

POR UNA FILOSOFÍA DE LA INFORMACIÓN

Luciano Floridi

Dipartimento di Scienze Filosofiche, Università degli Studi di Bari;

Faculty of Philosophy and Sub-Faculty of Computation, Information Ethics Group, Oxford
University.

Address for correspondence: Wolfson College, OX2 6UD, Oxford, UK;

luciano.floridi@philosophy.oxford.ac.uk

La filosofía de la Información (FI) es una nueva disciplina filosófica que trata acerca de:

A) la investigación crítica sobre la naturaleza conceptual y los principios básicos de la información, incluso de su *dinámica* (concretamente: la *computación*, es decir, la manipulación mecánica de los datos, y el *flujo informacional*, a saber, los diversos procesos de recogida, almacenamiento e intercambio de informaciones), su *utilización* (la denominada *information ethics* (o ética de la información') y sus ciencias; y

B) la elaboración de metodologías teórico-informacionales y computacionales susceptibles de ser aplicables a los problemas filosóficos.

Esta definición, introducida en Floridi [2002], es precisa aunque tal vez demasiado árida. Sugiere en efecto una idea acotada y tecnicista de la FI, sin explicar lo interesante, innovadora e influyente que puede resultar en muchos campos de la investigación filosófica. En este breve artículo intentaré restablecer el equilibrio¹.

Adentrémonos un poco más. La investigación computacional y teórico-informacional de los problemas filosóficos tuvo sus inicios muchos antes de la difusión comercial de los ordenadores, cuando a mediados de los años cuarenta del siglo XX se desarrollaron las seminales investigaciones de Alan Turing². Continuaron durante los años cincuenta dando lugar a la *filosofía de la inteligencia artificial*³, un ámbito de la investigación que durante las dos próximas décadas sería la cenicienta (cenerentola) de los estudios epistemológicos. A partir de segunda mitad de los años setenta tres importantes factores establecieron las bases

para el nacimiento de la FI⁴. El orden de presentación de los mismos no sigue ningún orden cronológico o de prioridad.

El primer factor es el sociológico. El fenómeno de la informática de masas, mediante el ordenador personal, la automatización industrial, Internet y la alfabetización informática, tiene como resultado la difusión de los *mass media* y de diversas ICT (o *Information and communication technologies*, pensemos por ejemplo en la telefonía móvil), dando lugar a la sociedad post-industrial de la información.

El segundo factor es el científico. Durante la segunda mitad de los años noventa la informática permeabiliza capilarmente el mundo del saber, estableciendo las bases para un nuevo ámbito de la investigación, como sucede por ejemplo en la bioingeniería, la genética, en las ciencias cognitivas, e las neurociencias experimentales, en la física de los materiales, en la lingüística computacional o en la arqueología. La investigación científica se ha visto totalmente influenciada por la revolución informática que resultaría ya imposible sin el soporte digital de la misma.

El último factor es el cultural. El mundo de la información digital y de los ordenadores, junto con todas estas nuevas formas conseguidas de experiencia, comunicación y socialización, ha dado lugar a una nueva lengua franca interdisciplinar, que incluye conceptos como *online*, *input*, *output*, *feedback*, *network*, sistemas digitales y etcétera.

La convergencia de estos y otros factores ha propiciado que, en pocos decenios, antiguos problemas filosóficos hayan sido transformados, mientras que ha aparecido una nueva serie de otros completamente nuevos. Nace de este modo la filosofía de la información⁵.

El tema de las relaciones entre mente y ordenador (central para la Inteligencia Artificial, o IA) es tan sólo uno de sus nuevos temas de investigación. Podemos pensar también en las dificultades presentes en la estética de los nuevos medios de comunicación de masas y del arte digital; o en la definición de quien o qué cosa puede ser calificado como un agente moral cuando es ya normal hablar de sistemas multi-agente, de agentes artificiales y de 'identidad' distribuida [Floridi & Sanders, 2004b]; o cómo debe entenderse la privacidad personal, un argumento clásico de la ética de la información. Los extendidos conceptos de (tele) *presencia* y *experiencia* deben reconsiderarse bajo la nueva luz no tanto de la realidad virtual (fenómeno todavía ligado más bien a la 'fantásticociencia' que no a la experiencia

cotidiana) como más bien en las relaciones surgidas de la virtualización de muchos aspectos de la realidad, como de la telemedicina a las comunidades electrónicas, o de la pornografía digital a la telerobótica. Incluso 'estar' quiere a menudo decir estar on-line o tele-estar [Goldberg, 2000; Floridi, 2004f], como en el caso de un cirujano que opera a distancia, de un soldado que trabaja en un campo minado controlando por control remoto un robot, o de un ingeniero que guía y controla un vehículo de exploración en Marte. Pensemos también en las rápidas transformaciones de las ideas acerca de la identidad personal y de las interacciones comunicativas en el contexto de la Red; o incluso en cómo podemos definir el umbral de lo existente y de lo vivo a través de los estudios sobre la *Artificial Life (Alife)*. Podemos fácilmente mencionar otros ejemplos, como el desarrollo de las lógicas no-monotónicas o de las lógicas modales gracias a las contribuciones de la inteligencia artificial (Benzi, 1997); las transformaciones sufridas en las prácticas de la escritura y en la teoría de la literatura (basta con pensar en los correos electrónicos o el advenimiento del hipertexto) sugieren el nacimiento de una epistemología informática y de metodologías experimentales *in silico* (mediante simulaciones computacionales) contempladas tanto desde la ética (Danielson, 1992), como de la epistemología (Grim *et al*, 1998).

Lo expresado hasta el momento constituye tan sólo un esbozo de la extraordinaria cantidad de ámbitos conflictivos y de los espacios de reflexión que se plantean desde la perspectiva de una filosofía de la información. A pesar de todo, creo que resulta necesario remarcar que las interacciones entre la filosofía, la informática y los fenómenos relacionados con las revoluciones informacionales han resultado ser siempre conceptualmente fecundas y han sido difundidas desde un punto de vista académico, produciendo en los últimos veinte años una amplia cantidad de resultados interesantes e importantes⁶.

Durante la segunda mitad del siglo XX, la informática ha terminado por transformar radicalmente la realidad que nos envuelve, nuestro modo de vivir, y el modo de conceptualizar el propio mundo. Jamás madrugadora, la filosofía, que ya ha dado cuenta de esta transformación a través de los múltiples campos teóricos que la componen - como la filosofía de la ciencia, la filosofía del lenguaje, la epistemología, la filosofía de la mente, la ontología, la ética o la estética - sería irreconocible sin el cuadro conceptual aportado por la revolución informática.

¿Ha devenido la cenicienta una princesa? Tal vez no. Resulta en efecto lícito

preguntarnos , si además de generar una nueva área de investigación, la informática ha puesto en realidad a disposición de la filosofía así de otras ciencias variadas (en realidad, pocas) únicamente instrumentos digitales (*software* para la didáctica, textos electrónicos y bases de datos, como podemos encontrar en Floridi, 1996b) o, por encima de todo, un potente metalenguaje a través del cual se traducen muchos de sus problemas clásicos y acaso aporta algunos nuevos. Se podría argumentar que aunque en la medicina los ordenadores resulten ya insustituíbles, ello no implica la existencia de una nueva área de la investigación médica.

La duda se encuentra justificada, aunque puede ser fácilmente disipada si consideramos, como han hecho Bynum y Moor (2000) que nuestra cultura está desarrollando un nuevo paradigma filosófico:

De vez en cuando se desarrollan en la filosofía importantes corrientes de pensamiento. Estas corrientes se inician con pocas y simples, aunque altamente fértiles, ideas – ideas que confieren a los filósofos una nueva perspectiva con la que analizar los problemas filosóficos. De forma gradual, los métodos y problemas filosóficos son redefinidos y analizados bajo estas nuevas nociones. Una vez que se obtienen nuevos e interesantes resultados filosóficos, el movimiento se expande como una onda intelectual que abarca la totalidad de la disciplina. Emerge un nuevo paradigma filosófico. (...) La informática aporta a la filosofía un conjunto de nociones simples aunque increíblemente fértiles – nuevos e *in fieri argumentos, métodos y modelos* para la investigación filosófica. La informática ofrece nuevas oportunidades y cambios a la actividad filosófica tradicional. (...) La informática está transformando el modo según el cuál los los filósofos comprendemos los conceptos fundacionales de la filosofía, tales como ‘mente’, ‘percepción’, ‘experiencia’, ‘razonamiento’, ‘conocimiento’, ‘verdad’, ‘ética’ o ‘creatividad’. Esta orientación de la investigación filosófica, que incorpora la computación dentro del ámbito de análisis de un problema, de su método y en tanto que un modelo para la filosofía, está adquiriendo una importancia cada vez mayor.

Como podemos ver, casi hemos llegado a la definición de nuestra nueva disciplina. La informática no se ha meramente tornado un nuevo instrumento para la filosofía, como las páginas electrónicas, los textos electrónicos, las enciclopedias en CD-ROM, o un nuevo vocabulario técnico. Podemos afirmar que ha generado un paradigma conceptual original,

dando lugar a un área de investigación que dispone de sus propios argumentos, modelos y métodos, que ofrecen innovación y fecundidad.

Para completar el cuadro, basta añadir en este punto un importante matiz, en lo que concierne a la esencia de la disciplina. Las diversas ramas de la filosofía se preguntan qué es *x*: el conocimiento, el ser, el bien, el significado y muchas otras cosas más. Pero, ¿cuál es la *x* del nuevo paradigma? Obviamente, la información. Lo que de hecho ha transformado la filosofía contemporánea no son tanto ciertas *máquinas* (fundamentalmente la máquina de Turing, el modelo implementados en cualquier PC, aunque también podemos pensar en las redes neuronales) con sus propias *arquitecturas* (básicamente la arquitectura de Von Neumann del clásico PC, aunque debemos recordar también la arquitectura paralela y la distribuída) o, de forma más genérica, la *computación* (entendida en un sentido amplio como la manipulación automática de los datos). Por encima de todo, es la propia información, antes que la escritura, la imprenta, los *mass media* e incluso la informática, la que ha hecho emerger de algún modo el ambiente ideal para el desarrollo de las sociedades postindustriales, resultando ser la principal ‘materia prima’ de nuestra actividad, deviniendo el paradigma fundamental para nuestra comprensión de la realidad y de las diversas interacciones que tenemos con la misma.

Bajo la luz de este deslizamiento de la centralidad de la computación hacia la de la información, se puede entender cómo la filosofía de la información es tan sólo la fase más reciente de un largo proceso de transformación del pensamiento filosófico. Hasta aquí un rápido bosquejo.

La revolución científica renacentista reorientó la atención de los filósofos sobre la naturaleza de la realidad cognoscible (mostrada tanto de forma potencial como intrínseca) hacia las relaciones epistémicas existentes entre la propia realidad y el sujeto cognoscente, reemplazando de este modo la metafísica con la epistemología en tanto que saber filosófico primario. La revolución de las comunicaciones ha implicado por lo tanto un desplazamiento ulterior, desde el conocimiento hacia sus condiciones de posibilidad, esto es, la actividad mental, el lenguaje y la semántica. El crecimiento continuado de la sociedad de la información y la aparición de la infoesfera (el ambiente semántico en el que cada vez más personas invertimos nuestro tiempo) ha hecho un paso más, desde el dominio representado de la memoria y del lenguaje del saber organizado (mediante los instrumentos que

constituyen la infoesfera), a su esencia, la propia información. La ‘Información’ se ha convertido por derecho propio en un objeto teórico digno del análisis filosófico, al mismo nivel que los de ‘ser’, ‘conocimiento’, ‘significado’, ‘bien moral’, ‘vida’ o ‘inteligencia’, conceptos todos ellos fundamentales con los que se han establecido nexos de proximidad conceptual.

La mayor ventaja de esta transformación progresiva consiste en que la ‘información’ constituye un objeto teórico menos *thick* (podemos decir que menos denso o rico) que otros, en cuanto permite operaciones de reducción y simplificación conceptual en los análisis (Floridi, 2004b, 2004e, 2004a). Esto justifica la esperanza en su utilidad, tras ser analizado de forma satisfactoria, como una ganzúa para disponer de un acceso privilegiado a los demás conceptos. En epistemología, por poner un ejemplo concreto, podemos servirnos de la filosofía de la información para aproximarnos a problemas, como el de Gettier⁷ que, debatidos en términos de conocimiento, resultan intratables (Floridi, 2004c). Tanto en ética como en filosofía de la mente podemos des-antropologizar y des-psicologizar nuestros análisis sirviéndonos de modelos en los cuáles agentes artificiales (habitualmente *software*) ponen a prueba teorías e intuiciones de otro modo demasiado ricas, resbaladizas o vagas para ser tratadas con la necesaria precisión y cura. Metodológicamente, la informática nos recuerda el valor de una clara especificación del *nivel de abstracción* (Floridi y Sanders, 2004a) al cual se conduce una investigación o se asume que está operando una teoría o un sistema, y de cómo el viciano *verum ipsum factum* y el baconiano *maker’s knowledge tradition* vuelven a mostrar su interés bajo la luz de una metodología que pone el acento sobre los modelos y las implementaciones diseñadas y construídas para la comprensión únicamente conceptual de los fenómenos.

La filosofía parece madurar mediante un doble proceso gradual de empobrecimiento, por un lado, y de enriquecimiento, por el otro. Ella misma delega progresivamente en otras disciplinas cuanto es matemática y empíricamente tratable, mientras se reserva para sí misma las cuestiones normativas y semánticas que se presentan en el curso de la historia humana, cuestiones que restan de hecho intrínseca y genuinamente abiertas a un desacuerdo honesto, razonable e informado. En el curso de tal dialéctica, el concepto de información, en sus tres principales acepciones de información sobre la realidad (*información semántica*, se piensa en una proposición verdadera), para la realidad (*información como instrucción*, si pensamos en

una receta) o como realidad (*información ambiental*, como un aspecto del mundo, si pensamos en las huellas digitales o el ADN), constituye hoy en día un nuevo grado de *minimalismo*. Se puede partir de lo primordial, de un concepto pobre o menos *thick*, con tal de iniciar la escalada hacia las grandes cimas del significado, del conocimiento, de la verdad, del ser y de los valores. Además de correr menores riesgos metafísicos de confundir la gimnasia con la magnesia (como locos con la cabeza en las nubes), partiendo de los fundamentos se dispone de la ventaja de gozar de una visión unificada de cómo los múltiples problemas filosóficos se encuentran trenzados entre sí, visiones que en general una vez nos encontramos en la cima de todas las cumbres. Ciertamente, se trata de una perspectiva demasiado humilde para quien tiene en mente fabulosos vuelos pindáricos y excelentes panorámicas a vista de pájaro. Tenemos por otra parte la ventaja de descender desde lugares nobles hasta los sótanos, descubriendo las raíces comunes – informacionales – de tantos problemas filosóficos. La filosofía de la información es una filosofía pobre, no tanto de ángeles inmaculados y puros, como de mecánicos que intentan construir y reparar sistemas y que se ensucian las manos con la grasa de los pernos de conceptos basilaes y aparentemente poco excitantes como los de los *datos* o la *computación*, de los *agentes* o de los *sistemas*, aunque también aspira a ser reconocida como *prima inter pares*. La Cenicienta emerge.

Agradecimientos

Deseo agradecer a Franca D'Agostini, Gian Maria Greco, Giovanni Mari y Gianluca Paronitti por sus útiles comentarios y por haber corregido algunas imprecisiones del texto original. Obviamente, soy el único responsable de los eventuales errores todavía presentes.

Traducción del italiano al castellano

Dr. Jordi Vallverdú. Prof. Filosofía (UAB).

Bibliografía

Todos los textos citados que están en proceso de publicación pueden encontrarse de forma completa en la siguiente dirección electrónica: www.wolfson.ox.ac.uk/~floridi.

- Anderson, A.R. (1984) *Controversia sobre mentes y máquinas*, Barcelona: Tusquets.
- Benzi, M. 1997, *Il ragionamento incerto. Probabilità e logica in intelligenza artificiale* (Milano: Franco Angeli).
- Boden, Margaret A. (1984) *Inteligencia artificial y hombre natural*, Barcelona: Tecnos.
- Bolter, J. D. 1985, *L'uomo di turing - la cultura occidentale nell'età del computer* (Parma: Pratiche).
- Borrajo, Juristo, N. *et al* (1993) *Inteligencia Artificial. Métodos y Técnicas*, Madrid: Centro de Estudios Universitarios Ramón Areces.
- Bynum, T. W., e Moor, J. H. (a cura di) 2000, *La fenice digitale: Come i computer stanno cambiando la filosofia* (Milano: Apogeo).
- Carlucci Aiello, L., e Cialdea Mayer, M. 1995, *Invito all'intelligenza artificiale* (Milano: Franco Angeli).
- Castelfranchi, Yuriy y Stock, Oliverio (2002) *Máquinas como nosotros, El desafío de la inteligencia artificial*, Madrid: Acento Editorial.
- Coello, Carlos A. (2003) *Breve historia de la computación y sus pioneros*, México: FCE.
- Colburn, T. R. 2000, *Philosophy and computer science* (Armonk, N.Y. ; London: M.E. Sharpe).
- Danielson, P. 1992, *Artificial morality : Virtuous robots for virtual games* (London:

- Routledge).
- Di Giandomenico, M., e Lepschy, A. (a cura di) 1991, *Epistemologia informatica* (Volume pubblicato come n. 5 di BioLogica).
- Fernández, J. (1992) *Vida Artificial*, de Alvaro, Madrid: Eudema.
- Floridi, L. (editor) 1996a, *Filosofia & informatica: Atti del primo incontro italiano sulle applicazioni informatiche e multimediali nelle discipline filosofiche* (Torino: Paravia).
- Floridi, L. 1996b, *L'estensione dell'intelligenza: guida all'informatica per filosofi* (Roma: Armando).
- Floridi, L. 1999a, "Information ethics: On the theoretical foundations of computer ethics", *Ethics and Information Technology*, 1(1), 37-56.
- Floridi, L. 1999b, *Philosophy and computing: an introduction* (London ; New York: Routledge).
- Floridi, L. 2002, "What is the philosophy of information?" *Metaphilosophy*, 33(1-2), 123-45.
- Floridi, L. (editor) 2003a, *The blackwell guide to the philosophy of computing and information* (Oxford, New York: Blackwell).
- Floridi, L. 2003b, "Two approaches to the philosophy of information", *Minds and Machines*, 13(4), 459-69.
- Floridi, L. 2004a, "Information, semantic conceptions of", *Stanford Encyclopedia of Philosophy*.
- Floridi, L. 2004b, "Is information meaningful data?" *Philosophy and Phenomenological Research*.
- Floridi, L. 2004c, "On the logical unsolvability of the Gettier problem", *Synthese*.
- Floridi, L. 2004d, "Open problems in the philosophy of information", *Metaphilosophy*. In corso di pubblicazione.
- Floridi, L. 2004e, "Outline of a theory of strongly semantic information", *Minds and Machines*.
- Floridi, L. 2004f, "The philosophy of presence: From epistemic failure to successful observability", *Presence: Teleoperators and Virtual Environments*.
- Floridi, L., e Sanders, J. W. 2004a, "The method of abstraction" in *Yearbook of the artificial. Nature, culture and technology. Models in contemporary sciences*, a cura di M. Negrotti (Bern: Peter Lang).

- Floridi, L., e Sanders, J. W. 2004b, "On the morality of artificial agents", *Minds and Machines*. In curso de publicacione.
- Ford, K. M., Glymour, C. N., e Hayes, P. J. (a cura di) 1995, *Android epistemology* (Menlo Park, Cambridge, Mass.: AAAI Press ; MIT Press).
- Ginsberg, M. (1993) *Essentials of Artificial Intelligence*, Morgan-Kaufmann.
- Goldberg, K. (a cura di) 2000, *The robot in the garden: telerobotics and telepistemology in the age of the internet* (Cambridge, Mass. ; London: MIT Press).
- Grim, P., Mar, G., e St. Denis, P. 1998, *The philosophical computer: exploratory essays in philosophical computer modeling* (Cambridge, Mass. ; London: MIT Press).
- Moor, J., e Bynum, T. W. 2002, *Cyberphilosophy: the intersection of philosophy and computing* (Oxford: Blackwell).
- Nilsson, N. (2000) *Inteligencia Artificial. Una nueva síntesis*, Madrid: McGraw-Hill, Madrid.
- Pagels, H. R. (editor) 1989, *La cultura dei computer* (Torino: Boringhieri).
- Prata, S. (1993) *Vida Artificial*, Madrid: Anaya.
- Rich, E. y Knight, K (1994) *Inteligencia Artificial*. McGraw-Hill.
- Russell, S. y Norvig, P. (1996) *Inteligencia Artificial. Un enfoque moderno*. Prentice Hall..
- Turing, A. M. 1992-2001, *Collected works of a.M. Turing*, IV vols. (Amsterdam-London: North-Holland).

¹ Para un desarrollo más amplio remito a Floridi (2003b).

² Ver Turing (1992-2001).

³ N. del T.: bajo la supervisión y consejo del professor Floridi, la bibliografía original sobre la IA en italiano ha sido substituida para beneficio del lector hispanoamericano por textos disponibles sobre el tema en castellano. En este idioma contamos con Anderson (1984), Boden (1984), Castellfranchi y Stock (2002), Nilsson (2000), Prata (1993), Rich y Knight (1994) y Coello (2003).

⁴ Sobre el impacto cultural de la informática véase Bolter (1985) y Pagels (1989).

⁵ N. del T. íbid nota a pie nº 3, en referencia a autores hispanoamericanos que han introducido la IA a sus investigaciones: Borrajo (1993), Fernández (1992). Ginsberg (1993) és un autor muy interesante, aunque únicamente disponible en inglés.

⁶ Para una reseña ver Ford et al. (1995) Bynum & Moor (2000), Moor & Bynum (2002), Colburn (2000) y Floridi (1999b); para una discusión de los problemas abiertos en la FI, vean Floridi (2004b); la más reciente y detallada introducción a la filosofía de la información se encuentra en Floridi (2003a).

⁷ N.del T. El problema de Gettier (1963) es un problema clásico de la epistemología filosófica. Fue publicado como artículo en 1963: "Is Justified True Belief Knowledge?", *Analysis*, 23: 121-123. La respuesta de Gettier al título de su artículo (el JTB, según el acrónimo de los analíticos para *Justified True Belief*) era que no era posible. Según la filosofía clásica, el conocimiento se identifica con la creencia, la verdad y la justificación epistémica. Mediante 2 contraejemplos, Gettier demostró que esta triádica identificación conceptual debía ser abandonada. En ambos contraejemplos, un sujeto *S* sostiene una creencia verdadera y justificada en un

contenido proposicional p , aunque no obstante no nos es posible afirmar que S sabe que p . Además, puede suceder que S crea que p , i que p tenga una relación con q . Si sucede que p es cierto, no tiene por qué serlo también q . Veamos los dos contraejemplos de Gettier: (1) Caso I. Juan se presenta a una entrevista de trabajo, aunque cree de forma justificada que “Pedro conseguirá el trabajo”. Juan tiene una creencia justificada en que “Pedro tiene 5 monedas en su bolsillo”. Consecuentemente, y por la transtitividad de la identidad, Juan cree en tercer lugar que “la persona que tiene 5 monedas en su bolsillo obtendrá el trabajo”. Finalmente, Pedro no es aceptado, y lo obtiene Juan, quien también tenía 5 monedas en el bolsillo. La tercera creencia de Juan, por lo tanto es verdadera y justificada, aunque no por ello es una fuente de conocimiento; (2) Caso II. Juan tiene la creencia justificada que “Pedro posee un bolígrafo”. Mediante la regla de la introducción de la disyunción, Juan cree que “Pedro tiene un bolígrafo o Pedro se encuentra en Barcelona”, aunque lo cierto es que Juan no sabe si Pedro se encuentra en Barcelona. De hecho, Pedro no tiene un bolígrafo, pero sí que se encuentra por casualidad en Barcelona. Una vez más, Juan tenía una creencia que es verdadera y justificada pero que no constituía conocimiento.