible to comprehend the consequences of digitalisation.**

In the 1980s, the first large-scale commercial applications of artificial intelligence (AI) began – although AI in those days was called 'expert systems'. Among those who worked in the business it was often joked that the acronym AI had different possible interpretations. For academics AI meant Artificial Intelligence; for hackers it meant Anything Impossible; for defence it meant Anything Invincible; and for marketing it meant Anything Interesting (and still does?). But there is a fifth and perhaps more appropriate interpretation: Amplified Intelligence.

### The digital dawn

We are today, for better or worse, forced to cope with digitalisation. Despite the frequency by which the term is used at present, digitalisation actually began in the 1960s with the first message sent over the ARPANET in 1969. The first expert systems saw the light of day in the early 1970s, well before the advent of the personal computer (Apple II in 1977 and the IBM PC in 1981). Digitalisation, meaning the conversion of text, pictures, or sound into a digital form and not least the processing thereof by computing machinery, had during the 1970s gradually been introduced in safety-critical process control (Netland & Hol, 1977) and in many service applications.

In the following decades digitalisation became ubiquitous and completely transformed daily life, first in industries and then in our private and public existences. It would therefore seem reasonable to assume that the problems of digitalisation by now have been completely solved. The reason why this is not the case, hence the motivation for this special issue of HindSight, is that the short-term benefits of digitalisation for human performance often have been outdone by the longer-term problems that digitalisation creates.

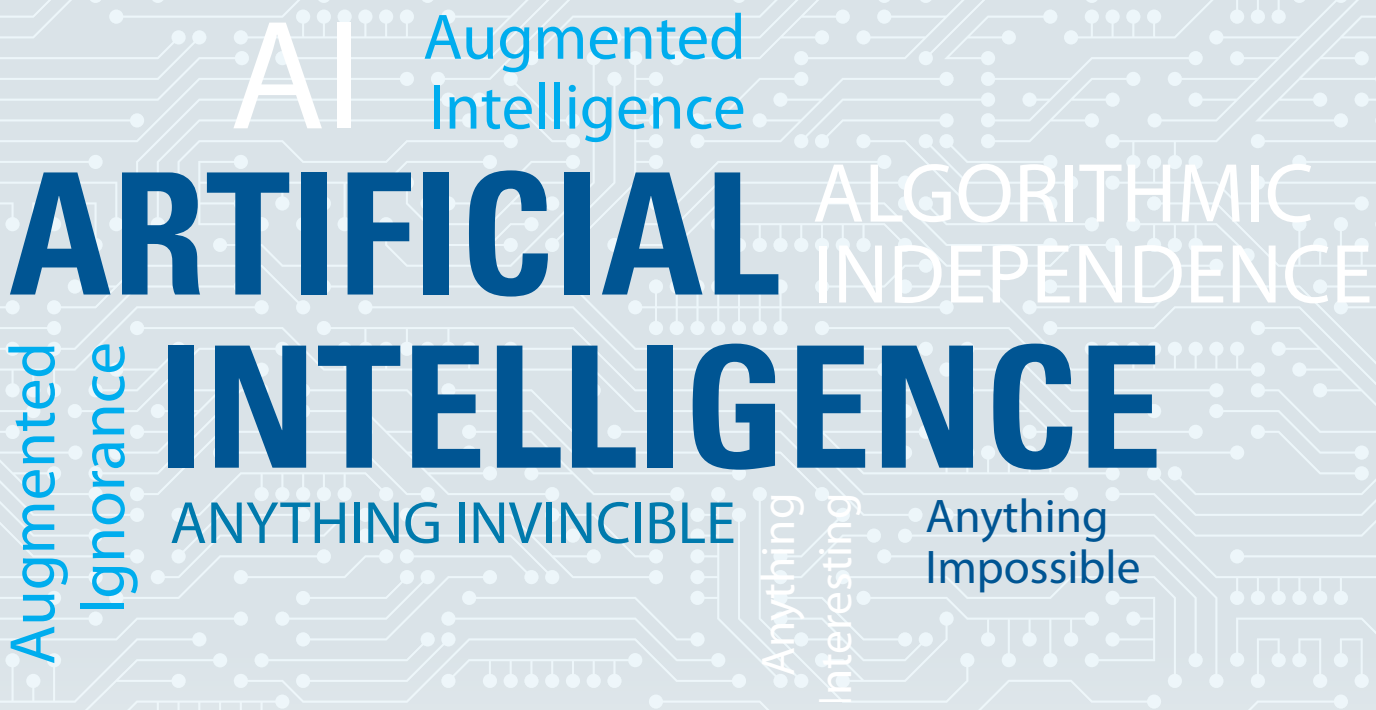
### Technological amplification of human abilities

Humans have always striven to make their lives easier by overcoming the limitations of their 'natural abilities'. The body and the brain, the human physiology and psychology, unmistakably limits what we are able to do. Finding ways to overcome these limits has been a human endeavour from the beginning of civilisations. Our natural intelligence has allowed us to develop technological 'amplifiers' and clever ways of using them. Physical abilities have been amplified with regard to power, speed, reach, precision, and endurance. Sensory abilities have been amplified with regard to size, distance, duration, scale, and thresholds. And mental or cognitive abilities have been amplified with regard to speed, quantity, and permanence. Early examples include the wheel, the lever, bow and arrow, the abacus, cuneiform writing, later followed by the telescope and microscope, various forms of engines, calculating machines, and computers.

Amplification has worked well for physical abilities but less so for cognitive abilities. This became obvious in the mid-1940s when the technologies were used to build partly self-controlling systems that were too fast and complicated for what the unaided human could manage. But the faster things happen, the more important it is to remain in control. This created a need to engineer for 'the human factor', often ironically by using even more technology as a substitute for what people could not do well enough. The dilemma was clearly stated by Paul Fitts:

"We begin with a brief analysis of the essential functions ... We then consider the basic question: Which of these functions should be performed by human operators and which by machine elements?" (Fitts, 1951).

In 2019, more than 70 years later, the FABEC (Functional Airspace Block Europe Central) expressed their




automation strategy in almost the same way: "Let ATCOs focus on the real, challenging work, to do what they are the best at, and leave the routine work to the machine".

System innovation is often driven by what humans cannot do rather than by what they can do. Technology is often used to replace or substitute human functions rather than to amplify or support them.

## Staying in control

Paul Fitts introduced the use of automation to replace human functions with technology. But this use failed to acknowledge the essential condition that it is necessary at all times to remain in control of what happens. Humans are aware of what they are doing and can imagine what the outcomes may be. Machines and technology can do neither. Digitalisation relies on highly effective but poorly understood algorithms, and AI is even worse in this respect. By replacing human functions with technology that is not fully comprehensible, control is gradually and irretrievably lost. Forty years ago, Earl Wiener noted that "It is highly questionable whether total system safety is always enhanced by allocating functions to automatic devices rather than human operators, and there is some reason to believe that flight-deck automation may have already passed its optimum point" (Wiener & Curry, 1980).

"Amplification has worked well for physical abilities but less so for cognitive abilities"

The problem with the substitution philosophy is that "the designer who tries to eliminate the operator still leaves the operator to do the tasks which the designer cannot think how to automate" (Bainbridge, 1983). The need to leave some parts of the work to humans is seen as a deplorable shortcoming of technological prowess, but also as something that soon will be remedied.

The unwavering technological optimism is one of the reasons why AI is the ultimate dream of automation and seen as the final (?) technological fix. Some even hope that we soon will reach 'the singularity' where machines will become truly intelligent and predict that Artificial General Intelligence will have arrived by 2040-50. Others see this as a hypothetical point in time at which technological growth irreversibly becomes uncontrollable.

But rather than using digitalisation as a substitution for what humans cannot do, it can also be used to amplify what they do well, to amplify intelligence. (This idea has a long history and was introduced as intelligence amplification by Ashby in 1961.) Instead of using digitalisation to replace what humans do badly, it should be used to support what humans do well, and stay in control.

## The pitfalls of technological solutioneering

Humans have always been attracted to promises of nice and easy solutions. There has been no shortage of these either in the context of work or in other areas of human endeavour. Human factors – or human factors engineering – may itself be seen as a 'nice and easy' solution, in the sense that it can overcome the problems arising from "the production of mechanical monstrosities which tax the capabilities of human operators" (Fitts, 1951, p. iv), or in today's terms be used to reduce human error, increase productivity, and enhance safety and comfort. Even if it usually is more difficult in practice than the theory. It is a sobering thought that Norbert Wiener – the creator of cybernetics – at the very beginning of digitalisation wrote about "gadget worshipers, who regard with impatience the limitations of mankind, and in particular the limitation consisting in man's undependability and unpredictability" (Wiener, 1964). Fifty years later the view was forcefully repeated when Morozov (2013) wrote about 'solutionism', defined as "an intellectual pathology that recognizes problems as problems based on just one criterion: whether they are 'solvable' with a nice and clean technological solution at our disposal".

The uncritical and overoptimistic belief in digitalisation is a form of solutionism, with the inevitable consequence that we will lose control of what we do. Problems are traditionally solved by breaking them into manageable parts which then are attacked and solved one by one, as if they could be dealt with in isolation. By doing so it becomes impossible to comprehend the consequences of what we are doing (Augmented Ignorance; Fujita, 2021), which in a vicious circle increases the need for 'nice and clean' solutions. Norbert Wiener characterised this situation as far back as 1954 by noting, "[W]e have modified our environment so radically that we must now modify ourselves in order to exist in this new environment".

## Algorithmic independence?

But the problem is not with digitalisation as such, since digitalisation is not even remotely intelligent. The problem is with how it is being used. Digitalisation basically relies on sophisticated algorithms that can solve well-defined problems with amazing ease. Because of the convenience they provide, we accept them one by one, enticed by the many advantages and oblivious of the drawbacks. (This is currently most obvious in the case of the social networks, but it is just as serious a concern in less conspicuous applications). So perhaps amplification of intelligence should be supplemented by a sixth interpretation of AI: Algorithmic Independence. Digitalisation all too easily creates algorithmic monstrosities that we then have to find yet another 'nice and easy' solution for.

"Rather than using digitalisation as a substitution for what humans cannot do, it can also be used to amplify what they do well, to amplify intelligence"

Erik Hollnagel is professor emeritus from LiU (S), ENSMP (F), and SDU (DK). His work focuses on unified system change and management. He is the author of more than 500 publications including articles from recognized journals, conference papers, and reports as well as twenty-eight books, and is still struggling to make sense of the blooming, buzzing confusion.

## References

Ashby, W. R. (1961). *An introduction to cybernetics.* Chapman & Hall Ltd.

Bainbridge, L. (1983). Ironies of automation. *Automatica, 19*(6), 775-779.

Fitts, P. M. (1951). *Human engineering for an effective air navigation and traffic control system.* Ohio State University Foundation Report, Columbus, OH

Fujita, Y. (2021). Personal communication.

Morozov, E. (2013). To save everything, click here: The folly of technological solutionism. New York: PublicAffairs.

Netland, K., & Hol, J. O. (1977). *Development of a computer and colour display based control room (OPCOM) for experimental operation of the Halden reactor (HPR-202)*. Halden, Norway: OECD Halden Reactor Project.

Wiener, N. (1954). *The human use of human beings*. Boston: Houghton Mifflin Co.

Wiener, N. (1964). *God and golem, Inc: A comment on certain points where cybernetics impinges on religion.* MIT Press.

Wiener, E. L. & Curry, R. E. (1980). Flight-deck automation: Promises and problems. *Ergonomics, 23*(10), 995-1011.