*Atti del XLIV Congresso Nazionale Sisfa - Firenze 2024* (2025), **43**_8, pp. 67–78
Doi:10.6093/978-88-6887-352-3-8

# Information transmission as artificial intelligence

Francesco Bianchini[1]

[1]University of Bologna, Bologna, francesco.bianchini@unibo.it.

*Abstract*: Artificial intelligence (AI) emerged from the convergence of several ideas and theories developed over a significant period in the first half of the twentieth century. Among these, information theory holds a prominent place, achieving its complete formulation with Claude Shannon in the 1940s. Shannon is considered one of the founding fathers of AI, being a key contributor to the 1956 Dartmouth seminar that officially launched the field. However, some aspects of this origin and the influence of information theory seem somewhat inconsistent from a broader perspective. Notably, information theory primarily deals with the transmission of signals rather than directly addressing the transmission of meaning, a key element of intelligence. At the same time, Nyquist's earlier work, which anticipated and facilitated the construction of a solid theory of information, concerned the transmission of intelligence, as per the author's terms. In my talk, I will attempt to trace the historical and theoretical connections between information theory and AI, exploring their mutual influence through the shared notion of "intelligence". I will demonstrate that while this shared usage does not signify a perfect conceptual overlap, it has nevertheless sparked numerous advancements in AI. These developments have far exceeded the initial expectations of those who first provided a technical and formal foundation for AI, and they have done so despite the great expectations generated in public opinion by early AI slogans.

*Keywords:* Artificial Intelligence, Communication Theory, Information, Transmission

## 1. Introduction

The history of artificial intelligence (AI) is a unique and complex discipline. Often, key aspects of AI's development cannot be fully understood by focusing solely on technical or theoretical milestones, nor by pinpointing specific moments in time. Much of AI's history is deeply rooted in theory and concerns periods before the concept of AI was even imagined or hypothesized. Furthermore, it requires addressing interdisciplinary questions, particularly those concerning theoretical frameworks from various fields. This blend of factors results in a diversity of ideas and methodologies, and given AI's relatively recent and rapid advancement compared to other scientific-technological disciplines, many relevant elements and their nuances have yet to be fully explored. The historical and theoretical evolution of AI, therefore, remains an area ripe for further investigation, offering insights into the epistemological present of the field while allowing for speculation on its future. The discovery of the artificial (Cordeschi, 2002) is, ultimately, a continuous rediscovery of the artificial – a perspective worth keeping in mind.

One of the major contributions to the development of AI is Information Theory, which reached maturity in the years leading up to AI's official inception. Less commonly known is the fact that the concept of transmitting intelligence had been explored at least two or three decades prior, within the research that culminated in Claude Shannon's complete formulation of information theory. Notably, some of Harry Nyquist's writings explicitly reference the transmission and communication of intelligence through technical systems. Recognizing the influence of information theory on AI highlights a fascinating bridge between these fields, especially when viewed from less traditional perspectives. This connection between information and intelligence offers a rich basis for further exploration, particularly as we seek to understand this union in light of contemporary advancements in AI and the increasing pervasiveness of the information notion.

In this chapter I will explore concepts related to intelligence and theories of communication and transmission, examining both pre-AI perspectives and ideas from around the inception of the field. In section 2, I will discuss Harry Nyquist's insights on the transmission of intelligence, focusing on the particular notion of intelligence at play. In section 3, I will consider the transition from Nyquist to Shannon from their background perspective, showing the different influences they had in relation to their notion of intelligence. In section 4, I will review Claude Shannon's reflections on the mathematical theory of communication and its connection to intelligence and AI. Finally, in section 5, I will present final remarks on how the concept of intelligence transmission continues to shape the evolution and analysis of AI technologies.

## 2. Nyquist and the transmission of intelligence

The earliest connection between the concepts of transmission and intelligence does not emerge in Shannon's systematic and foundational work on communication theory but rather appears a couple of decades earlier in seminal writings on communication by Harry Nyquist. While interest in telecommunications began even earlier, with Guglielmo Marconi's pioneering studies, it was only in the mid-20th century that communication theory received formal mathematical systematization through Claude Shannon's work. Nonetheless, these intermediate steps remain fundamental and appear to be underexplored, as the focus has often centered on Shannon" theories and his parallel involvement in the early stages of AI—two intertwined yet distinct paths in his thought, as we will later discuss.

Harry Nyquist, an electrical engineer with a doctorate in physics, made significant contributions to this field in several of his writings from the 1920s, particularly in *Certain Factors Affecting Telegraph Speed* (1924) and *Certain Topics in Telegraph Transmission Theory* (1928). One of Nyquist's major achievements in these works is his formulation of the theorem stating that the maximum number of independent pulses that can be transmitted through a telegraph channel in a given time unit is limited to twice the channel's bandwidth. This result, later adopted by Shannon, eventually became known as the Nyquist–Shannon sampling theorem (or Whittaker-Nyquist–Shannon sampling theorem). More relevant to our discussion, however, is that Nyquist explicitly addresses the concept of "transmission of intelligence" in these works. Although this notion of intelligence differs from the one central to AI, certain elements, on which I will focus, decisively anticipate its development.

In his 1924 paper, Nyquist begins by addressing "two fundamental factors entering into the maximum speed of transmission of intelligence by telegraph" (Nyquist, 1924, p. 324). These two factors, signal shaping and code selection, are crucial in determining the speed of intelligence transmission. The selection of the optimal wave shape for the transmission medium aims to maximize transmission speed while minimizing interference, a common concern in telecommunications theory. Of particular interest for our purposes is the second focus of the paper, which centers on code selection. Nyquist links this directly to the speed of intelligence transmission. To support this, he introduces a formula that relates line speed to the number of used values:

> A formula will first be derived by means of which the speed of transmitting intelligence, using codes employing different numbers of current values, can be compared for a given line speed, *i.e.,* rate of sending of signal elements. Using this formula, it will then be shown that if the line speed can be kept constant and the number of current values increased, the rate of transmission of intelligence can be materially increased. (Nyquist, 1924, p. 332)

The formula Nyquist introduces to capture the relationship between transmission speed and the number of values used is:

$$W = K \log m \tag{2.1}$$

where $W$ is the speed of transmission of intelligence, $m$ denotes the number of discrete values or symbols used in the transmission, and $k$ is a constant. This equation highlights the principle that increasing the diversity of symbols can enhance the transmission rate, provided the system can efficiently encode and decode these symbols.

Why is this formula interesting? It introduces an unconventional notion of "intelligence" in the context of information theory. As Nyquist describes it: "By the speed of transmission of intelligence is meant the number of characters, representing different letters, figures, etc., which can be transmitted in a given length of time assuming that the circuit transmits a given number of signal elements per unit time" (Nyquist, 1924, p. 332). This concept of intelligence, while distant from the one at the core of AI, is nonetheless related to it. Nyquist's formulation emphasizes a key aspect of information theory: the disregard for meaning. In this framework, what is transmitted is independent of the content or meaning of the message. Intelligence is reduced to the quantity of symbols and the rate at which they are transmitted, with speed explicitly tied to the volume of symbols.

Yet, it is precisely this connection with symbols that lends some legitimacy to referring to this as "intelligence". Nyquist appears to recognize, well before major advances in logic and computation theory, the deep link between intelligence and symbolic representation, even if his approach remains quantitative and indifferent to the symbols' meaning. This assumption, focused purely on the transmission rate and quantity of symbols, will later become a hallmark of Shannon's classical communication theory.

Using the proposed formula, Nyquist quantifies the amount of information that can be transmitted with a given number of signal elements based on the number of possible current values. This allows for a tabulated approach to explore how transmission speed can be increased as the number of current values grows:

| **Number of current values employed** | **Relative amount of intelligence which can be transmitted with a given number of signal elements** |
|---|---|
| 2 | 100 |
| 3 | 158 |
| 4 | 200 |
| 5 | 230 |
| 8 | 300 |
| 16 | 400 |

**Tab. 1:** possibilities of speeding up the transmission of intelligence (adapted from Nyquist, 1924, p. 333)

In broader terms, intelligence is defined quantitatively: as the number of current values increases, a greater number of signal elements can be transmitted. This operational definition aligns well with a discrete information transmission system. Additionally, it facilitates evaluating the optimal number of current values, balancing the benefit of transmitting more signal elements against the complexity introduced by increasing the current values:

> It should also be noted that whereas there is considerable advantage in a moderate increase in the number of current values, there is little advantage in going to a large number... If the line is subject to fluctuations so that the stronger currents at certain times become less in magnitude than the weaker currents at other times, it will be impossible to discriminate between the different current strengths making up the code, particularly if the fluctuations are rapid. (Nyquist, 1924, p. 333-335)

As the number of letters that can be transmitted increases, the speed of information transmission also rises, particularly when fewer signal elements are needed for each letter, due to a larger number of employed current values. However, increasing the number of current values yields diminishing returns, as illustrated in Table 1, and introduces technical risks, such as fluctuations that blur distinctions between current levels, interference, and limitations on usable power or voltage. Additionally, Nyquist states that "in general, whenever more than two current values are employed it is necessary to make the sending and receiving means more complicated and expensive" (Nyquist, 1924, p. 335).

Lastly, Nyquist's discussion of power, as it relates to intelligence in the closing section of his 1924 article, further clarifies the characteristics of the intelligence concept he uses. In evaluating other theories of information transmission via telegraphy and other means, Nyquist critiques certain unfounded assumptions, one being the idea that a waveform is ideally suited for both power and intelligence transmission. He argues this is inaccurate because "the transmission of intelligence inherently involves rapid and unpredictable changes in the current, whereas the transmission of power is best brought about by steady current, either direct or alternating. These two conditions are, of course, incompatible" (Nyquist, 1924, p. 340). This incompatibility underscores the fundamental distinction between power and intelligence. Since intelligence entails sudden and unpredictable shifts in current (enabled by the wave and its shape), whereas power is best transmitted through stable currents, intelligence emerges as qualitatively distinct from power. It is deeply intertwined with the unpredictability of signal changes and depends directly on the signal elements, indicating that intelligence is more an intrinsic aspect of the transmission process rather than merely something transmitted. This perspective anticipates a view of intelligence that aligns closely with the developments in artificial intelligence.

## 3. From Nyquist to Shannon: a different background

Although Nyquist and Shannon both built their work on similar foundational assumptions within the field of telecommunications, their different historical contexts shaped their approaches. The developments separating Nyquist's era from Shannon's are critical for understanding the transition toward a more comprehensive theory of information and the evolving role of the concept of intelligence. Despite these differences in background, notable continuities can be identified, not only in their contributions to information theory but also in their treatment of the notion of intelligence.

At the time of Nyquist, what was already known, his "past" and the groundwork for his contributions, had already been laid by earlier developments in radio and wireless transmission, most notably those of Guglielmo Marconi. While Marconi's innovations lacked a comprehensive mathematical framework, they paved the way for advancements that Shannon would later formalize into a general theory of information. Additionally, in the Anglo-Saxon world, particularly in the context of military practices, signal intelligence was already a well-established field. This involved the interception of telecommunications signals to gather data and infer enemy plans, a practice refined during the First World War (Winkler, 2009). While this use of "intelligence" differed from its broader conceptualizations, it nonetheless engaged with symbolic elements such as signals, codes, and codebreaking.

Simultaneously, the notion of intelligence was gaining prominence in other domains, particularly in psychology, where it was understood either as measurable through intelligence tests or as a set of diverse cognitive abilities applicable to various tasks. Perhaps even more influential, however, was the ongoing international debate surrounding the foundations of mathematics, spearheaded by Hilbert's program. This movement sought to establish a firm axiomatic basis for all mathematics, reflecting a period of great confidence in logic and formal systems. The elements of mathematical logic, which George Boole had earlier described as the "laws of thought", remained central to this intellectual climate. These converging elements - information, intelligence testing, and logical "laws of thought" - likely informed Nyquist's conceptualization of intelligence as he began to work on the problem of intelligence transmission.

What Nyquist could not yet foresee - his "future", available instead to Shannon - was a series of developments that would later prove pivotal for the emergence of AI. Among these was Gödel's Incompleteness Theorems of 1931, which dealt a critical blow to the ambitions of Hilbert's program. Gödel demonstrated that any formal system expressive enough to include arithmetic is inherently incomplete and that its consistency cannot be proven within itself. These results sparked widespread debate and introduced significant challenges in logic and mathematics, particularly concerning the reliability of mathematics and computation.

A key question that arose from Gödel's findings was: given these limitations, what functions can actually be computed within such systems? This problem eventually led to a groundbreaking solution, most famously articulated by Alan Turing in his 1936-1937 theory of computation. Turing showed that not all functions are computable; only those that can be computed by a Turing Machine, a theoretical device capable of mechanically executing a set of algorithmic instructions, fall within this scope. This concept of computation laid the groundwork for modern computer science and the eventual development of AI. The Turing thesis (or more precisely, the Church-Turing thesis, named also after Alonzo Church, who provided an equivalent formulation) is not a proven theorem but rather an accepted hypothesis due to the absence of counterexamples. It asserts that any function that can be computed algorithmically is computable by a Turing machine. While this is not the place to explore the topic in detail, it is worth emphasizing that the concept of intelligence underlying Turing's work, foundational to theoretical computer science, is rooted in the idea of a mechanical calculator or "computer", which could also describe a human performing similar tasks mechanically.

Turing reflected on these topics extensively, eventually formulating one of the foundational questions of AI: the question of "thinking machines" (Turing, 1950), or machines capable of demonstrating a form of intelligence (particularly through linguistic interaction). Two key elements emerge from his work. First, Turing revolutionized the concept of intelligence and thought, advocating for a mechanical understanding that could encompass all their aspects. He did so even before the formal emergence of AI, positioning himself as one of its most pivotal precursors. Second, Turing's expertise extended to codebreaking. During World War II, he worked at Bletchley Park, where he contributed to breaking the German Enigma code. To achieve this, Turing and Gordon Welchman developed "bombes", electromechanical devices designed to decrypt these codes. Furthermore, the first digital computers developed within two decades after Nyquist's groundbreaking work. They were electromechanical machines, such as the Z3 and Z4, or in other cases, like Colossus and ENIAC, were fully electronic. These early machines were eventually succeeded by the first computers capable of storing programs, marking the transition to modern computing.

Nyquist was unaware of many of these developments, and his background was far from the fields of theoretical and applied computer science, let alone the nascent field of AI. The idea of embedding intelligence in machines was entirely foreign to Nyquist's cultural and intellectual milieu – unlike Shannon, whose work laid foundational principles for AI. Nevertheless, we can identify key areas of overlap between their contributions. Both Nyquist and Shannon focused on signals and the theory and practice of transmission. Signals inherently involve a kind of symbolism, which, as we will see, takes on a more explicitly mathematical meaning in Shannon's work. Both also operated in a domain where coding played a central role – not just in transmitting information, but also in the context of wartime codebreaking, an activity deeply tied to decoding processes. Given these shared interests and overlapping elements, despite their differing backgrounds, one might reasonably expect them to have arrived at a somewhat similar notion of intelligence. However, as we shall see, this hypothesis is ultimately disproven. In this divergence in their conceptualizations, one could find significant implications for the future development of AI.

## 4. The intelligence of Shannon

Claude Shannon systematically addresses the problem of information transmission by proposing a comprehensive mathematical framework in his seminal essay *A Mathematical Theory of Communication* (Shannon, 1948). Three elements from the essay are particularly relevant to our discussion: first, the issue of *measuring* information in communication; second, the emphasis on *transmitting* information; and third, the notable absence – or near absence – of any reference to *intelligence*. After all this is consistent with Shannon explicitly stating in the introduction that, while the messages being communicated often carry meaning, the semantic aspects of communication are irrelevant to the field of communication engineering (Boden, 2006). In other words, the meaning of a message is considered extraneous to the problem at hand: its successful transmission.

> The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set* of possible messages. (Shannon, 1948, p. 379)

Regarding the question of measurement, Shannon begins by addressing the possibility of treating information quantitatively. This foundational inquiry underpins the subsequent mathematical treatment of communication and is encapsulated in his pivotal question: “Can we define a quantity which will measure, in some sense, how much information is ‘produced’ by such a process, or better, at what rate information is produced?” (Shannon, 1948, p. 389). Shannon’s mathematical formalization of information measurement ultimately leads to his well-known endorsement of the binary code as an effective system for transmitting information. This two-valued code aligns closely with the proposal previously advanced by Nyquist, who had argued for its efficiency in communication systems:

> In such a case one can construct a fairly good coding of the message on a 0, 1 channel by sending a special sequence, say 0000, for the infrequent symbol *A* and then a sequence indicating the *number* of *B*’s following it. This could be indicated by the binary representation with all numbers containing the special sequence deleted. All numbers up to 16 are represented as usual; 16 is represented by the next binary number after 16 which does not contain four zeros, namely 17 10001, etc. (Shannon, 1948, p. 398)

Binary digits would later play a crucial role in establishing the standard for digital computers, which served as the practical implementation of the Universal Turing Machine (Turing, 1936). These digits permeate computer science broadly, including AI, particularly in its quest to emulate certain aspects of brain function through algorithms and programs capable of exhibiting some sort of intelligent behavior. The parallel between digital electronic computers and the brain arises from their shared foundation in binary processing. The brain, like a computer, can be conceptualized as operating with binary circuits, based on the fundamental observation that a neuron either fires or does not fire. This insight underpins the view that the brain too, not just the mind, functions as a computational organ. It also forms the basis for the development of early artificial neural network prototypes, as exemplified in the seminal work of McCulloch and Pitts (1943).

In connection with the simulation of neurons, another foundational aspect of the development of neural networks - particularly the contemporary models that have been prominent since the 1980s (Rumelhart *et al.*, 1986) - is already anticipated in a structured manner within Shannon’s theory: the issue of coding. As we have seen, this theme was already present in Nyquist’s work, as it is fundamental to the transmission of information via symbolic systems. In Shannon's formulation, the concept of coding takes on characteristics that would later become central to neural networks. Specifically, neural networks can be understood as systems for information transmission, where incoming data must first be encoded and then decoded at the conclusion of the transmission process. This principle mirrors the mechanisms underlying modern neural networks, highlighting their dual role as computational models and information processing systems.

> We have yet to represent mathematically the operations performed by the transmitter and receiver in encoding and decoding the information. Either of these will be called a discrete transducer. The input to the transducer is a sequence of input symbols and its output a sequence of output symbols. The transducer may have an internal memory so that its output depends not only on the present input symbol but also on the past history. (Shannon, 1948, p. 394)

The fact that neural networks serve a dual purpose - as computational models of brain processes and as information-processing systems - often overshadows an important aspect: the mathematical and stochastic nature of information processing involves the transmission of information within the network. This concept was central to the field of cybernetics during the period when Shannon developed his mathematical theory of communication. This often-overlooked fact - that neural networks also function as information transmission systems (Herzog *et al.*, 2008) highlights the parallels between neural networks and systems of information transmission, including the brain itself. Within this framework, two key points emerge: the shared absence of inherent meaning in transmitted signals and the critical importance of encoding and decoding processes, which correspond to input and output functions. Such coding-related processes have become especially significant in the development of generative AI models, particularly large language models. In these systems, encoding, decoding, and autoencoding mechanisms underpin their impressive linguistic capabilities, enabling the high performance for which these models are now widely recognized (Vaswani *et al.*, 2017).

Another critical element that closely connects neural networks with information transmission systems is their treatment of noise, a further topic explicitly addressed by Shannon in his seminal 1948 paper. In telecommunications, noise interferes with messages, obstructing their accurate reconstruction during reception. Shannon proposed methods to mitigate noise and reduce uncertainty in transmission: "if the channel is noisy, it is not in general possible to reconstruct the original message or the transmitted signal with *certainty* by any operation on the received signal E. There are, however, ways of transmitting the information which are optimal in combating noise" (Shannon, 1948, p. 399). A similar approach to noise treatment is applied in deep neural networks to enhance their performance as the amount of information processed, i.e. transmitted within them, increases (Semenova *et al.*, 2022).

A notable aspect of Shannon's seminal 1948 paper is the absence of any reference to the concept of intelligence. This omission seems to mark a departure from Nyquist's earlier work, suggesting that the evolution from Nyquist to Shannon resulted in the deliberate exclusion of intelligence from the theory of communication. In his paper, Shannon mentions *intelligibility* only as a criterion for evaluating the likelihood of misinterpreting words by the recipient of a message[1] . However, this notion of intelligibility is distinct from the concept of intelligence and bears no direct relation to it.

Shannon was certainly not opposed to the concept of intelligence or its application in scientific and practical domains. In fact, he was one of the four organizers of the 1956 Dartmouth Seminar, widely regarded as the founding event of AI (McCarthy *et al.*, 1995). This seminar outlined key problems and areas of application for AI, including:

[1] "The structure of the ear and brain determine implicitly an evaluation, or rather a number of evaluations, appropriate in the case of speech or music transmission. There is, for example, an "intelligibility" criterion [referred to] the incorrectly interpreted words when message $x(t)$ is received as $y(t)$" (Shannon, 1948, p.428).

1. Automatic computers
2. How a computer can be programmed to use language
3. Neuron nets
4. The theory of computational complexity
5. Self-improvement
6. Abstractions
7. Randomness and creativity

Topics 3 (neuron nets) and 4 (computational complexity) seem particularly aligned with Shannon's interests. However, this raises an intriguing question: why is Shannon considered one of the founding fathers of AI when the concept of intelligence was conspicuously absent from his mathematical treatment of communication theory?

Shannon played a pivotal role in the foundational stages of AI, as highlighted in the text of the proposal. This was due to his groundbreaking work in information theory, his interest in learning machines, and his co-editorship with John McCarthy of *Automata Studies*, a volume in *Annals of Mathematical Studies*[2]. In the early 1950s, Shannon developed chess-playing programs and constructed the Theseus Machine, a mechanical mouse capable of navigating a maze and learning (on the basis of many relays) the route to the exit through trial and error. These contributions placed Shannon at the forefront of the emerging field of research that would later be named "artificial intelligence". Even before the term itself was coined, this field focused on leveraging newly invented calculators to perform tasks traditionally associated with human intelligence (Cordeschi, 2007). Interestingly, and consistent with Shannon's approach to communication theory, the word *intelligence* is notably absent even in his seminal 1950 paper on chess and computers (Shannon, 1950).

The absence of the term *intelligence* does not diminish Shannon's role in the foundation of AI. Instead, it highlights how the focus of early research centered on nascent computers and thinking (to borrow from Turing's conceptual framework) was not on intelligence as a whole, but rather on its specific components: reasoning, language, memory, and learning. Information theory was already recognized as integral to these studies. In the Dartmouth proposal, Shannon outlined two primary research directions: exploring the approach to automata by modeling the relationship between the environment and the brain; and applying concepts from information theory to computing machines and brain models. Regarding the first theme, which was the second one he proposed, Shannon referred to *mechanized intelligence* signaling that the time had come to tackle the problem of artificial systems through the lens of intelligence:

> Often in discussing mechanized intelligence, we think of machines performing the most advanced human thought activities – proving theorems, writing music, or playing chess. I am proposing here to start at the simple and when the environment is neither hostile (merely indifferent) nor complex, and to work up through a series of easy stages in the direction of these advanced activities (McCarthy *et al.*, 1995).

Regarding the first topic – the application of concepts from information theory to computing machines and brain models, a field in which Shannon was undoubtedly most renowned – it is particularly noteworthy that he begins with the question of information transmission and progresses toward the concept of information theory networks. In these networks, the flow – the transmission – of information is not linear but instead complex and loop-based. This transition directly connects the idea of information transmission to the concept of (neural) networks, which are capable of modeling aspects of brain function:

[2] "C. E. Shannon, Mathematician, Bell Telephone Laboratories. Shannon developed the statistical theory of information, the application of propositional calculus to switching circuits, and has results on the efficient synthesis of switching circuits, the design of machines that learn, cryptography, and the theory of Turing machines. He and J. McCarthy are co-editing an *Annals of Mathematics Study* on "The Theory of Automata" ". (McCarthy *et al.*, 1995)

> A basic problem in information theory is that of *transmitting information* reliably over a noisy channel. An analogous problem in computing machines is that of reliable computing using unreliable elements. This problem has been studied by von Neumann for Sheffer stroke elements and by Shannon and Moore for relays, but there are still many open questions. The problem for several elements, the development of concepts similar to channel capacity, the sharper analysis of upper and lower bounds on the required redundancy, etc. are among the important issues. Another question deals with the theory of information networks where information *flows* in many closed loops (as contrasted with the simple one-way channel usually considered in communication theory). Questions of delay become very important in the closed loop case, and a whole new approach seems necessary. This would probably involve concepts such as partial entropies when a part of the past history of a message ensemble is known. (McCarthy *et al.*, 1995)

It is interesting to observe how Shannon connects the artificial modeling of the brain to the concept of information, even though information alone seems to be not enough. What must be taken into account is the dynamic process of *information transmission*.

Retracing Shannon's path, we can affirm that information theory, as expanded in the book co-authored with Weaver the following year (Shannon & Weaver, 1949), is pivotal for understanding information systems - both artificial systems, such as computers, and intelligent natural systems, including humans and animals, as viewed through the lens of cybernetics. Moreover, "information devices", such as computers, represent our most advanced tools for simulating natural information systems. The shared conceptual foundation for these endeavors lies not merely in the concept of information itself but in the notion of *information transmission*. Shannon's groundbreaking contributions to the mathematical standardization and modeling of information theory were essential for establishing a science of the artificial, one that conceptualizes intelligent systems as information transmission systems. This work laid the groundwork for developments that were unimaginable at the time, even if, as we have seen, somehow Shannon mentioned them.

## 5. Conclusion: Shannon, Nyquist, and the symbol problem

The history of AI frequently emphasizes the significance of Shannon's contributions to information theory (Cordeschi, 2002; Boden, 2006). While some authors give less attention to the role of information theory, they highlight other equally important aspects of Shannon's intellectual legacy. For instance, as early as 1937, in his thesis at Yale University, Shannon demonstrated that binary arithmetic and Boolean algebra could simplify telephone switching circuits and, conversely, that such circuits could perform Boolean operations – a breakthrough that proved fundamental for computer design (Nilsson, 2010).

From a more contextual perspective, one might note that Shannon's contribution to the founding of AI was perhaps driven more by his interest in developing programs capable of playing chess and his maze-solving mechanical mouse system. Nonetheless, Shannon's ideas on information became pervasive across fields like computer science and psychology. This is evident in approaches such as information-processing psychology (IPP), where the concept of information flow is central. Furthermore, information remains foundational in contemporary theories on consciousness, such as Chalmers' work on consciousness and Tononi's Integrated Information Theory (IIT).

It is important to remember that, according to Shannon, information primarily referred to its transmission, independent of meaning – an aspect that seems to distance his work from the early developments in AI, which were highly symbolic and less mathematical in focus. Nonetheless, information has remained a central concept in the framework connecting computers and brains. This connection is grounded in the idea that, before being systems of information processing, both are fundamentally systems of information transmission. This transmission occurs according to specific rules and yields significant outcomes, one of which is intelligence. In these systems – whether computers or brains – the transmission of information effectively becomes the transmission of intelligence, which brings us back to what could be described as the “Nyquist symbol problem”.

Nyquist’s problem focused on increasing the speed of information transmission in telecommunication systems. He recognized that information, through the careful selection of signals, could be seen as imbued with intelligence, as the optimal choice of signals was linked to the most effective capacity for transmitting symbols. Modern, advanced neural networks are algorithms capable of processing vast quantities of data. In fact, they remain systems in which information is transmitted after being encoded and before being decoded. In the future, it might become possible to determine, in a more precise, explainable, and ultimately symbolic manner, the nature of the information transmitted and processed by these networks. Such advancements could revisit and build upon Nyquist’s insights, viewing this transmission as a transfer of intelligence.

## Acknowledgments

I would like to extend my thanks to Flavia Marcacci for her excellent suggestion regarding the topic addressed in this chapter and for the numerous fruitful discussions that greatly enriched this work.

## Bibliography

Boden, M. (2006). *Mind as Machine: a History of Cognitive Science*. Oxford: Clarendon Press.

Cordeschi, R. (2002). *The Discovery of the Artificial*. Dordrecht: Kluwer Academic Publishers.

Cordeschi, R. (2007). “AI turns fifty: revisiting its origins”, *Applied Artificial Intelligence*, 21(4–5), pp. 259–279.

Herzog, A. *et al.* (2008), “Information transmission between recurrent neural networks by sparsely electrical connections”, *BMC Neuroscience*, 9(Suppl 1), P17.

McCarthy, J. *et al.* (1995). *A proposal for the dartmouth summer research project on artificial intelligence*. Available at jmc.stanford.edu (accessed 16 November 2024).

McCulloch, W. & Pitts, W. (1943). “A logical calculus of the ideas immanent in nervous activity”, *Bulletin of Mathematical Biophysics*, 5, pp. 115-133.

Nilsson, N. J. (2010). *The Quest for Artificial Intelligence: a History of Ideas and Achievements*. Cambridge: Cambridge University Press.

Nyquist, H. (1924). “Certain factors affecting telegraph speed”, *The Bell System Technical Journal*, 3(2), pp. 324-346.

Nyquist, H. (1928). “Certain Topics in Telegraph Transmission Theory”, *Transactions of the American Institute of Electrical Engineers*, 47(2), pp. 617-644.

Rumelhart, D. E., McClelland, J. L. & the PDP Research Group (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge: MIT Press.

Semenova, N., Larger, L. & Brunner, D. (2022). “Understanding and mitigating noise in trained deep neural networks”, *Neural Networks*, 146, pp. 151-160.

Shannon, C.E. (1948). “A mathematical theory of communication”, *The Bell System Technical Journal*, 27(3), pp. 379-423.

Shannon, C.E. (1950). “XII. Programming a computer for playing chess”, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 41(314), pp. 256–275.

Shannon, C.E. & Weaver, W. (1949). *The Mathematical Theory of Communication*. Champaign: University of Illinois Press.

Turing, A.M. (1936). "On Computable Numbers, with an Application to the Entscheidungsproblem", *Proceedings of the London Mathematical Society*, 42(1), pp. 230–265.

Turing, A.M. (1950). "Computing Machinery and Intelligence", *Mind*, 59, pp. 433–460.

Vaswani, A. *et al.* (2017). "Attention is all you need", in *Proceedings of 31$^{s}$t International Conference on Neural Information Processing Systems*, Red Hook: Curran Associates, pp. 6000-6010.

Winkler, J.R. (2009). "Information Warfare in World War I", *The Journal of Military History*, 73, pp. 845-867.