Información de Shannon: Bases teóricas, articulaciones conceptuales e interpretaciones

Shannon's Information: Theoretical Basis, Conceptual Articulations and Interpretations

Olimpia Lombardi

CONICET – Universidad de Buenos Aires (Argentina)
https://orcid.org/0000-0003-2204-7902
olimpiafilo@gmail.com

Resumen: Aunque el término 'información' apareció bajo diferentes formas a lo largo de la historia de la humanidad, fue durante el siglo XX que adquirió sus connotaciones actuales, vinculadas a la comunicación y la tecnología. En este contexto, Claude Shannon es considerado el padre fundador de la ciencia de la información tal como se la conoce actualmente. En el presente artículo nos centraremos no sólo en los aspectos técnicos de la teoría de Shannon, sino también en algunas articulaciones de la teoría, algunas de las cuales han sido relevantes en los debates filosóficos sobre la interpretación del concepto de información. El propósito es ofrecer un panorama equilibrado que incluya los múltiples rostros de la noción de información, una noción que se ha convertido en un hito de nuestra cultura contemporánea.

Palabras clave: Shannon, entropía, complejidad algorítmica, información cuántica, interpretaciones del concepto de información.

Abstract: Although the term 'information' appeared under different forms all along the history of mankind, it is during the twentieth century that it acquired its present-day connotations, linked to communication and technology. In this context, Claude Shannon is rightly considered the founding father of the science of information as known nowadays. In the present article, we will focus not only on the technical aspects of Shannon's theory, but also on some articulations of the theory, some of which have been relevant in the philosophical debates about the interpretation of the concept of information. The purpose is to offer a balanced panorama that includes the many faces of the notion of information, which has become a landmark of our contemporary culture.

Keywords: Shannon, entropy, algorithmic complexity, quantum information, interpretations of the concept of information.

Ápeiron. Estudios de filosofía, monográfico «Diez ensayos de filosofía de la información», n.º 12, 2020, pp. 37–51,
Madrid-España (ISSN 2386 – 5326)
http://www.apeironestudiosdefilosofia.com/

Recibido: 02/11/2019 Aceptado: 10/01/2020

1. Introducción

Si bien el término 'información' aparece repetidamente en la historia del pensamiento desde la Antigüedad (véase Adriaans 2013), es recién en el siglo XX que adquiere sus connotaciones actuales, ligadas a los conceptos de datos, comunicación y tecnología. No obstante, la fuerte presencia de la noción de información en nuestros días no conlleva acuerdo en cuanto al significado del concepto. En efecto, existen al menos dos modos generales de concebir la información: desde una perspectiva semántica y desde una perspectiva matemática. El concepto de *información semántica* (Bar-Hillel y Carnap 1953, Bar-Hillel 1964, Floridi 2011, 2013) se relaciona estrechamente con nociones como referencia, significado y representación: la información semántica tiene contenido e intencionalidad. Sin embargo, la información puede ser cuantificada con independencia de su contenido, convirtiéndose en un recurso de gran utilidad en las disciplinas científicas matematizadas.

El trabajo seminal para el concepto *matemático* de información es el artículo donde Claude Shannon (1948) introduce un formalismo preciso diseñado para resolver ciertos problemas tecnológicos específicos en la ingeniería de comunicaciones (véase también Shannon y Weaver, 1949). En términos generales, la entropía de Shannon se refiere a las propiedades estadísticas de un sistema y a las correlaciones entre los estados de dos sistemas, independientemente del significado y de cualquier contenido semántico de esos estados. Actualmente, la teoría de Shannon es un ingrediente básico en la capacitación de los ingenieros en comunicaciones.

2. Los orígenes históricos de la teoría de Shannon

Ciertamente, Shannon merece ser considerado el padre fundador de la teoría de la información; pero esto no significa que el concepto mismo de información, en su sentido matemático y cuantitativo, haya aparecido por primera vez en su trabajo de 1948.

Ralph Hartley fue un investigador en electrónica estadounidense, que trabajó en los Bell Laboratories, la división de investigación y desarrollo de la Bell Telephone Company, en ese entonces el principal proveedor de servicios telefónicos en EE. UU. En 1927, en el Congreso Internacional de Telegrafía y Telefonía, Hartley presentó un trabajo titulado "Transmission of Information", que se centraba explícitamente en cuantificar la información y establecer la capacidad de los sistemas físicos para transmitir información. Esa presentación se convirtió en un artículo, publicado en 1928 en el *Bell System Technical Journal*, donde Hartley emprende explícitamente la tarea de definir una "medida cuantitativa de información [...] basada en consideraciones físicas en lugar de psicológicas" (Hartley 1928: 535). En particular, considera un sistema que produce s estados distinguibles en secuencias de longitud n, de modo tal que el número de secuencias distinguibles es s^n. Sobre esta base, propone "tomar como nuestra medida práctica de la información el logaritmo del número de posibles secuencias de símbolos" (Hartley 1928: 540), es decir,

$$H = \log s^n = n \log s$$

Como veremos en la próxima sección, esta medida de información es proporcional a la entropía de Shannon de una fuente con estados equiprobables, cada uno de ellos con una probabilidad $1/n$. Como el propio Shannon más tarde reconocerá en la primera página de su famoso artículo de 1948, esto significa que su formalismo generaliza la propuesta de Hartley al caso de una fuente que produce estados no equiprobables.

Shannon nació en Michigan en 1916, nieto de un inventor y primo lejano de Thomas Edison. Estudió ingeniería eléctrica y matemáticas en la Universidad de Michigan, graduándose en 1936, y luego se convirtió en asistente de investigación en el departamento de ingeniería eléctrica del Instituto de Tecnología de Massachusetts. Su tesis de maestría, *A Symbolic Analysis of Relay And Switching Circuits*, publicada en 1938, mostró la posibilidad de resolver problemas simplemente manipulando dos símbolos, 0 y 1, en un circuito eléctrico;

en particular, demostró que el funcionamiento de interruptores y relés en circuitos electrónicos puede representarse mediante un álgebra booleana.

Después de obtener su título superior en ingeniería eléctrica y un doctorado en matemáticas en el Instituto de Tecnología de Massachusetts en 1940, en 1941 el Comité de Investigación de Defensa Nacional contrató a Shannon para trabajar en los Bell Laboratories. Durante la Segunda Guerra Mundial, su investigación se centró en criptografía. Al final de la guerra, en septiembre de 1945, preparó un memorando clasificado para los Bell Laboratories titulado "A Mathematical Theory of Cryptography". Este trabajo, cuando fue desclasificado, se publicó en 1949 como el documento "Communication Theory of Secrecy Systems", que ya incorporaba muchos de los conceptos y formulaciones matemáticas del artículo clásico de 1948. De hecho, la investigación de Shannon sobre sistemas de cifrado inspiró su Teoría de la Información cuando se dio cuenta de que, así como los códigos protegen la información del espionaje, los códigos también pueden proteger la información del ruido indeseable. Varios años después, Shannon reconoció la fuerte influencia que este trabajo durante la guerra ejerció sobre sus resultados posteriores acerca de la comunicación de información.

Según Timothy Glander (2000), la agenda del gobierno de EE. UU. financió y dominó la investigación en comunicaciones durante la guerra y la posguerra. Este hecho ayuda a explicar la amplia repercusión del trabajo de Shannon: es muy plausible suponer que una teoría que prometía un tratamiento matemático de la información fuera bienvenida en la atmósfera de la Guerra Fría, donde la preservación y la gestión de la información se habían convertido en preocupaciones prioritarias.

3. Bases teóricas de la teoría de Shannon

Antes de la teoría de la información, para la comunicación a distancia se utilizaban señales analógicas: los mensajes se transformaban en señales de voltaje variable y se enviaban a lo largo de un cable; esas señales se medían en el receptor y para reconvertirse nuevamente en el mensaje original. Aunque este método se puede utilizar en la comunicación a través de distancias cortas, para largas distancias se torna inutilizable: una señal eléctrica analógica que viaja a lo largo de un cable se debilita y queda enmascarada por el ruido aleatorio del entorno. Estas dificultades pueden superarse si se utilizan señales digitales y el mensaje se convierte, letra por letra, en un código digital, por ejemplo, un código formado por 0s y 1s, donde cada 0 está representado por una señal corta de bajo voltaje y cada 1 está representado por una señal corta de alto voltaje. Por supuesto, estas señales digitales sufrirán los mismos problemas que una señal analógica, es decir, debilitamiento y ruido. Sin embargo, dado que son estados muy diferentes, se pueden distinguir incluso después de un deterioro importante, por lo cual a pesar de ello es posible recuperar el mensaje original. Esta es la idea tecnológica que motivó la Teoría de la Información.

Con su trabajo "The Mathematical Theory of Communication" (1948), Shannon ofreció resultados precisos sobre los recursos necesarios para una codificación óptima y una comunicación sin errores. Este documento fue seguido inmediatamente por muchos trabajos de aplicación a campos como radio, televisión y telefonía. La teoría de Shannon fue luego matemáticamente axiomatizada en el ámbito de la teoría de la probabilidad (Khinchin 1957). Según Shannon (1948; véase también Shannon y Weaver 1949), un sistema de comunicación general consta de cinco partes:

— Una *fuente S*, que genera el *mensaje* que se recibirá en el destino.
— Un *transmisor T*, que convierte el mensaje generado en la fuente en una *señal* apta para la transmisión. En los casos en que la información se codifica, este sistema también implementa la codificación.
— Un *canal CH*, es decir, el medio utilizado para transmitir la señal del transmisor al receptor.
— Un *receptor R*, que reconstruye el mensaje a partir de la señal.
— Un *destino D*, que recibe el mensaje ya reconstruido.

$$\boxed{S} \xrightarrow{\text{mensaje}} \boxed{T} \xrightarrow{\text{señal}} \boxed{CH} \xrightarrow{\text{señal}} \boxed{R} \xrightarrow{\text{mensaje}} \boxed{D}$$

La fuente *S* es un sistema con un rango de posibles estados $s_1,...,s_n$ generalmente llamados *letras*, cuyas probabilidades de ocurrencia son $p(s_1),...,p(s_n)$ repectivamente (aquí se considera el caso discreto, pero todas las definiciones pueden extenderse al caso continuo; véase, por ejemplo, Cover y Thomas 1991). Un *mensaje* es una secuencia de estos estados. Análogamente, el destino *D* es un sistema con un rango de estados posibles $d_1,...,d_m$, con probabilidades $p(d_1),...,p(d_m)$ respectivamente. Shannon define la *entropía de la fuente S*, $H(S)$, y la *entropía del destino D*, $H(D)$, como

$$H(S) = -\sum_{i=1}^{n} p(s_i) \log p(s_i) \qquad\qquad H(D) = -\sum_{j=1}^{m} p(d_j) \log p(d_j)$$

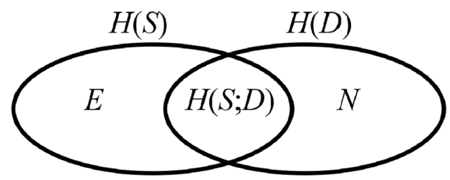
En su artículo original, Shannon (1948: 349) explica la conveniencia del uso de una función logarítmica en la definición de las entropías: su utilidad reside en que muchos parámetros importantes en ingeniería varían linealmente con el logaritmo del número de posibilidades; a su vez, es intuitivo porque solemos medir magnitudes por comparación lineal con unidades de medida; y es matemáticamente conveniente porque muchas operaciones matemáticas en términos de logaritmos son más simples que en términos de números de posibilidades. A su vez, la elección de una base logarítmica equivale a la elección de una unidad para medir la información. Si se usa base 2, como sucede casi siempre, la unidad resultante se denomina *bit* (contracción de *binary unit*) Un bit es la cantidad de información obtenida cuando se especifica una de dos alternativas igualmente probables. Pero también puede utilizarse el logaritmo natural, y en este caso la unidad de medida es el *nat* (contracción de *natural unit*). Y cuando se usa el logaritmo en base 10, la unidad es el *Hartley*.

Por el modo en que son definidas −como promedios ponderados por la probabilidad−, las entropías de la fuente $H(S)$ y del destino $H(D)$ denotan *cantidades medias* de información por letra generada en la fuente y recibida en el destino, respectivamente. Esto no significa que no puedan definirse cantidades de información individuales, correspondientes a la ocurrencia de un estado particular de la fuente. Sobre esta base, $H(S)$ y $H(D)$ pueden expresarse como

$$H(S) = \sum_{i=1}^{n} p(s_i)\, I(s_i) \qquad\qquad H(D) = \sum_{j=1}^{m} p(d_j)\, I(d_j)$$

donde $I(s_i) = \log(1/p(s_i))$ y $I(d_j) = \log(1/p(d_j))$ son las cantidades individuales de información correspondientes a la ocurrencia del estado particular s_i en el origen y del estado particular d_j en el destino, respectivamente (véase, por ejemplo, Lombardi 2004, Bub 2007, Lombardi, Holik y Vanni 2016a). Desde esta perspectiva, una sola letra de la fuente es un tipo particular de mensaje y transmite información. Por ejemplo, Hartley afirma explícitamente que, en el caso de alternativas equiprobables, "[l]a información asociada con una sola selección es el logaritmo del número de símbolos disponibles" (Hartley 1928: 541). No obstante, no debe olvidarse que el interés de la teoría de Shannon se centra en cantidades medias de información y no en cantidades de información individuales.

La relación entre las entropías de la fuente *S* y del destino *D* puede representarse de un modo sencillo en el siguiente diagrama (véase, por ejemplo, Cover y Thomas 1991: 20):



donde:

- $H(S;D)$ es la *información mutua*: la cantidad media de información generada en la fuente *S* y recibida en el destino D.
- E es la *equivocidad*: la cantidad media de información generada en *S* pero no recibida en *D*.
- N es el *ruido*: la cantidad media de información recibida en *D* pero no generada en *S*.

Como muestra el diagrama, la información mutua se puede calcular como

$$H(S;D) = H(S) - E = H(D) - N$$

Del diagrama también puede observarse que la equivocidad *E* y el ruido *N* son medidas de la dependencia entre la fuente *S* y el destino *D*:

- Si *S* y *D* son completamente independientes, los valores de *E* y *N* son máximos ($E = H(S)$ y $N = H(D)$), y el valor de $H(S;D)$ es mínimo ($H(S;D) = 0$).
- Si la dependencia entre *S* y *D* es máxima, los valores de *E* y *N* son mínimos ($E = N = 0$), y el valor de $H(S;D)$ es máximo ($H(S;D) = H(S) = H(D)$).

Los valores de la equivocidad *E* y del ruido *N* son funciones no sólo de la fuente *S* y del destino *D*, sino también del canal de comunicación *CH*, que está definido por la matriz $\left[p(d_j/s_i) \right]$, donde $p(d_j/s_i)$ es la probabilidad condicional de que ocurra d_j en el destino *D* dado que ocurrió s_i en el fuente *S*, y los elementos en cualquier fila suman 1. A su vez, la capacidad *C* del canal se define como

$$C = \max_{p(s_i)} H(S;D)$$

donde el máximo se toma sobre todas las distribuciones $p(s_i)$ posibles en la fuente. *C* mide la máxima cantidad de información que se puede *transmitir* a través del canal de comunicación *CH*.

Los dos resultados más importantes obtenidos por Shannon son los teoremas conocidos como Primer Teorema de Shannon y Segundo Teorema de Shannon. Para comprenderlos, es necesario introducir la etapa de codificación en el proceso de transmisión de la información. El transmisor *T* codifica los mensajes de la fuente *S*, y esto equivale a realizar la conversión entre el alfabeto de la fuente, $A_S = \{s_1, ..., s_n\}$, y el alfabeto-código del transmisor que, en general, suele ser un alfabeto binario, $A_T = \{0,1\}$, cuyos miembros se denominan *símbolos*. La secuencia de símbolos producidos por el transmisor *T* y que ingresan al canal *CH* es la *señal*. El alfabeto-código del transmisor puede implementarse físicamente por medio de sistemas clásicos de dos estados, que algunos autores han denominado *cbits* (Caves y Fuchs 1996).

En el contexto de la teoría de Shannon, la *codificación* es un mapeo desde el alfabeto de la fuente A_S al conjunto de *cadenas de símbolos* de longitud finita del alfabeto-código A_T −en general, cadenas de 0s y 1s−, también llamadas *palabras de código*. En general, las palabras de código no tienen la misma longitud: la palabra de código w_i, correspondiente a la letra s_i, tiene una longitud l_i. Esto significa que la codificación es un mapeo desde ítems de longitud fija a ítems de longitud variable. Por lo tanto, la *longitud promedio L* de las palabras de código en el proceso de codificación se puede definir como

$$L = \sum_{i=1}^{n} p(s_i)\, l_i$$

L indica la compacidad del código: cuanto menor es L, más eficiente es la codificación, es decir, se necesitan menos recursos para codificar los mensajes que produce la fuente. El Primer Teorema de Shannon, también llamado *Noiseless-Channel Coding Theorem*, demuestra que, para mensajes suficientemente largos, existe un proceso de codificación óptimo tal que la longitud promedio L de las palabras de código es igual al valor de la entropía de la fuente $H(S)$. En otras palabras, $H(S)$ mide el valor óptimo compresión de los mensajes de la fuente. Esto sugiere una estrategia natural para la codificación: las letras más probables deben codificarse con palabras de código −cadenas de 0s y 1s− más cortas, mientras que las letras menos probables pueden codificarse con palabras de código más largas. De este modo, los mensajes típicos quedan codificados por una secuencia binaria de longitud $NH(S)$, en general más corta que la longitud N del mensaje original.

Por otro lado, a principios de la década de 1940 se pensó que el aumento de la velocidad en la transmisión de información a través de un canal de comunicación siempre aumentaría la probabilidad de error. El Segundo Teorema de Shannon, también conocido como *Noisy-Channel Coding Theorem*, demostró que ese supuesto no era correcto, siempre que la velocidad de comunicación se mantuviera por debajo de la capacidad C del canal. La capacidad del canal, por lo tanto, mide la cantidad máxima de información que se puede transmitir a través del canal y que puede recuperarse en el destino con una probabilidad de error muy baja.

4. Articulaciones conceptuales de la teoría de Shannon

Si bien la teoría de Shannon es el *locus classicus* para el tratamiento matemático de la información, el concepto se ha diversificado en otros formalismos, cada uno correspondiente a un concepto diferente (por ejemplo, información de Fisher, información de Rényi, etc.). Aquí mencionaremos sólo dos casos, el primero por su importancia conceptual y el segundo por su relevancia física.

4.1. Información algorítmica

En la teoría de Shannon, las cantidades de información son promedios: sólo se pueden calcular cuando los mensajes son elementos de un conjunto y cada elemento tiene su propia probabilidad. En su famoso artículo, Shannon enfatiza que su noción se refiere a la comunicación de mensajes seleccionados entre un grupo de mensajes producidos por una fuente; por lo tanto, la fuente "debe estar diseñada para operar para cada selección posible, no sólo la que realmente se elegirá, ya que esto es desconocido en el momento del diseño" (Shannon 1948: 379). En otras palabras, la información de Shannon está determinada por las características de la fuente, y no por las características de los objetos que son sus resultados. En cambio, la información algorítmica o información de Kolmogorov, también llamada complejidad algorítmica (Solomonoff 1964, Kolmogorov 1965, 1968, Chaitin 1966), brinda una medida de información para mensajes individuales, tomados en sí mismos. Como afirma Andréi Kolmogorov, su objetivo es complementar el trabajo de Shannon: "Nuestra definición de la cantidad de información tiene la ventaja de que se refiere a objetos individuales y no a objetos tratados como miembros de un conjunto de objetos con una distribución de probabilidad dada en él. La definición probabilística se puede aplicar de manera convincente a la información contenida, por ejemplo, en un flujo de telegramas de felicitación. Pero no estaría claro cómo aplicarlo, por ejemplo, a una estimación de la cantidad de información contenida en una novela o en la traducción de una novela a otro idioma en relación con el original." (Kolmogorov 1983: 29).

En la teoría de la complejidad algorítmica, el interés es encontrar el número mínimo de bits a partir del cual se puede reconstruir efectivamente un mensaje particular: la cuestión básica es entonces la compresión final de mensajes individuales. La idea que subyace a la teoría es que la descripción de algunos mensajes se puede comprimir considerablemente, si exhiben suficiente regularidad. Sobre esta base, dada una cadena finita

x de letras, la *complejidad algorítmica* de x se define como la longitud del programa de computadora más corto que, en una máquina Turing, imprime x y luego se detiene. Intuitivamente, la secuencia

010101010101010101010101010101010101010101010101010101010101

es simple: puede ser construida por el programa "Imprimir 01 treinta veces; detener". Por el contrario, supongamos que la secuencia

11011110011101010001011001010100010110111100010111001010011

es verdaderamente aleatoria, generada por lanzamientos de monedas o por otro fenómeno completamente azaroso. En este caso, no se puede comprimir, es decir, no hay mejor programa para construirlo que simplemente "Imprimir ...", donde los puntos representan la secuencia completa.

Vale la pena insistir en la diferencia entre la información de Shannon y la complejidad algorítmica: para cualquier fuente con dos estados, la entropía de Shannon es como máximo 1 bit, pero los estados se pueden elegir con una complejidad algorítmica arbitrariamente alta. De hecho, la complejidad algorítmica considera sólo el mensaje en sí mismo para determinar el número de bits en la versión completamente comprimida, independientemente de la forma en que se generó el mensaje. El precio a pagar por esta estrategia se deriva del hecho de que no se cuenta con una función recursiva de cálculo: la complejidad algorítmica no es computable en general. Esto significa que no existe ningún programa de computadora que, cuando recibe cualquier secuencia arbitraria como entrada, emite la complejidad algorítmica de esa secuencia y luego se detiene. Por lo tanto, aunque podemos comprimir una secuencia de manera óptima almacenando o transmitiendo el programa más corto que la genera, sólo es posible encontrar dicho programa en casos particulares.

A pesar de la diferencia conceptual entre la entropía de Shannon y la complejidad algorítmica, existe una relación significativa entre las dos magnitudes: dada una secuencia extraída al azar de una fuente que tiene un valor dado de entropía de Shannon, el valor esperado de su complejidad algorítmica es cercano a ese valor. En otras palabras, el promedio ponderado de las cantidades individuales de información medidas por la complejidad algorítmica tiende al valor de la entropía de Shannon en situaciones genéricas. Por esta razón, algunos teóricos de la información, especialmente los informáticos, consideran la complejidad algorítmica como más fundamental que la entropía de Shannon en tanto medida de la información (véase, por ejemplo, Cover y Thomas 1991: 3).

4.2. Información cuántica

Si bien muchos trabajos sobre información cuántica precedieron al artículo de Benjamin Schumacher (1995) "Codificación Cuántica" (véase, por ejemplo, Ingarden 1976), este trabajo se considera la primera formalización precisa de la Teoría de la Información Cuántica. El objetivo principal del artículo es demostrar un teorema para la codificación cuántica análogo al Primer Teorema de Shannon (*Noiseless-Channel Coding Theorem*).

Al igual que en el caso de Shannon, el sistema de comunicación consta de cinco partes: una fuente S, un transmisor T, un canal CH, un receptor R y un destino D. La fuente es un sistema de n letras s_i, cada una con su propia probabilidad $p(s_i)$; por lo tanto, aquí también la entropía de Shannon de la fuente es $H(S)$. La diferencia se centra en la etapa de codificación. El transmisor T mapea el alfabeto de la fuente, $A_S = \{s_1, ..., s_n\}$, ya no sobre un alfabeto binario implementado por sistemas clásicos de dos estados, sino sobre un conjunto de n estados $|a_i\rangle$ de un *sistema cuántico M*, $A_T = \{|a_i\rangle, ..., |a_n\rangle\}$, que pueden no ser ortogonales. La mezcla de estados del sistema M puede ser representada por un operador densidad

$$\rho = \sum_{i=1}^{n} p(a_i)\left|a_i\right\rangle\left\langle a_i\right|$$

Sobre dicho operador puede calcularse la llamada *entropía von Neumann* como

$$S(\rho) = Tr(\rho \log \rho)$$

En el caso particular de que los estados cuánticos sean mutuamente ortogonales, la entropía de von Neumann es igual a la entropía de Shannon. En el caso general, en cambio, $S(\rho) \leq H(A)$. De este modo, cada mensaje de longitud *N* producido por la fuente *S* se codifica mediante palabras de código de longitud *L*, físicamente implementadas por un número *L* de sistemas cuánticos de dos estados, usualmente denominados *qubits*. Análogamente al caso de Shannon, *L* indica la compacidad del código: cuanto menor es el valor de *L*, mayor es la eficiencia de la codificación, es decir, se necesitan menos qubits para codificar los mensajes. El Teorema Cuántico de Codificación de Canal sin Ruido (*Quantum Noiseless-Channel Coding Theorem*) de Schumacher demuestra que, para mensajes suficientemente largos, el número óptimo de qubits necesarios para transmitir los mensajes generados por la fuente es proporcional, ya no a la entropía de Shannon, sino a la entropía de von Neumann.

Durante las últimas décadas, la llamada "Teoría de la Información Cuántica" se ha convertido en un campo en constante crecimiento, dedicado a explotar los recursos cuánticos para realizar tareas informativas. En particular, el entrelazamiento cuántico (*entanglement*) se ha utilizado como un recurso muy importante en la comunicación. La *teleportación* es uno de los ejemplos paradigmáticos de comunicación asistida por entrelazamiento (*entanglement assisted communication*). Si bien se trata de un resultado directo de la mecánica cuántica, la teleportación aparece como un fenómeno extraño cuando se lo describe como un proceso de transmisión de información. En términos generales, un estado cuántico desconocido se transfiere entre dos agentes, Alice y Bob, con la ayuda de un par de partículas compartidas preparadas en un estado entrelazado y de dos bits clásicos enviados por Alice a Bob (la descripción del protocolo se puede encontrar en cualquier libro de texto sobre el tema; véase, por ejemplo, Nielsen y Chuang 2010). La perplejidad se deriva del hecho de que una gran cantidad (estrictamente infinita) de información se transfiere de Alice a Bob enviando solo dos bits, y sin ningún enlace físico que no sea el canal clásico a través del cual se transmiten los bits clásicos.

Sin duda, Teoría de la Información Cuántica merece ser mencionada como una importantísima articulación de la teoría de la información formulada por Shannon debido a su enorme relevancia científica y tecnológica actual. No obstante, además de este interés, como veremos la irrupción de la información cuántica condujo a una profunda revisión del ya instalado debate acerca de la interpretación del concepto de información.

5. Interpretaciones del concepto de información

A pesar del éxito inmediato de la Teoría de la Información de Shannon y de sus múltiples aplicaciones, el concepto de información fue objeto de diferentes interpretaciones, incluso sobre la base del acuerdo respecto del formalismo (Lombardi, Fortin y Vanni 2015). A su vez, la Teoría de la Información Cuántica introdujo nuevas aristas en la discusión (véase Lombardi, Holik y Vanni 2016b).

5.1. Interpretación epistémica

La interpretación epistémica de la información concuerda con la visión cotidiana, que adjudica al conocimiento un papel central en la constitución del concepto: la información es algo que proporciona conocimiento, algo que modifica el estado del conocimiento de quienes lo reciben. Esta idea puede sugerir que el vínculo entre información y conocimiento es una característica de la noción cotidiana de información, pero no del concepto técnico (véase Timpson 2004, 2013). Sin embargo, la bibliografía sobre el tema pone de manifiesto que dicho vínculo se encuentra con frecuencia tanto en filosofía como en ciencia.

Por ejemplo, en el campo de la filosofía, Fred Dretske sostiene que: "la información es una mercancía que, dado el destinatario adecuado, es capaz de producir conocimiento" (1981: 47). Algunos autores dedicados a las ciencias especiales también consideran que el significado central del concepto de información, incluso en su sentido técnico, está vinculado al concepto de conocimiento; en esta tendencia, Jon M. Dunn define la información como "lo que queda del conocimiento cuando se quita la creencia, la justificación y la verdad" (2001: 423). Cuando se trata de información, también los físicos con frecuencia hablan sobre lo que sabemos o podemos saber. Por ejemplo, Anton Zeilinger incluso identifica información y conocimiento cuando afirma que "[tenemos] conocimiento, es decir, información de un objeto solo a través de la observación" (1999: 633) o, con Časlav Brukner, "[p]or conveniencia no usaremos aquí una medida de información o conocimiento, sino más bien su opuesto, una medida de incertidumbre o entropía."(Brukner y Zeilinger 2009: 681-682). En un libro de texto tradicional sobre la teoría de Shannon aplicada a la ingeniería, también se puede leer que la información "se mide como una diferencia entre el estado del conocimiento del destinatario antes y después de la comunicación de la información" (Bell 1957: 7). Aunque no se refiere a la teoría de Shannon, sino al contexto cuántico, Christopher Fuchs adhiere al bayesianismo respecto de las probabilidades y, en consecuencia, también aboga por una interpretación epistémica de la información (véase Caves, Fuchs y Schack 2002).

Si bien desde la perspectiva epistémica la información no es un elemento físico, en general se supone que la posibilidad de adquirir conocimiento sobre la fuente de información al acceder al estado del sistema destino está enraizada en la conexión nómica entre fuente y destino, es decir, en la legalidad de las regularidades subyacentes a toda la situación. En particular, las probabilidades condicionales utilizadas para calcular las magnitudes relevantes de la teoría deben determinarse mediante leyes naturales, que establecen directa o indirectamente los vínculos entre el origen y el destino. De hecho, si esas probabilidades condicionales representaran correlaciones accidentales, meramente de facto, los estados en el destino no nos dirían nada acerca de lo que sucedió en la fuente.

A pesar de su plausibilidad, la interpretación epistémica no se encuentra exenta de dificultades conceptuales. Mientras que la información mutua $H(S;D)$ puede interpretarse fácilmente como una medida del conocimiento sobre la fuente obtenida en el destino, el ruido N y la equivocidad E no miden conocimiento, sino que, por el contrario, son obstáculos para la adquisición de conocimiento. No es fácil ver cómo el ruido, que puede producirse fuera del sistema de comunicación sin relación alguna con la fuente de información (piénsese, por ejemplo, en ruido blanco en un receptor de radio), podría concebirse como algo que transporta o brinda conocimiento. Una solución a este problema consistiría en suponer que sólo las entropías de la fuente $H(S)$ y del destino $H(D)$ y la información mutua $H(S;D)$, pero no el ruido N y la equivocidad E, pueden conceptualizarse de manera significativa como medidas de conocimiento. Pero esta respuesta llevaría a admitir la posibilidad de sumar y restar magnitudes que refieren a diferentes tipos de ítems, en este caso conocimiento y algo diferente del conocimiento (como, por ejemplo, lo admite el diagrama que relaciona las diferentes entropías), una práctica absolutamente prohibida en las ciencias matematizadas.

Otro problema de la interpretación epistémica se refiere a la relación entre información y comunicación. Consideremos un transmisor de televisión TV que transmite una señal recibida por dos aparatos de televisión S y R: aunque no hay interacción física entre los dos aparatos, las correlaciones entre sus estados no son accidentales, sino que resultan de la dependencia física de esos estados respecto de los estados del transmisor TV. Por lo tanto, a partir de una interpretación epistémica, nada nos impide admitir la existencia de un enlace informativo entre los dos aparatos: podemos definir un canal de comunicación entre S y R porque es posible saber algo sobre S mirando a R y viceversa: "desde un punto de vista teórico, [...] el canal de comunicación puede considerarse simplemente como el conjunto de relaciones de dependencia entre [un sistema] S y [un sistema]

R. Si las relaciones estadísticas que definen equivocidad y ruido entre *S* y *R* son apropiados, entonces hay un canal entre estos dos puntos, y la información pasa entre ellos, incluso si no hay un enlace físico directo que una *S* con *R*." (Dretske 1981: 38). Aunque consistente, algo hay en esta situación que suena extraño cuando se considera que la información está relacionada con la comunicación. En efecto, la comunicación implica que, en algún lugar, alguien hace algo que tiene consecuencias en otro lugar, y gracias a ello el enlace comunicacional se establece. Pero en el caso de los dos aparatos de televisión, nada puede hacerse, por ejemplo, en el extremo *S* que afecte lo que sucede en el extremo *R*. En otras palabras, el cambio del estado del aparato *S* no puede utilizarse para controlar el estado del aparato *R* y, de este modo, *enviar* información. Esto significa que, desde la perspectiva puramente epistémica, algo todavía se pierde respecto de la concepción habitual de comunicación.

Estas dificultades de la interpretación epistémica de la información no implican que tal interpretación sea completamente errónea. Por el contrario, puede aplicarse de un modo fructífero en psicología y en ciencias cognitivas, donde el concepto de información puede usarse para conceptualizar las habilidades humanas de adquirir conocimiento (véase, por ejemplo, Tononi 2004, 2017). La interpretación epistémica también podría servir como base para los intentos, motivados filosóficamente, de suplementar una teoría formal de la información con una dimensión semántica (MacKay 1969, Nauta 1972, Dretske 1981).

5.2. Interpretación física

Según la interpretación física, la información es algo que existe en la naturaleza, independientemente de cómo la midamos: "La información es física" (Landauer 1991: 23). Ésta es la posición de muchos físicos y la mayoría de los ingenieros, para quienes el vínculo con el conocimiento no es un tema central, ya que la transmisión de información puede usarse sólo con fines de control, como, por ejemplo, operar un dispositivo en el extremo de destino actuando sobre el estado de la fuente. En este contexto interpretativo, la información generalmente se compara con la energía, que ingresó en el dominio de la física como una mera herramienta para describir lo que podemos hacer con los sistemas físicos "para realizar el trabajo", pero gradualmente se convirtió en un elemento abstracto, que desempeña un papel central en la unificación de la física: la energía es un elemento esencialmente presente en todas las teorías físicas contemporáneas. A la luz de la fuerte presencia del concepto de información en la física actual, varios autores (por ejemplo, Stonier 1990, 1996, Rovelli, comunicación personal) consideran que dicho concepto está siguiendo una trayectoria histórica análoga a la seguida por el concepto de energía durante el Siglo XIX. Sobre la base de esta caracterización general, sin embargo, hay diferentes maneras de abogar por la interpretación física.

La interpretación física *fuerte* está vinculada con la idea expresada por el conocido slogan "no hay información sin representación": la transmisión de información entre dos puntos del espacio físico requiere necesariamente una señal portadora de información, es decir, un proceso físico que se propaga desde un punto al otro. Rolf Landauer es un explícito defensor de esta posición cuando afirma que "[la] información no es una entidad abstracta incorpórea; siempre está ligado a una representación física. Se representa mediante el grabado en una tableta de piedra, un giro, una carga, un agujero en una tarjeta perforada, una marca en un papel o algún otro equivalente." (1996: 188; véase también Landauer 1991, Rovelli 2016). Esta interpretación también es adoptada por algunos filósofos de la ciencia; por ejemplo, Peter Kosso afirma que "la información se transfiere entre estados a través de una interacción" (1989: 37). Esta interpretación física fuerte es la más común en los campos de la física y la ingeniería: la necesidad de una señal portadora suena natural a la luz de la idea genérica de que las influencias físicas sólo pueden transferirse a través de interacciones. Sin embargo, incluso en el contexto de esta versión fuerte, se pueden distinguir dos puntos de vista ontológicos diferentes, en general implícitos en las discusiones sobre el tema.

Según una primera perspectiva, la información pertenece a la categoría ontológica de *sustancia*, es decir, es objeto de predicación que no puede predicarse de nada más, y es portadora de propiedades (véase Robinson 2014). En este sentido, la característica esencial de la información consiste en su capacidad de "fluir" a través del espacio físico, es decir, de generarse en un punto y transmitirse a otro punto; también se puede acumular, almacenar y convertir de una forma a otra. Desde otra perspectiva, en cambio, la información es ontológicamente una *propiedad*; en particular, es una propiedad de la señal portadora. Por lo tanto, incluso si las propiedades no fluyen, la imagen del "flujo" de información puede tener cierto sentido: hay una propagación de la señal física que une el transmisor y el receptor, y la información es una propiedad de dicha señal.

Si bien generalizada en las comunidades de la física y la ingeniería, la interpretación física fuerte se enfrenta a un desafío muy serio cuando entra en juego la mecánica cuántica, en particular, la comunicación asistida por entrelazamiento. Como ya fue señalado, en la teleportación se transfiere un estado cuántico desconocido entre dos puntos del espacio con la ayuda de un par de partículas compartido, preparado en un estado entrelazado, y de dos bits clásicos enviados de uno de los puntos al otro. Sin embargo, no hay señal física que lleve el estado teleportado: no hay una interacción que sirva de conexión entre los dos lugares del espacio. Por esta razón, muchos físicos cuánticos intentan encontrar un vínculo físico entre los dos puntos del espacio, por ejemplo, considerando que la información viaja hacia atrás en el tiempo hasta el evento en el que se produjo el par entrelazado, y luego viaja hacia el futuro hasta el momento en el que se recupera el estado teleportado (Penrose 1998, Jozsa 1998, 2004), o suponiendo que la información es transportada por los bits clásicos, pero de forma oculta (Deutsch y Hayden, 2000). Pero el defensor de la interpretación física fuerte se encuentra en serios problemas si se mantiene reacio a aceptar este tipo de explicaciones de un alto grado de artificialidad.

A diferencia de la interpretación fuerte, la interpretación física *débil* no requiere una interacción física entre dos puntos del espacio para admitir la existencia de transmisión de información entre ellos. Al permanecer agnóstica sobre su categoría ontológica, la única afirmación de la postura débil es que la información es un elemento que pertenece al mundo de la física. En otras palabras, el concepto de información no es un concepto meramente formal: el término 'información' se refiere a algo que es parte del mobiliario de la realidad tal como la describen las ciencias físicas. Por lo tanto, incluso si la cuestión de su categoría ontológica aún no se decide, el comportamiento de la información se rige por las leyes de la física. Gracias a su perspectiva ontológicamente neutral, la interpretación física débil no se ve desafiada por la comunicación asistida por entrelazamiento, si bien sigue en deuda respecto de la decisión acerca de la categoría ontológica de la información.

5.3. Interpretación formal o deflacionista

Aunque la interpretación física fuerte ha sido la más habitual en los libros de enseñanza tradicionales para la formación de ingenieros, esta situación ha cambiado en los últimos tiempos: en general, los textos actuales presentan la teoría de la información de manera formal, sin mencionar fuentes, receptores o señales, y los conceptos básicos se introducen en términos de variables aleatorias, distribuciones de probabilidad sobre sus posibles valores y correlaciones entre ellas (véase, por ejemplo, Cover y Thomas 1991). Esta tendencia conlleva implícitamente una interpretación formal, según la cual el término 'información' no refiere a algo que existe en el mundo, pero tampoco tiene que ver con el conocimiento. Estrictamente hablando, la palabra no pertenece a la ciencia empírica sino a la *ciencia formal*: no tiene una referencia extralingüística en sí misma; su "significado" tiene sólo una dimensión sintáctica. Por lo tanto, información es un concepto puramente formal y la teoría de la información es un capítulo de la *teoría de la probabilidad* (véase, por ejemplo, Khinchin 1957, Reza 1961).

En cierto sentido, el enfoque deflacionista de la información, defendido por Christopher Timpson (2004, 2013) y Armond Duwell (2008), puede verse como una variante de la interpretación formal: "[u]na vez que se reconoce que 'información' es un sustantivo abstracto, entonces está claro que no hay más preguntas que responder sobre cómo se transmite la información en la teleportación que va más allá de proporcionar una descripción de los procesos físicos involucrados en lograr el objetivo del protocolo. Eso es todo lo que «¿Cómo se transmite la información?» puede significar de manera inteligible; porque no se trata de que la información sea una sustancia o entidad que se transporta, ni de que 'información' sea un término con referencia." (Timpson 2006: 599). En consecuencia, desde esta perspectiva, las teorías de la información (clásica y cuántica, y también la teoría de la computación) son "teorías acerca de lo que podemos hacer usando sistemas físicos" (Timpson 2013: 69). Además, puesto que el éxito de la transmisión de información depende de una correcta correlación entre los estados de la fuente y los estados del destino, la transmisión de información sólo depende de las características formales de la disposición comunicacional, independientemente de cómo se implementen estas características por medios físicos.

A la base de la interpretación formal o deflacionista de la información se encuentra la neutralidad física de la información. Más allá de cómo se codifiquen los mensajes en el transmisor (de modo clásico o cuántico), la fuente y el destino de la información se definen de un modo independiente de su sustrato físico: las letras que produce la fuente no son estados físicos, sino que están *implementados* por estados físicos, que pueden ser de naturaleza muy variada. Y lo mismo puede decirse sobre el canal, que encarna las correlaciones entre el origen y el destino: no importa cómo se establezcan y se "materialicen" físicamente esas correlaciones; lo único relevante es que vinculen los estados de la fuente y los estados del destino. Esto significa que la teoría de Shannon "[s]e puede aplicar a cualquier sistema de comunicación independientemente de si sus partes se describen mejor por la mecánica clásica, la electrodinámica clásica, la teoría cuántica o cualquier otra teoría física" (Duwell, 2003: 480).

Por supuesto, el enfoque formal está libre de las dificultades que desafían la interpretación física fuerte: si la palabra 'información' no tiene referencia en el mundo, los problemas de explicar cómo se transmite la información en la teleportación o de decidir la categoría ontológica de información desaparecen de un plumazo. Sin embargo, concebir la información como un concepto meramente formal conduce al mismo problema que acecha a la interpretación epistémica: existiría transmisión de información aun cuando la comunicación es imposible. Y aún más: si la información es un concepto puramente formal, las correlaciones involucradas en la teoría matemática de la información pierden incluso su ingrediente nómico. Por ejemplo, la información mutua entre dos variables aleatorias puede definirse incluso si no existe una relación legal entre ellas y sus probabilidades condicionales expresan sólo correlaciones de facto. Por lo tanto, la interpretación formal no incorpora el elemento de producción implícito en la comunicación, ese elemento que permite que algunas acciones tomadas en la fuente de información tengan un efecto en el destino de la información (véase Lombardi, Fortin, y López 2016). En otras palabras, si bien ganando en generalidad, la visión formal de la información pierde fuerza conceptual en el contexto comunicacional.

Precisamente para resolver este problema, algunos autores han propuesto suplementar la idea de correlación entre fuente y destino con un ingrediente causal: de algún modo, las modificaciones en el estado de la fuente deben ser responsables de las modificaciones en los estados del destino. En particular, Cristian López y Olimpia Lombardi (2019) toman como base la visión deflacionista de Timpson, pero la complementan mediante el *enfoque manipulabilista de la causación* en su versión intervencionista (Woodward 2003, 2007, Pearl 2009). Desde esta perspectiva, la información no implica simplemente la existencia de correlaciones estadísticas entre la fuente y el receptor, como se entiende desde una visión puramente formal. La comunicación es un fenómeno *asimétrico* que involucra nociones *causales*: el sistema destino debe poder ser manipulado causalmente, interviniendo en la fuente, para una comunicación exitosa. Sobre esta base, López y Lombardi acuñan el lema: "no hay comunicación sin manipulación".

6. Consideraciones finales: los múltiples rostros de la información

El concepto de información ingresa al ámbito científico durante el siglo XX de la mano de diversos autores, entre los cuales se destaca Claude Shannon. Y si bien la Teoría de la Información de Shannon fue formulada inicialmente con fines estrictamente tecnológicos, su alcance rápidamente desbordó los cauces originales para introducirse en prácticamente todos los ámbitos de la ciencia. Por otra parte, otras teorías y formalismos fueron enriqueciendo el panorama de las ciencias de la información desde distintas perspectivas. El acelerado crecimiento de estas temáticas condujo a la reflexión filosófica acerca del concepto mismo de información:

durante las últimas décadas se ha producido un interesante trabajo de análisis que ha alimentado un debate interpretativo profundo y fructífero.

Frente a este panorama, tal vez sea hora de dejar de lado las posturas monistas acerca del concepto de información y adoptar una posición pluralista, según la cual los diferentes formalismos y las diferentes interpretaciones ya no son rivales, sino que representan distintos miembros de una única familia conceptual, distintas versiones específicas de una misma idea general. Cada una de estas versiones es legítima en la medida en que su aplicación resulta de utilidad en un determinado campo científico o tecnológico. Esta postura pluralista no sólo recoge la amplia y fuerte presencia del concepto de información en todas las actividades humanas contemporáneas, con significados que van más allá de la teoría de Shannon, sino que también se encuentra en resonancia con la posición del propio Shannon. De hecho, a pesar de haber sido el autor del formalismo más famoso para tratar la información, Shannon mismo reconoció que el concepto de información es altamente polisémico y que no hay razón alguna para esperar o desear una unificación futura:

> La palabra 'información' ha recibido diferentes significados por diversos autores en el campo general de la teoría de la información. [...] No es de esperar que un solo concepto de información explique satisfactoriamente las numerosas aplicaciones posibles de este campo general. (Shannon 1993: 180)

Referencias

Adriaans, P. (2013) "Information", en E. N. Zalta (ed.), The Stanford Encyclopedia of Philosophy (Fall 2013 Edition). URL = <http://plato.stanford.edu/archives/fall2013/entries/information/>.

Bar-Hillel, Y. (1964) *Language and Information: Selected Essays on Their Theory and Application*. Reading, MA: Addison-Wesley.

Bar-Hillel, Y. y R. Carnap (1953) "Semantic Information". *The British Journal for the Philosophy of Science*, 4: 147-157.

Bell, D. (1957) *Information Theory and its Engineering Applications*. London: Pitman & Sons.

Brukner, Č. y A. Zeilinger (2009) "Information Invariance and Quantum Probabilities". *Foundations of Physics*, 39: 677-689.

Bub, J. (2007) "Quantum Information and Computation", en J. Butterfield y J. Earman (eds.), *Philosophy of Physics. Part B*. Amsterdam: Elsevier, pp. 555-660.

Caves, C. M. y C. A. Fuchs (1996) "Quantum Information: How Much Information in a State Vector?", en A. Mann y M. Revzen (eds.), "The dilemma of Einstein, Podolsky and Rosen - 60 Years Later", *Annals of the Israel Physical Society*, 12: 226-257. URL = <https://arxiv.org/pdf/quant-ph/9601025.pdf>

Caves, C. M.; C. A. Fuchs y R. Schack (2002) "Unknown Quantum States: The Quantum de Finetti Representation". *Journal of Mathematical Physics*, 43: 4537-4559.

Chaitin, G. (1966) "On the Length of Programs for Computing Binary Sequences". *Journal of the Association for* Computing *Machinery*, 13: 547-569.

Cover, T. y J. A. Thomas (1991) *Elements of Information Theory*. New York, NY: John Wiley & Sons.

Deutsch, D. y P. Hayden (2000) "Information Flow in Entangled Quantum Systems". *Proceedings of the Royal Society of London A*, 456: 1759-1774.

Dretske, F. (1981) *Knowledge and the Flow of Information*. Cambridge, MA: The MIT Press.

Dunn, J. M. (2001) "The Concept of Information and the Development of Modern Logic", en W. Stelzner (ed.), Non-Classical Approaches in the Transition from Traditional to Modern Logic. Berlin: de Gruyter, pp. 423-427.

Duwell, A. (2003) "Quantum Information Does Not Exist". *Studies in History and Philosophy of Modern Physics*, 34: 479-499.

Duwell, A. (2008) “Quantum Information Does Exist”. *Studies in History and Philosophy of Modern Physics*, 39: 195-216.
Floridi, L. (2011) The Philosophy of Information. Oxford: Oxford University Press.
Floridi, L. (2013) “Semantic Conceptions of Information”, en E. N. Zalta (ed.), The Stanford Encyclopedia of Philosophy (Spring 2013 Edition). URL =
<http://plato.stanford.edu/archives/spr2013/entries/information-semantic/>.
Glander, T. (2000) *Origins of Mass Communications Research during the American Cold War: Educational Effects and Contemporary Implications*. Mahwah, NJ: Lawrence Erlbaum Associates.
Hartley, R. (1928) “Transmission of Information”. *Bell System Technical Journal*, 7: 535-563.
Ingarden, R. (1976) “Quantum Information Theory”. *Reports on Mathematical Physics*, 10: 43-72.
Jozsa, R. (1998) “Entanglement and Quantum Computation”, en S. Huggett, L. Mason, K. P. Tod, S. T. Tsou y N. M. J. Woodhouse (eds.), *The Geometric Universe*. Oxford: Oxford University Press, pp. 369-79.
Jozsa, R. (2004) “Illustrating the Concept of Quantum Information”. *IBM Journal of Research and Development*, 4: 79-85.
Khinchin, A. (1957) *Mathematical Foundations of Information Theory*. New York, NY: Dover.
Kolmogorov, A. (1965) “Three Approaches to the Quantitative Definition of Information”. *Problems of Information Transmission*, 1: 4-7.
Kolmogorov, A. (1968) “Logical Basis for Information Theory and Probability Theory”. *Transactions on Information Theory*, 14: 662-664.
Kolmogorov, A. (1983) “Combinatorial Foundations of Information Theory and the Calculus of Probabilities”. *Russian Mathematical Surveys*, 38: 29-40.
Kosso, P. (1989) *Observability and Observation in Physical Science*. Dordrecht: Kluwer.
Landauer, R. (1991) “Information is Physical”. *Physics Today*, 44: 23-29.
Landauer, R. (1996) “The Physical Nature of Information”. *Physics Letters A*, 217: 188-193.
Lombardi, O. (2004) “What is Information?”. *Foundations of Science*, 9: 105-134.
Lombardi, O.; S. Fortin y C. López (2016) “Deflating the Deflationary View of Information”. *European Journal for Philosophy of Science*, 6: 209-230.
Lombardi, O.; S. Fortin y L. Vanni (2015) “A Pluralist View about Information”. *Philosophy of Science*, 82: 1248-1259.
Lombardi, O.; F. Holik y L. Vanni (2016a) “What is Shannon Information?”. Synthese, 193: 1983-2012.
Lombardi, O.; F. Holik y L. Vanni (2016b) “What is Quantum information?”. *Studies in History and Philosophy of Modern Physics*, 56: 17-26.
López, C. y O. Lombardi (2019) “No Communication Without Manipulation: A Causal-Deflationary View of Information”. *Studies in History and Philosophy of Science*, 73: 34-43.
MacKay, D. (1969) *Information, Mechanism and Meaning*. Cambridge MA: The MIT Press.
Nauta, D. (1972) *The Meaning of Information*. The Hague: Mouton.
Nielsen, M. y I. Chuang (2010) *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.
Pearl, J. (2009) *Causality: Models, Reasoning and Inference*. Cambridge: Cambridge University Press.
Penrose, R. (1998) “Quantum Computation, Entanglement and State Reduction”. *Philosophical Transactions of the Royal Society of London A*, 356: 1927-1939.
Reza, F. (1961) *Introduction to Information Theory*. New York, NY: McGraw-Hill.
Robinson, H. (2014) “Substance”, en E. N. Zalta (ed.), The Stanford Encyclopedia of Philosophy (Spring 2014 Edition). URL =
<http://plato.stanford.edu/archives/spr2014/entries/substance/>.
Rovelli, C. (2016) “Meaning = Information + Evolution”. Publicado en arXiv. URL = < https://arxiv.org/abs/1611.02420>
Schumacher, B. (1995) “Quantum Coding”. *Physical Review A*, 51: 2738-2747.
Shannon, C. (1948) “The Mathematical Theory of Communication”. *Bell System Technical Journal*, 27: 379-423.
Shannon, C. (1993) *Collected Papers*, N. Sloane y A. Wyner (eds.). New York, NY: IEEE Press.

Shannon, C. y Weaver, W. (1949) *The Mathematical Theory of Communication*. Urbana y Chicago, IL: University of Illinois Press.

Solomonoff, R. (1964) “A Formal Theory of Inductive Inference”. *Information and Control*, 7: 1-22, 224-254.

Stonier, T. (1990) *Information and the Internal Structure of the Universe: An Exploration into Information Physics*. New York, NY & London: Springer.

Stonier, T. (1996) “Information as a Basic Property of the Universe”. *Biosystems*, 38: 135-140.

Timpson, C. G. (2004) *Quantum Information Theory and the Foundations of Quantum Mechanics*. PhD dissertation, University of Oxford.

Timpson, C. (2006) “The Grammar of Teleportation.” *The British Journal for the Philosophy of Science*, 57: 587-621.

Timpson, C. (2013) *Quantum Information Theory and the Foundations of Quantum Mechanics*. Oxford: Oxford University Press.

Tononi, G. (2004) “An Information Integration Theory of Consciousness”. *BMC Neuroscience*, 5: 42.

Tononi, G. (2017) “Integrated Information Theory of Consciousness. Some Ontological Considerations”, en M. Velmans y S. Schneider (eds.), *The Blackwell Companion to Consciousness*. Hoboken NJ: Wiley-Blackwell, pp. 621-633.

Woodward, J. (2003) *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

Woodward, J. (2007) “Causation with a Human Face”, en H. Price y R. Corry (eds.), *Causation, Physics, and the Constitution of Reality: Russell’s Republic Revisited*. Oxford: Oxford University Press, pp. 66-105.

Zeilinger, A. (1999) “A Foundational Principle for Quantum Mechanics”. *Foundations of Physics*, 29: 631-643.