

## APUNTES EPISTEMOLÓGICOS A LA E-CIENCIA

Jordi Vallverdú  
Universitat Autònoma de Barcelona\*  
jordi.vallverdu@uab.es.

### Resumen

En los inicios del siglo XXI está desarrollándose una e-ciencia, una ciencia electrónica y altamente computarizada que exige un replanteamiento sobre la epistemología científica. A través del ejemplo de la Bioinformática y las Biotecnologías, el autor muestra algunas características de esta nueva e-ciencia e indica algunos de los problemas con los que deben enfrentarse los filósofos de la ciencia contemporáneos.


PALABRAS CLAVE: e-ciencia, epistemología, computación, bioinformática, cognición.

### Abstract

*Right at the beginning of the 21st century an e-Science is emerging, a highly computerized electronic science which demands a new way of understanding scientific epistemology. Through Bioinformatics and Biotechnologies, the author analyzes some of the characteristics of this new e-Science and shows several problems which contemporary philosophers of science must face.*

KEYWORDS: e-Science, epistemology, computation, bioinformatics, cognition.

### 1. Introducción

 El 16 de febrero del año 2001, la revista *Science* publicó un artículo especial que llevaba por título “The Sequence of the Human Genome” y que significó la presentación del primer mapa del genoma humano, constituido por más de 3 000 millones de bases (A, C, G, T). El artículo estaba firmado por 275 científicos procedentes de 14 instituciones científicas repartidas por todo el mundo. Los costes finales del proyecto pueden evaluarse en un número de centenas de millones de euros, muy elevado. Al tratarse de experimentación con seres humanos, se tuvo que respetar las

\* Esta investigación ha sido parcialmente financiada por la Fundació Víctor Grífols i Lucas, dentro del proyecto “e-Biotecnología: Simbiosis de valores”.

políticas de diversas agencias norteamericanas, así como de la World Medical Association, especialmente de la Declaración de Helsinki (revisada el año 2000). A fin de tener criterios unificados, se creó el Institutional Review Board (IRB) para que contribuyera en el diseño del protocolo para la obtención del ADN y que contemplara los derechos de privacidad y confidencialidad de los donantes. Celera Genomics obtuvo un Certificado de Confidencialidad del Department of Health and Human Services y pudo empezar sus investigaciones. Se tomaron muestras de diversos voluntarios, de los que finalmente se escogieron 5 (dos hombres y tres mujeres): un afroamericano, un asiático chino, un hispano mexicano y dos caucásicos. De esta manera se pretendía no hacer ninguna discriminación de género ni de raza.

El proyecto, además de estas cuestiones de tipo ético, tenía tres grandes retos: por un lado, secuenciar rápida y eficientemente, por otro, almacenar estos datos y, finalmente, encontrar sentido a la información obtenida y catalogada. En estos tres pasos tenemos la presencia completa y absoluta de sistemas computacionales o robotizados: ABI 3700 para la primera fase, GenBank y el algoritmo BLAST para la segunda y el sistema experto Otto para la tercera (identificaba y caracterizaba genes del genoma humano). Todas ellas no son tareas meramente mecánicas, algunas de ellas requieren el desarrollo de un pensamiento artificial similar al del ser humano, aunque más rápido, fuerte y preciso. *Citius, altius, fortius*, pero para los sistemas computacionales<sup>1</sup>. Los seres humanos estamos descubriendo nuestra esencia genética con la ayuda de máquinas que empiezan a ser inteligentes. Al mismo tiempo, el almacenamiento y trabajo con las bases de datos existentes requería de *software* y *middleware* específico. Bienvenidos a la e-ciencia.

## 2. Definiendo a la e-ciencia

Cuando hablamos de e-ciencia, y según la Fundación Española de Ciencia y Tecnología (FECYT)<sup>2</sup>, nos referimos “a las actividades científicas a muy gran escala que deberán desarrollarse cada vez más mediante colaboraciones globales distribuidas y accesibles a través de Internet. Este tipo de actividad científica cooperativa<sup>3</sup> requerirá acceso a bancos de datos muy voluminosos y a recursos de computación de muy gran escala, además de prestaciones de visualización de alta calidad y otro tipo de herramientas”.

<sup>1</sup> O lo que viene a ser lo equivalente en el ámbito de la ciencia computacional: faster, smaller, cheaper. (Vaughan 1996, p. 28).

<sup>2</sup> <http://www.fecyt.es/e-ciencia/>.

<sup>3</sup> (Gainer 2002) incluye las herramientas computacionales en una ciencia más efectiva e incluso añade un aspecto más a tener en cuenta para los espacios físicos donde hay importante concentración de investigadores: los arquitectónicos. Entre los centros de investigación que han planeado mejor los espacios físicos encontramos el Weatherall Institute of Molecular Medicine (WIMM), en el John Radcliffe Hospital (Headington, Oxford, UK).

La pregunta clave, tras lo expuesto en este capítulo, es si la e-ciencia es simplemente ciencia clásica conectada a la red, o si más bien nos encontramos ante nuevos modos de producir, gestionar y analizar la información científica a través de innovadoras formas de trabajo a distancia.

Una mirada a las argumentaciones realizadas hasta este momento, nos impulsan a afirmar que nos encontramos, efectivamente, ante una e-ciencia, una nueva forma de crear, evaluar y difundir el conocimiento. Con ello no pretendemos afirmar que es algo radicalmente nuevo y que no existe ligazón alguna con el período previo. Lo que importa es admitir que la relación entre mundo-conocimiento-investigadores ha sido radicalmente transformada. La epistemología, el estudio de los modos de adquirir y crear conocimiento válido, debe incluir estas nuevas variables en nuestros análisis, por lo que, extendiendo el símil, estamos ante una e-epistemología.

Toda nuestra reflexión gira en torno a la idea de ‘información’, el concepto clave de las teorías sociológicas<sup>4</sup> que analizan la dirección universal tomada a finales del siglo XX e inicios del XXI, que remite a diversos fenómenos a considerar sobre la naturaleza de la ciencia contemporánea:

<sup>4</sup> (Castells 2000). En el ámbito de la Filosofía de la Ciencia, remito a la especial obra de (Floridi 2002, 2003), quien ha desarrollado una Filosofía de la Información interesante desde un punto de vista epistemológico.

Proceso	Concreciones
Creación /Descubrimiento	Data Tsunami: Petabytes de datos. Instrumentos virtuales. Ontologías <sup>5</sup> . Inteligencia Artificial (IA).
Gestión: búsqueda- acceso-traslado- manipulación- minería	Bases de datos: Complejas, jerárquicas, dinámicas, inconsistentes. <i>Software</i> . <i>Middleware</i> .
Comprensión	Modelización computarizada <sup>6</sup> . Visualización (entornos amigos). Integración de la información (gen- sociedad).
Evaluación	Computacional, abierta.
Comunicación	Revistas electrónicas gratuitas <sup>7</sup> .
Trabajo	Deslocalizado. En red. Cooperativo. Dinámico. Interoperatividad. Cognición socialmente distribuida <sup>8</sup> .
Financiación	Pública-privada.
Control	Evitando regulaciones nacionales.

<sup>5</sup> Siempre y cuando estén unificadas, son más útiles para los investigadores (Ashburner et al. 2000).

<sup>6</sup> Para un estudio referente a las biotecnologías y el uso de modelos, remito a (Kohl et al. 2000), (Noble D. et al. 1999), (Noble & Colatsky 2000) y (Kitano 2002).

<sup>7</sup> (Malakoff 2004) cita la página electrónica de los Washington DC Principles for Free Access to Science [www.dcpinciples.org](http://www.dcpinciples.org). Sobre los problemas inherentes a este planteamiento, del tipo 'quien paga la infraestructura', 'pagar por publicar', etc., remito a los interesantes artículos de (Butler 2003) y (Owens 2003). La Public Library of Science (PLOS: <http://www.publiibraryofscience.org>) es un ejemplo de ello (Terry 2005). La PLoS ha sido desarrollada por una fundación privada, una charity británica, el Wellcome Trust. La idea, es que la investigación con financiación pública (gobierno) o de vocación pública (fundaciones privadas, ONG) pueda servir a la comunidad social, mediante un acceso libre y gratuito. Es decir, que la información científica sea de dominio público y no protegida mediante suscripciones comerciales, como sucede con las revistas científicas más prestigiosas. Con el modelo propuesto, el *peer review* se mantiene, otorgando credibilidad a los artículos ofrecidos, y el canal mediante el cual estos artículos llegan al público no es el papel sino el bit: la Red abarata los costes. Podríamos decir que nos encontramos ante una revolución en los procesos de comunicación similar al que supuso la imprenta de Gutenberg. Bajos costes y facilidad de reproducción implicó e implica una mayor difusión de la información, lo que se traduce en mayor actividad y, a la postre, en cambios significativos en los modelos sobre el mundo. Inició sus propias revistas en los meses de octubre 2003, *PLoS Biology*, y 2004, *PLoS Medicine*, respectivamente. El modelo anterior, en el que el autor cedía gratuitamente el copyright de su

Es un hecho, la cantidad de información con la que deben trabajar los investigadores de nuestros días es de una magnitud tal, que requiere de nuevas formas de trabajar con los datos<sup>9</sup>.

Esta e-ciencia no es solo un deseo pro-tecnológico, sino más bien la respuesta a las necesidades de la investigación en múltiples disciplinas de finales de los años noventa del siglo XX. No es por azar que Tim Berners-Lee inventó la World Wide Web en el CERN: ya por entonces, los físicos se encontraban con grandes cantidades de datos que necesitaban coordinar con colegas de otros países, becarios post-doctorales desplazados, técnicos especializados no residentes en Suiza,... Sin embargo, tampoco debemos pensar que la creación de estos servicios es algo habitual en la estructura de la ciencia. Como en tiempos anteriores, lo que beneficia curricularmente a un investigador es el trabajo propio en el laboratorio, y no la creación de herramientas compartibles en red. Pero por lo visto hasta este momento, somos capaces de advertir la crucial importancia de la creación de *soft-* y *middleware*. Es necesario, pues, volver a plantearnos las características recompensables en la carrera científica, siempre desde un punto de vista social, es decir, considerando la ciencia como un proyecto comunitario que debe su evolución y progreso a muchas variables.

En junio de 2005, Vinton G. Cerf y Robert E. Kahn recibieron el premio A.M. Turing, el premio sobre computación más importante a nivel internacional. Aunque se trata de un premio importante, el A.M. Turing Award no tiene el mismo prestigio que un Nobel. Y lo cierto es que lo que hicieron Cerf y Kahn, crear en 1973 el Transmission Control Protocol (TCP), la base de Internet (que no llegó al dominio público hasta la creación de un programa, Mosaic, el cual permitió a no expertos navegar por la red) ha contribuido más por la transformación de la ciencia y las sociedades contemporáneas que la suma del trabajo de muchos especialistas de una miríada de pequeñas disciplinas. Y no se puede caer en la demagogia, diciendo que por la misma razón deberíamos premiar a los ingenieros que diseñaron las carreteras por las que circulan

---

artículo a las revistas, quienes utilizan esta información para crear materiales que venden casi de forma exclusiva a las bibliotecas (prácticamente todas de tipo público o, si no, pertenecientes a entidades con fines sociales), con lo que cubren gastos y obtienen un beneficio, incluso del 30%, que nunca retorna a los autores o los expertos consultados durante el proceso de revisión. Además, al ser un mercado en el que tanto productores como consumidores tienen la sensación que no hay costes (porque las suscripciones las pagan las instituciones a las que pertenecen) ni beneficios (porque no cobran por producir la información contenida), las editoriales científicas han aumentado drásticamente los precios de las publicaciones. En 10 años, las suscripciones en el Reino Unido han aumentado de precio un 200%. Y con los libros de editoriales como Blackwell, Elsevier o Kluwer, entre muchas otras, ha sucedido lo mismo: se publican libros en ediciones de por sí muy limitadas y a unos precios exorbitantes que los investigadores no pueden adquirir de forma particular y que acaban por encargarlos a las bibliotecas de sus universidades o centros de investigación.

<sup>8</sup> Estany 2000.

<sup>9</sup> Ver gráfico de petabytes de datos en creciente aumento en <http://binfo.ym.edu.tw/edu/seminars/seminar-041002.pdf>.

los científicos, los aviones que les sirven para asistir a congresos o los sistemas eléctricos que utilizan a lo largo del día. Es cierto, estas son necesidades diarias que constituyen la infraestructura básica de nuestros días, pero en el caso de Internet y la computación, no es una característica más, sino la base completa de una nueva generación de prácticas científicas.

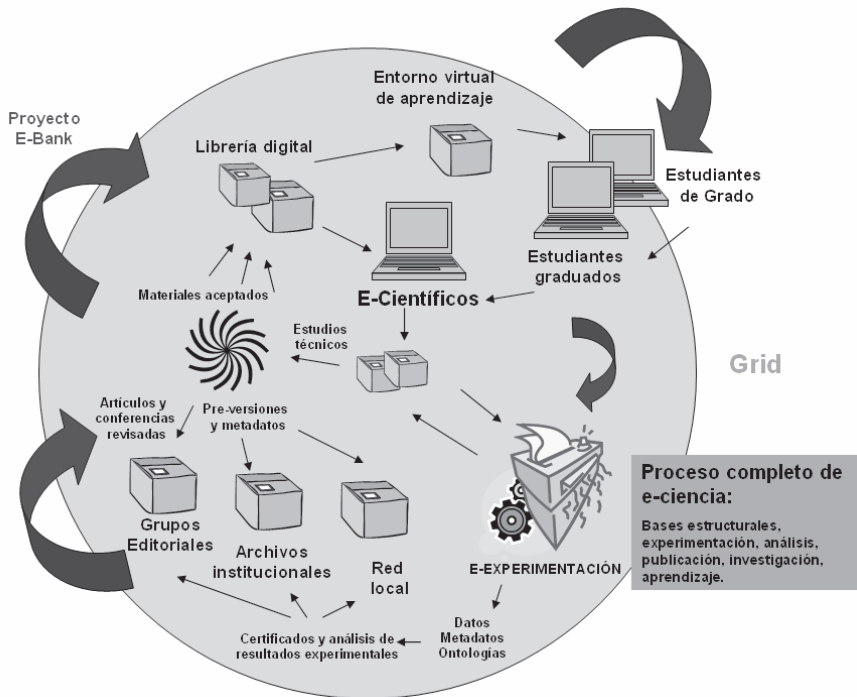
La e-ciencia implica optimizar el proceso completo de la producción científica, desde el laboratorio, hasta la difusión, la formación de nuevos especialistas. Por este motivo, los científicos dedicados a la creación de estas herramientas específicas de comunicación deben ser considerados piezas fundamentales del mecanismo de la actividad científica. Además, tampoco debemos caer en el viejo prejuicio de: “eso son ingenieros/técnicos”<sup>10</sup> que realizan un proceso meramente mecánico. Nada más alejado de la realidad: los especialistas en bioinformática son al mismo tiempo informáticos y biólogos, y deben pensar, por no decir construir, las formas de analizar la realidad biológica más eficientes: ontologías, software, middleware, bases de datos, lenguajes específicos, ... es una tarea titánica que exige grandes conocimientos. Sobre esta base se construye la realidad de la ciencia contemporánea y sus continuos descubrimientos<sup>11</sup>.

Es por eso que en el ámbito de las biotecnologías, es tan importante reflexionar sobre el papel y la formación necesaria de los profesionales<sup>12</sup>. Veamos un esquema extraído del proyecto británico eBank (revista *Science*, vol. 308, p. 820):

<sup>10</sup> Es lo que ha sucedido desde las disciplinas más abstractas, como las matemáticas y la física teórica, hacia las prácticas, como la ingeniería. Remito por ejemplo a las opiniones despreciativas de un matemático como J.F. Nash hacia los ingenieros de la RAND (Nasar 1999). Al fin y al cabo, unos sin otros no consiguen nada.

<sup>11</sup> Langley (2000) especifica en el punto 4 de su artículo (“Computer-aided scientific discoveries”) algunos ejemplos de ello. Sus reflexiones sobre las relaciones entre computación y descubrimiento científico son muy relevantes para nuestro caso, además de su trabajo en inteligencia artificial.

<sup>12</sup> Wickware (2000, p. 683), por ejemplo, nos dice: “the ideal next-generation biologists should develop both wet-lab expertise and software-writing ability”.



### Proceso de e-ciencia, desde la e-investigación hasta el e-aprendizaje

La e-ciencia está considerada aquí como un ciclo, un modelo de optimización del trabajo más eficiente que el anterior, al tiempo que implica un cambio en los modos de considerar la propia realidad, ahora acotada por los modelos informáticos. En el caso de la biología *in silico*, no siempre se parte de un laboratorio, sino que se realizan experimentos virtuales con la información existente en las bases de datos, a los que se aplican análisis computacionales, de manera que todo el proceso es virtual: los datos y las herramientas son bits de un espacio electrónico. Más fácil, más rápido, más barato. Por estas tres razones, situadas en su correcto contexto, la e-ciencia es el nuevo paradigma. En el caso de la biología, ha conducido a una biología cuantificada, experimental, modelizable y predictiva<sup>13</sup>, además de, o debido a ser profundamente computarizada. Y este esquema abarca tanto a la biotecnología pública como la privada. La dirección de Aventis<sup>14</sup> contempla la creación de un sistema global que permita

<sup>13</sup> Nölting et al. (2004) relaciona la programación evolutiva computarizada de las secuencias de proteínas y las predicciones de estructuras.

<sup>14</sup> www.aventis.com, una multinacional resultado de la fusión entre la alemana Hoechst AG y la francesa Rhone-Poulenc S. A.

a los científicos de sus empresas dispersados a lo largo y ancho del mundo acceder a las herramientas y datos (heterogéneos) necesarios de manera eficiente. Una sección de IBM, IBM Life Sciences Solutions, ha desarrollado una arquitectura de soporte para ello, además de *middleware* especial, el IBM DiscoveryLink.

Pero no por ello está la e-ciencia exenta de problemas. Uno de los más acuciantes de esta era bioinformática es el de la reproducibilidad de los resultados computacionales. En los modelos clásicos de comprensión de la actividad científica, se ha considerado desde el Renacimiento que la ciencia es algo experimental y que su éxito explicativo y predictivo pasa por la posibilidad de reproducir los experimentos que demuestran alguna hipótesis o teoría. En el caso de las biotecnologías, con una alta implementación computacional, es difícil tener la oportunidad de poder reproducir un experimento virtual. Y los motivos son múltiples: el uso de estándares no unificados de almacenamiento de los datos, algoritmos de cálculo y procesamiento variopintos, un sistema simple y eficaz para poder compartir a distancia estos datos, tanto los iniciales como los finalmente resultantes... Pero encontramos casos de falsaciones realizadas en la ciencia computacional, como en el caso del célebre experimento de Knight & Leveson (1986) y su análisis de las probabilidades de error de los programas multiversión. La teoría clásica predice que la probabilidad de error de un programa multiversión era el producto de las probabilidades de error de las versiones individuales. Sin embargo, Knight & Leveson observaron que los programas multiversión tenían probabilidades de error significativamente más elevadas, lo que falsaba la idea clásica que afirmaba que los errores de las versiones de un programa son estadísticamente independientes<sup>15</sup>.

Además, la ciencia experimental física y la virtual no se excluyen, sino más bien complementan la una a la otra<sup>16</sup>. En algunos casos, no obstante, la creciente automatización y robotización de las prácticas humanas parece que desea extenderse hacia el ámbito de la ciencia<sup>17</sup>.

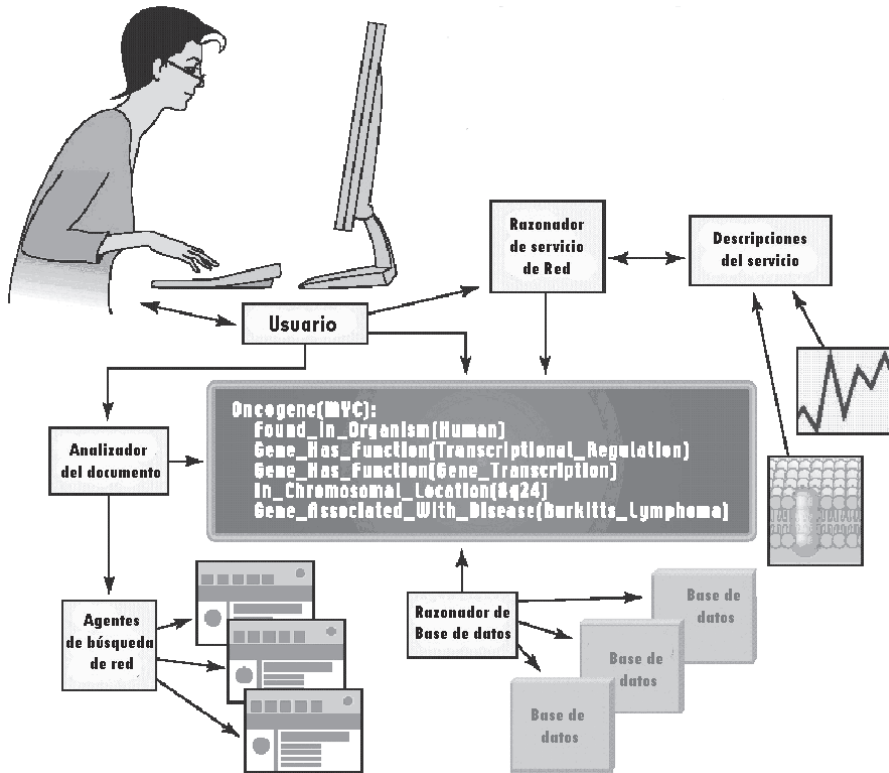
La e-ciencia supone un cambio en los modos de trabajar que ha tornado el proyecto científico más rápido y competitivo. Veamos el ejemplo de Hendler (2003, p. 520):

<sup>15</sup> Tichy 1998, p 33.

<sup>16</sup> Long et al. (2003) estudia los casos de formación de nuevos genes mediante análisis tanto empíricos como computacionales.

<sup>17</sup> Langley, Simon, Bradshaw & Zytkow 1987.





La usuaria/científica está disfrutando del sueño de Vannevar Bush, el pionero del hipertexto y de la idea de interconectividad y acceso inmediato a diversos tipos de datos. En el número de julio de 1945 de la revista *The Atlantic Monthly*, Bush estableció en su artículo “As We May Think”<sup>18</sup> (Deemer 1994) los conceptos fundamentales del hipertexto. Bush se refería a una máquina imaginaria, *memex*, en la que un individuo podría almacenar sus libros, discos y otros documentos, de forma mecanizada que permitiera su rápida recuperación para su uso. Memex parecía un escritorio con dos monitores de acceso táctil y una superficie de escáner. Toda la información podía estar indexada de forma clara y simple.

<sup>18</sup> Descargable en: <http://www.ps.uni-sb.de/~duchier/pub/vbush/vbush-all.shtml>. Existe, no obstante, su versión en papel: Nelson (1972).

El software de visualización<sup>19</sup>, *imaging* en el argot internacional, está produciendo el mismo efecto en la biología molecular de finales de los noventa, como hizo por otras disciplinas como las matemáticas<sup>20</sup> o la medicina<sup>21</sup>. Al aproximar el funcionamiento de las máquinas a las características de los seres humanos, éstas son incorporadas con mayor rapidez, facilidad y eficacia<sup>22</sup>. Debemos pensar que el 60% de la información que recibe nuestro cerebro procede de la visión<sup>23</sup> y que, por lo tanto, la visualización de los datos que obtenemos del mundo condiciona nuestras teorías sobre el mismo. Proyectos internacionales como Physiome Project<sup>24</sup> intentan desarrollar meta-lenguajes de programación que permitan trabajar de una forma más intuitiva a los programadores/usuarios y que sean de fácil conectividad entre sí. CellML<sup>25</sup> o SBML<sup>26</sup> son ejemplos de lo comentado.

En el ejemplo que Hender nos ofrece, la máquina es al mismo tiempo física (hardware y software del PC) y virtual (middleware, bases de datos, Internet)<sup>27</sup>, y permite a la investigadora efectuar diversos procesos necesarios para la consecución de nuevo conocimiento sin tener que desplazarse continuamente a espacios físicos alejados y con funcionamientos propios que exigen una formación exhaustiva. La eficiencia es una optimización máxima de las prácticas científicas, lo que permite generar no solo nuevo conocimiento a un menor coste económico y personal, sino también nuevas formas de crear el propio conocimiento.

<sup>19</sup> Una empresa dedicada a software para visualización es BioImaging Systems ([www.uvp.com](http://www.uvp.com)). Remito a la obra de Wolf y Yaeger 1993.

<sup>20</sup> Honrad Polthier, de la Technische Universität de Berlin ha diseñado el *Java View* que puede funcionar en cualquier PC con conexión a Internet, que es freeware y que permite visualizar las matemáticas.

<sup>21</sup> Desco y Vaquero 2004.

<sup>22</sup> Remito a los estudios sobre visualización y comprensión de Robert Horn, que pueden encontrarse en la dirección [www.macrovu.com](http://www.macrovu.com). Asimismo, recomiendo su libro *Visual Language. Global Communication for the 21st Century*. Y, aunque antiguo, no debemos olvidar el agudo trabajo de Latour (1986).

<sup>23</sup> Humphreys 2004, p. 112.

<sup>24</sup> [www.physiome.org](http://www.physiome.org).

<sup>25</sup> [www.cellml.org](http://www.cellml.org).

<sup>26</sup> [www.cds.caltech.edu/erato](http://www.cds.caltech.edu/erato).

<sup>27</sup> Si bien es cierto que la distinción físico-virtual se desvanece bajo la consideración de la completa fisicalidad de los eventos del universo, no está de más poder distinguir entre los soportes físicos (hardware) y los sistemas moldeables del manejo de la información de los mismos (software). Tal distinción significó un hito el desarrollo de máquinas programables, bajo la guía de Turing, y es útil todavía para la comprensión del funcionamiento de las computadoras y el control y asignación de los errores de su funcionamiento.

### 3. El caso de la bioinformática

Si bien el concepto de información puede parecer algo ajeno a la biología, tan solo presente a través de acumulación de datos en las bases electrónicas, tiene un profundo significado biológico<sup>28</sup>: la idea central de la biología contemporánea es la de la información. El ADN es un conjunto de datos químicos que, transmitido de generación en generación, contribuye decisivamente al desarrollo de los seres vivos individuales y a la evolución de las poblaciones formadas por ellos.

Si analizamos los términos que utilizan los biólogos contemporáneos, veremos que la idea de información ha calado muy a fondo en el conocimiento científico: transcripción, traducción, código, redundancia, sinónimo, mensajero, editar, corrección, biblioteca... son términos que aparecen continuamente en los textos biológicos contemporáneos. Y las analogías, como formas de estimular las mentes científicas siempre son bienvenidas si conducen a resultados palpables e importantes.

Se podría criticar el uso de un término físico en el ámbito de lo biológico argumentando que la teoría de la información de Shannon no se puede aplicar en este contexto. En 1948, C.E. Shannon, un joven y brillante ingeniero eléctrico y matemático que trabajaba para una empresa privada que ofreció algunos premios Nobel a la par que grandes beneficios económicos, los laboratorios Bell, publicó en el *Bell System Technical Journal* un artículo que llevaba por nombre, "A mathematical theory of communication"<sup>29</sup>. Allí exponía su teoría de la información. Para Shannon, la medida de la cantidad de información,  $\Sigma P \log P$  es la medida de la capacidad de un canal para transmitir información, a partir del número de mensajes que puede transmitir.

Sus ideas han tenido una amplia difusión en la informática y la física, incluso en neurobiología y raramente en genética. L.L. Gatlin escribió en el año 1972 un libro la teoría de Shannon a la genética<sup>30</sup>, teniendo claro que el problema no es la transmisión de la información sino más bien su significado. Pero en cierto punto de este debate nos adentramos en un mundo de debates académicos que se escapan del objetivo de la presente obra. Volvamos a la noción básica de información biológica, como datos creados por los investigadores a partir de sus interacciones con los objetos de estudio, en nuestro caso, el mundo de lo vivo desde el ADN hasta las sociedades. Existe, no obstante, un tipo de recursos tecnológicos que han determinado el curso de la disciplina: se trata de la computación, que incluso ha dado el nombre a una nueva disciplina, la *bioinformática*. La bioinformática<sup>31</sup> consiste en aproximaciones computacionales o algorítmicas al análisis e integración de datos genómicos, proteómicos o

<sup>28</sup> (Smith 2000). Gran parte de las argumentaciones siguientes se fundamentan en este agudo artículo.

<sup>29</sup> vol. 27, páginas 379-423 y 623-656, July and October, 1948.

<sup>30</sup> Gatlin, L.L. 1972, *Information Theory and the Living System*, Columbia University Press, NY.

<sup>31</sup> <http://bioinformatics.org/>, es una página 'clásica' al respecto. Ver Simon 2005.

químicos que se encuentran archivados en bases de datos, de acuerdo con Studt (2003). El problema fundamental de la biología contemporánea: la gran cantidad de datos a tener en cuenta o síndrome HAOD<sup>32</sup>. Obtener sentido de tal información requiere nuevas formas de pensar e, incluso, nuevas metodologías de trabajo, afirma Hadley (2004). La bioinformática ha permitido la cuantificación de la vida y la transformación de los datos en modelos (computarizados), al tiempo que permite la integración mayor entre disciplinas anteriormente separadas, como la informática de la biodiversidad (salud, genoma, ...) la neuroinformática<sup>33</sup>.

Uno de los problemas al analizar la naturaleza de la bioinformática es hasta qué punto la bioinformática es una disciplina autónoma o no se trata de nada más que de la aplicación de herramientas informáticas a problemas biológicos, donde primaría el ámbito de la biología como disciplina fundamental<sup>34</sup>. Mi tesis es que la bioinformática<sup>35</sup> es una disciplina autónoma que ha seguido los pasos de otras disciplinas clásicas y ha propiciado un cambio radical en la forma de hacer ciencia y, por ello, en los increíbles resultados obtenidos. Por lo que refiere a la comunicación entre grupos de investigación alejados entre sí, el cambio ha sido absoluto.

La computarización de la biología ha conllevado asimismo un cambio en las formas de pensar, actuar y decidir de la comunidad biológica. En cierto modo, estamos ante una nueva biología, centrada en la obtención y procesamiento cuantificado de datos biológicos, propicia a análisis inductivos y a la simulación sistemática de fenómenos vivos. También es una disciplina en la que es difícil distinguir entre teorías y experimentos, entre ciencia básica y ciencia aplicada, estando las divisiones clásicas fuera de lugar<sup>36</sup>. En resumen, una biología mucho más teórica y rica, al tiempo que inextricablemente unida a las herramientas informáticas y estadísticas<sup>37</sup>. Los tubos de ensayos son virtuales, como *Edna*, y trabajan con eDNA, o ADN electrónico<sup>38</sup>, en una nueva computación biomolecular<sup>39</sup>. La simple posibilidad de ‘visualizar’

<sup>32</sup> Acrónimo de *Huge Amount Of Data* (acrónimo del autor). Debemos pensar que la cantidad de datos a procesar en la genómica contemporánea es inmensa. En el campo de la física, el Large Hadron Collider del CERN, cerca de Génova, produce 15 *petabytes* (15 millones de *Gigabytes*) de datos al año. Con la finalidad de poder analizar estos datos se ha creado el LHC Computer Grid (<http://lcg.web.cern.ch/LCG>).

<sup>33</sup> Heidorn 2003. Consultar el número completo.

<sup>34</sup> Ouzounis, 2002.

<sup>35</sup> Se la puede denominar también, ciberbiología, biología computacional, e-biología, biología digital, *web-based biology*... justamente por esa multiplicidad de nombres estamos ante un paradigma emergente.

<sup>36</sup> Keller 2000.

<sup>37</sup> Baldi & Brunak 2001. Por lo que respecta al tratamiento informático de los datos biológicos, existe un dominio claro del uso del modelo bayesiano (Modelos Markov,...) en detrimento del frecuentista. Y sin demasiados debates de tipo conceptual, tan solo basados en planteamientos práctico-operativos.

<sup>38</sup> Garzon et al. 2003.

<sup>39</sup> Algunos ejemplos de laboratorios virtuales: <http://jersey.uoregon.edu/vlab>, <http://curry.edschool.virginia.edu/go/frog>. Sobre la naturaleza de tales experimentos, es decir, si se trata de

tridimensionalmente proteínas y genes ha permitido pensar la realidad desde una perspectiva agradable a la mayor parte de los seres humanos<sup>40</sup>. Aunque pueda parecer paradójico lo que voy a decir, *la incesante computarización de la ciencia y, por inclusión, de la biología, ha supuesto una humanización de la misma*. Diseñamos estas herramientas para que sean sustitutos de nuestros sentidos más directos, y propiciamos la creación de simulaciones que nos permiten ‘ver’ el micromundo para pensarlo mejor. Somos una especie basada en la visión, primero, y en el lenguaje simbólico, después. Nuestras metáforas sobre los límites del conocimiento son siempre visuales. Y no es una forma actual de pensar: un clásico de la Edad Media, como el argumento ontológico de San Anselmo<sup>41</sup> es un ejemplo de lo expuesto. Además, las metáforas biológicas en el mundo computacional son múltiples: virus informáticos, *bugs*, autómatas celulares, computadoras genéticas<sup>42</sup>, algoritmos genéticos, el sistema regulatorio genético como computadora química<sup>43</sup> el ADN como ‘programa de ordenador’ (Schrödinger lo comparó con el código morse)<sup>44</sup> ...<sup>45</sup>

La biología computacional ha demostrado su utilidad al cumplir con creces una de las condiciones de toda teoría correcta (o útil) sobre el mundo: la predictibilidad. La bioinformática actual ofrece una capacidad de predicción tal que ha permitido

---

ayudas a la investigación, de soporte a los experimentos físicos, si su naturaleza es real o si son meros *Gedankenexperimente* útiles para la reflexión, remito a los estudios sobre las consideraciones epistemológicas de las simulaciones de Varenne (2001), Winsberg (2003) y Frigg & Hartmann (2006).

<sup>40</sup> En el caso de la unión entre visualización en 3D, anatomía y bioinformática, ha aparecido lo que se ha denominado la ‘Anatomica’ (*anatomics*), una anatomía computarizada muy eficiente para la formación y precisión de investigadores (Bard 2005). Algunos enlaces importantes: [www.opengalen.com](http://www.opengalen.com), [www.physiome.org](http://www.physiome.org), [www.xspan.org](http://www.xspan.org)...

<sup>41</sup> San Anselmo fue un teólogo del siglo XI que intentaba conciliar fe y razón, adoptando el *Credo ut intelligam* (creo para comprender) de San Agustín de Hipona. Intentando demostrar racionalmente la existencia de Dios escribió dos obras, el *Monologium* y el *Proslogion*. Fue en esta última obra, que desarrolló su argumento ontológico, es decir, su razonamiento sobre el origen (divino) de todas las cosas. El argumento consiste en pensar aquello más grande que puede venir a la mente. Pues o bien eso es Dios o bien hay algo más grande que eso pensado en un primer momento, que es Dios. Se trata del paso de la idea de algo a su existencia. Según San Anselmo, no era posible que la idea más perfecta de algo no existiera, porque entonces, argumentaba, ¿de dónde podría haber salido? Es uno de los tantos experimentos visuales-mentales de los pensadores universales.

<sup>42</sup> Keller (2000, p. S77) expone esta metáfora sobre el promotor *Endo16* aparecida en un artículo de la revista *Science*. La idea de fondo de Keller es que si bien hasta hace poco estas metáforas eran tan solo eso, metáforas, en estos momentos el análisis del funcionamiento de los mecanismos básicos de la vida están atrayendo la atención de especialistas en computación y robótica, lo que produce a su vez una ‘maquinación de la vida’ o una implementación de lo vivo en lo mecánico. Los biochips (o los *gene chips*) son un ejemplo de lo que estoy diciendo.

<sup>43</sup> Richardson (1997, p. S259). El autor va incluso más lejos al considerar los genes, al presentar disposiciones de activo/inactivo, como nodos de una red Booleana.

<sup>44</sup> Fogle 1995, Van Dijk 1998, Condit 1999. Todos estos autores respasan las metáforas y significados públicos del genoma humano. Respecto a aquellas relativas a las relaciones entre humanos y máquinas remito a Cela y Marty (1994).

<sup>45</sup> Bentley 2002. O incluso en el de la sociobiología, con los *meme machines* de Blackmore (1999).

economizar la cantidad de experimentos necesarios para obtener conocimiento de lo vivo, lo que se traduce en un progreso más rápido y consistente.

Pero, aunque normalmente así lo parece, la cooperación entre ordenadores y científicos de la vida no es algo reciente (Ouzounis & Valencia 2003). Durante la última década de los noventa del siglo XX se produjo un crecimiento explosivo en la disciplina, pero debemos remontarnos hasta los años cincuenta y sesenta para encontrar sus orígenes. Fue entonces cuando los problemas básicos de la biología molecular mostraron su naturaleza algorítmica: la estructura del ADN, la codificación de la información genética por parte de las proteínas, las propiedades estructurales de las moléculas proteínicas y otros niveles de análisis eran susceptibles de ser tratados mediante modelos e instrumentos computacionales. Por aquel entonces estaba formándose la teoría de la computación por Chaitin (1966)<sup>46</sup>, la teoría de la información de Shannon y Weaver (1962), la definición de gramáticas de Chomsky (1959), la teoría de juegos de Neumann y Morgenstern (1953), con el *minimax*, y la idea del equilibrio de Nash (1950) y de los autómatas celulares por Neumann (1966).

Fue entonces durante los años sesenta que aparecieron numerosos desarrollos que justifican el nacimiento de la biología computacional<sup>47</sup>. El hecho es que cada vez más experimentos de la biología son llevados a cabo en modelos computarizados<sup>48</sup> o, una vez obtenidas ciertas predicciones simuladas en ordenadores son realizados experimentos reales con seres vivos.

Ordenadores y genoma constituyen un binomio inseparable. La genómica no sería posible sin ordenadores, no solo por la cantidad de datos que se tienen que procesar (ahora automatizados), sino por la dificultad cognitiva intrínseca a la comprensión de los mismos. Las máquinas nos ayudan a crear modelos que permiten el trabajo con los datos. De hecho, con las máquinas informatizadas secuenciamos, analizamos automáticamente, creamos el modelo y visualizamos los resultados, que son dispuestos en bases de datos abiertas a la red. Al mismo tiempo, la interconexión de estos ordenadores en red permite nuevos modelos de dinámica científica, en la cual la

<sup>46</sup> En las referencias temporales citadas hasta el final del párrafo he omitido la citación de las obras originales por tratarse de clásicos fácilmente accesibles y para no engrosar de manera indiscriminada el aparatado de referencias bibliográficas.

<sup>47</sup> Todavía hoy en día hay quien discute término preciso para denominar la disciplina. Si bien habitualmente se la denomina 'bioinformática', hay quien cree que este término hace referencia a una nueva especialidad propia, que debe diferenciarse de otra denominación, a saber, la de 'biología computacional', que se referiría más bien a la biología que recurre a técnicas informáticas para su desarrollo. A mi parecer, ambas expresiones son correctas e intercambiables, aunque debe recordarse que en cierta literatura se establece tal distinción.

<sup>48</sup> Remito a la explicación de Richardson (1997, p. S263) sobre los modelos simulados de la bacteriófago lambda ( $\lambda$ ) llevados a cabo a mediados de los noventa por Harley McAdams y Lucy Shapiro. Según Richardson, la simulación automatizada de un organismo permite a los investigadores utilizar de forma más eficiente el conocimiento actual, al tiempo que permite extender los límites del mismo hasta nuevos y más ricos resultados. ***El modelado computacional es útil para los experimentos, para los descubrimientos y para la formación académica.***

distancia y la accesibilidad a los datos son variables radicalmente transformadas de la práctica de los investigadores. Sin ordenadores, esta ciencia simplemente no sería posible. La computación es algo intrínseco ya a la investigación científica contemporánea en muchas disciplinas, no un mero instrumento o ayuda.

Podemos hablar incluso de ‘biología *in silico*’, es decir, sistemas biológicos virtuales creados a partir del uso de algoritmos computacionales, que nos permiten emular virtualmente mecanismos moleculares, células u organismos complejos con el subsiguiente beneficio en costos, simplificación del estudio y tiempo dedicado<sup>49</sup>. A nivel empresarial, estos mecanismos de computación han representado la reducción del tiempo de desarrollo de un nuevo fármaco en un 25% del total anterior, generando un recorte de gastos de 250 millones de dólares por medicamento<sup>50</sup>. La bioinformática no solo es útil, es altamente lucrativa.

Es lo que Gilks (2004) ha denominado como ‘nuevo paradigma’: la bioinformática permite nuevas formas de pensar la información. El clásico debate entre frecuentistas y bayesianos<sup>51</sup> se está diluyendo ante la potencia de los ordenadores y la mejor sintonía del bayesianismo en el ámbito de las biotecnologías. En consecuencia, las técnicas de Monte Carlo vía cadenas de Markov cada vez son más habituales. Incluso las redes bayesianas permiten automatizar mejor las tareas de descubrimiento causal automatizado<sup>52</sup>.

Otro aspecto a tener en cuenta es que en bioinformática los algoritmos son considerados más importantes que los propios modelos. Y, por encima de todo, es una cultura de acceso abierto. Según Hunter (2005), los biólogos dedican cada vez más tiempo en el laboratorio a la pantalla de un ordenador: tanto para acceder a la literatura reciente en bases de datos informatizadas, como en el diseño de los experimentos o el análisis de sus resultados. Por ello, los biólogos contemporáneos deben dominar el modelado matemático, el análisis estadístico o los algoritmos de procesamiento de imágenes y de visualización en 3D<sup>53</sup>. Al mismo tiempo, los informáticos y matemáticos que trabajan codo con codo con estos biólogos en tales laboratorios deben tener una formación básica en biología, lo que produce mejores y más rápidos resultados. Del mismo modo que los físicos encargados de las fotografías por cristalografía de rayos X contribuyeron al desciframiento del ADN, los matemáticos e informáticos del siglo XXI están contribuyendo al análisis profundo de lo vivo. Pero, por lo que respecta a la visualización, es cierto que Watson y Crick se beneficiaron de las imágenes de Wilkins y Franklin y que su representación tridimensional en modelo casi escolar resultó útil para el desarrollo de las investigaciones sobre la estructura del ADN,

<sup>49</sup> Álvarez et al. 2001, Hunter 2005.

<sup>50</sup> Tufts Center for the Study of Drug Development, <http://csdd.tufts.edu/>.

<sup>51</sup> Vallverdú 2004.

<sup>52</sup> Boden 1990.

<sup>53</sup> Marshall et al. 2002.

pero lo que sucede hoy en día es que las representaciones gráficas no son simplemente útiles, sino que se han tornado *necesarias* para el progreso científico<sup>54</sup>.

#### 4. *Mente extendida*<sup>55</sup> y *e-cognición*

El impacto y revolución causados por las tecnologías de la información y la computación en la ciencia contemporánea es algo que está fuera de toda duda. Del mismo modo que artefactos como el telescopio o el microscopio revolucionaron, respectivamente, la física y la biología, las diversas implementaciones computacionales están modificando sustancialmente numerosas ramas del saber. Podemos encontrar ejemplos en el desarrollo de la genómica, la astronomía y los múltiples proyectos de computación distribuida, los simuladores climáticos (Blue Gene, por poner un ejemplo), la inteligencia artificial, las bases de datos abiertas en línea o los modelos bayesianos computarizados en los estudios clínicos.

Las tecnologías de la información y computación están modificando el modo de producir conocimiento, al mismo tiempo que ofrecen nuevas formas de trabajo cooperativo (trabajo en red a distancia, instrumentos virtuales, computación distribuida, *middleware*...). En suma: están desarrollando un papel fundamental en los procesos actuales de dinámica científica, en los que la sociedad civil quiere y desea participar cada vez con más intensidad<sup>56</sup>. Gracias al cambio en el paradigma informacional y los modelos comunicativos, esto es ya posible.

Por decirlo de algún modo, nuestras mentes están extendiéndose a través de las herramientas computacionales, según Humphreys (2004), transformando la naturaleza misma de los procesos mentales considerados como un conjunto de operaciones que en estos momentos se encuentran distribuidas en espacios diversos (a caballo entre lo biológico y lo computacional). Nuestro conocimiento es creado a través de una red compleja de elementos que implican extensiones de nuestros sentidos (telescopios, coches, lápices,...) y nuestras mentes (herramientas computacionales).

Existe una larga serie de autores que han contribuido con sus modelos al desarrollo de un marco cognitivo que nos permita entender esta nueva racionalidad: Hutchins (1991, 1995), Clark (2003), Perkins (1986, 1992), Norman (1985, 1990, 1992, 1993, 1997, 2004) y Hutchins & Norman (1988). Al mismo tiempo existe un fuerte interés por la reflexión sobre el impacto de las ciencias de la computación en la filosofía y el pensar humano: Feigenbaum & Feldman (1963), Bynum & Moor (1998),

<sup>54</sup> Humphreys (2004, p. 113). El autor nos ofrece un claro ejemplo: un modelo sobre el comportamiento de un gas cerca de un agujero negro requiere de unas 10.000 operaciones que generan 1.25 mil millones de valores numéricos. Sin técnicas computerizadas de visualización, ninguna mente humana es capaz de considerar estos datos y realizar razonamientos con ellos.

<sup>55</sup> Término desarrollado por Clark & Chalmers 1998.

<sup>56</sup> Sunstein 2003.



Ford, Glamour & Hayes (1995), Mitcham (1994) en las relaciones filosofía-tecnología, Sloman (1978), Thagard (1988), Giere (2003) y Grim (2004).

En estos modelos, como el de Norman (2004), se consideran aspectos relativos a las emociones. Para realizar un buen diseño cognitivo sobre los procesos racionales, debemos contemplar el papel de las emociones. Parte de este cambio radical en el planteamiento del papel de las emociones lo debemos a los trabajos de neurofisiólogos como Damasio<sup>57</sup>. De hecho, en sus investigaciones, los científicos están expuestos a los que Paul Thagard ha denominado ‘variables cognitivas calientes’, tales como motivaciones o emociones<sup>58</sup>. Incluso un autor como Freedman argumenta que todavía debemos explicar por qué los científicos modifican los condicionantes cognitivos que, a su vez, dan lugar a las representaciones mentales, un proceso en el que los factores sociológicos desarrollan un importante papel, lo que exigiría una nueva axiología de la ciencia<sup>59</sup>: se ha podido demostrar la clásica falacia *ad hominem* en las percepciones que tienen los científicos de los resultados de las investigaciones de sus principales competidores, a los que rebajan los niveles de evidencia de los resultados.

Cuando nos enfrentamos al análisis de la e-ciencia, ¿no nos encontramos ante actitudes negativas o ingenuamente continuistas en un aspecto del análisis metodológico propicias a una falacia *ad machinam*?

##### 5. Concluyendo: ¿Hacia una e-epistemología?

¿Permiten todos estos indicios relativos a la generación, transformación, comunicación y evaluación de la información científica el plantearnos una transformación en el marco conceptual epistemológico? A mi parecer, esto sería lo más razonable. Y lo sería por diversos motivos:

- (1) Las características de la nueva e-ciencia mostradas en la tabla del inicio del artículo exigen una reconsideración de los modos de producción, control y difusión de la información científica.
- (2) El uso de herramientas computacionales nos remite a la pregunta sobre el valor del experimento virtual y de los mismos instrumentos virtuales que participan en el mismo: cuestiones como la replicabilidad, la unificación en el procesamiento de los datos, la extrapolación cognitiva desde los bloques de datos en bruto hacia los resultados bajo *imaging*, el ajuste de la coma flotante...

<sup>57</sup> Damasio, 1994. Picard (1997) remite el análisis de las emociones en el ámbito de la computación, algo en lo que la investigadora del MIT (Breazeal, 2002) es una experta.

<sup>58</sup> Thagard (1992). Cuando escribí este ensayo estaba por aparecer un libro de Thagard (2006) sobre la “razón emocional”.

<sup>59</sup> Echeverría 2002.

- (3) El nuevo desarrollo de las revistas abiertas y las implicaciones para el control y la validación del conocimiento.
- (4) Los problemas relativos a la conservación y control de la información utilizada, para su posterior evaluación y utilización en caso de conflicto, además de sus implicaciones cognitivas, según Rapp (2003). Pensemos que en muchos casos estamos hablando de bases de datos privadas con acceso previo pago, lo que dificulta la contrastación de la información. Los conflictos entre la investigación pública y privada son muy importantes por lo que respecta al control de la información.
- (5) La evaluación de los procesos computacionales no transparentes que se escapan a la copia simple secuencial del razonamiento humano, de acuerdo con Norman (1997) y Valdés-Pérez (1999).
- (6) La existencia de problemas inherentes a la difícil unificación en las herramientas computacionales, como sucede en el caso de la computación distribuida<sup>60</sup>.

Por esta limitada lista de motivos, que podría ser ampliada con creces, debemos plantearnos la necesidad de una epistemología computacional que nos ofrezca un marco conceptual válido para entender el desarrollo de la ciencia contemporánea. Mi propuesta consiste en indicar en este texto algunos de los puntos débiles observables en estos momentos y en el planteamiento de un exigente trabajo futuro hacia la construcción de un modelo epistémico computacional que nos permita analizar eficientemente y de forma completa la actividad científica.

### Referencias bibliográficas

- Álvarez, A. et al. (2001), "DARWIN: An Evolutionary program for nonlinear modelling of chaotic time series", *Computer Physics Communications*. **136**: 334-339.
- Ashburner, M. et al. (2000), "Gene ontology: tool for the unification of biology. The Gene Ontology Consortium", *Nature Genetics*. **25**: 25-29.
- Baldi, P. & Brunak, S. (2001), *Bioinformatics: the Machine Learning Approach*. Cambridge, MA: MIT Press.
- Bard, J. B. L. (2005), "Anatomics: the intersection of anatomy and bioinformatics", *J. Anat.* **206**:1-16.

<sup>60</sup> Un problema inicialmente serio para la computación distribuida fue el descubrir que un mismo modelo de análisis producía diversos resultados. Pronto se descubrió un problema matemático inesperado: la falta de acuerdo internacional para redondeo de errores en funciones exponenciales, produjo que las pequeñísimas diferencias entre miles de ordenadores personales provocaran desajustes finales importantes. Esto ha puesto en peligro interesantes proyectos como Foldinghome, Einstein@home, Seti@home o muchos otros que suponen la existencia del *altruismo computacional* y la participación social indirecta en la creación de conocimiento científico.

- Bentley, P. J. (2002), *Digital Biology*. New York: Simon & Schuster.
- Blackmore, S. (1999), *The Meme Machine*. Oxford: Oxford University Press.
- Boden, M. A. (ed.) (1990), *The philosophy of artificial intelligence*. Oxford: Oxford University Press.
- Breazeal, C. (2002), *Designing Sociable Robots*. Cambridge, MA: MIT Press.
- Butler, D. (2003), "Scientific Publishing: Who will pay for open access?", *Nature*, **425**: 554-555.
- Bynum, T. W. & Moor, J. H. (1998), *The Digital Phoenix. How Computers Are Changing Philosophy*. Oxford: Blackwell Publishers.
- Castells, M. (2000), *La era de la información*. 3 Vol. Madrid: Alianza.
- Cela, C. J y Marty, G. (1994), "Vida, mente, máquina: medio siglo de metáforas", *Ludus Vitalis*, vol.II, no. 2, pp. 25-38.
- Clark, A. (2003), *Natural-born cyborgs. Minds, technologies, and the future of human intelligence*. Oxford: Oxford University Press.
- Clark, A. & Chalmers, D. (1998), "Embodied, situated and distributed cognition", en Bechtel. & Graham (eds.), *A Companion to Cognitive Science*. Malden (MA): Blackwell Publishers.
- Condit, C. (1999), *The Meaning of the Gene: Public Debates about Human Heredity*. Madison: University of Wisconsin Press.
- Damasio, R. (1994), *Descartes' Error: Emotion, Reason, and the Human Brain*. London: Harper.
- Deemer, Ch. (1994), "What is hypertext?" [Online] [Accedido: 22 de octubre de 2004]. Dirección: <http://www.ibiblio.org/cdeemer/hypertext.htm>.
- Desco, M. y Vaquero, J.M. (2004), "Más de un siglo de imagen médica", *Arbor*, vol. CLXXVII, no. 698, pp. 337-364.
- Echeverría, J. (2002), *Ciencia y Valores*. Barcelona: Ediciones Destino.
- Estany, A. (2000), "Ventajas epistémicas de la cognición socialmente distribuida", *Contrastes*, **6**: pp. 351-375.
- Feigenbaum, E. A. & Feldman, J. (eds.), (1963), *Computers and Thought*. New York: McGraw-Hill.
- Floridi, L. (2002), "What is the philosophy of information?", *Metaphilosophy*, vol. 33, no. 1-2, pp. 123-45.
- \_\_\_\_\_ (2003), "Two approaches to the philosophy of information", *Minds and Machines*, vol. 13, no. 4, pp. 459-69.
- Fogle, Th. (1995), "Information Metaphors and the Human Genome Project", *Perspectives in Biology and Medicine*, vol. 38, no.4, pp. 536.
- Ford, K. M., Glamour, C., Hayes, P.J. (eds.) (1995), *Android Epistemology*. Cambridge, MA: MIT Press.
- Frigg, R. & Hartmann, S. (2006), "Models in Science", *The Stanford Encyclopedia of Philosophy* (Spring 2006 Edition), in S. Sarkar et al. (eds.), *The Philosophy of Science: An Encyclopedia*, Vol. 2. New York: Routledge, pp. 740-749.

- Gainer, A. (2002), "Interactive Labs", *TRENDS in Genetics*, vol. 18, no. 12, pp.651-653.
- Garzon, M. et al. (2003), "Self-Assembly of DNA-like Structures *In Silico*", *Genetic programming and Evolvable Machines*. 4: 185-200.
- Giere, R. (2003), "The role of computation in scientific cognition", *Journal of Experimental & Theoretical Artificial Intelligence*. 15: pp.195-202.
- Gilks, W. (2004), "Bioinformatics: new science-new statistics?", *Significance*, March, pp. 6-9.
- Grim, P. (2004), "Computational modeling as a philosophical methodology", e Floridi (ed.), *Philosophy of computing and information*, Oxford: Blackwell, pp. 337-349.
- Hadley, C. (2004), "Biologists think bigger", *EMBO reports*, vol. 5, no. 3, pp. 236-238.
- Heidorn, B. (2003), "Biological Informatics: A Comparison of Biodiversity Informatics and Neuroinformatics", *Bulletin of the American Society for Information Science and Technology*, Oct/Nov 2003, vol. 30, no. 1.
- Hendler, J. (2003) January 24th, "Science and the Semantic Web", *Science*. 299: 520-521.
- Humphreys, P. (2004), *Extending Ourselves. Computational Science, Empiricism and Scientific Method*. Oxford: Oxford University Press.
- Hunter, Ph. (2005), "Laptop biology", *EMBO reports (European Molecular Biology Organization)*. Vol. 6, no. 3, pp. 208-210.
- Hutchins, E. (1991), "Organizing Work by Adaptation", *Organizational Science*. 2: 14-39.
- \_\_\_\_\_ (1995), *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Hutchins, E. & Norman, D. A. (1988), *Distributed cognition in aviation: a concept paper for NASA* (Contract No. NCC 2-591). Department of Cognitive Science. San Diego: University of California.
- Keller, E. F. (2000), "Models Of and Models For: Theory and Practice In Contemporary Biology", *Philosophy of Science*. 67: S72-S86.
- Kitano, H. (2002), "Systems biology: towards systems-level understanding of biological systems", en Kitano (ed.) *Foundations of Systems Biology*. Cambridge, MA: MIT Press.
- Knight, J. C. & Leveson, N.G. (1986), "An Experimental Evaluation of the Assumption of Independence in Multiversion Programming", *IEEE Trans. Software Eng.* Jan, pp. 96-109.
- Kohl, P. et al. (2000), "Computacional modelling of biological systems: tools and visions", *Philosophical Transactions of the Royal Society London*. A 358: 579-610.
- Langley, P., Simon, H. A., Bradshaw, G.L., Zytkow, J. (1987), *Scientific Discovery*. Cambridge, MA: MIT Press.
- Langley, P. (2000), "The computational support of scientific discovery", *International Journal of Human-Computer Studies*. Vol. 53, no. 3, pp. 393-410.

- Latour, B. (1986), "Visualization and Cognition: Thinking with Eyes and Hands", *Knowledge and Society*. **6**: 1-40.
- Long, M. et al. (2003), "Origin of new genes: evidence from experimental and computational analyses", *Genetika*. **118**: 171-182.
- Malakoff, D. (2004), "Scientific Societies Lay Out 'Free Access' Principles", *Science*. **303**: 1959.
- Marshall, K.E. et al. (2002), "*In silico* discovery of gene-coding variants in murine quantitative trait loci using strain-specific genome sequence databases", *Genome Biology*, vol. 3, no. 12, pp. 1-9.
- Mitcham, C. (1994), *Thinking Through Technology. The Path between Engineering and Philosophy*. Chicago: Chicago University Press.
- Nasar, S. (1999), *A Beautiful Mind: A biography of John. Forbes Nash Jr.* New York: Simon & Schuster.
- Nelson, Ted 1972, "As We Will Think", en James M. Nyce & Paul Kahn (eds.) 1991, *From Memex to Hypertext: Vannevar Bush and the Mind's Machina*, Academic Press, Boston.
- Noble, D. et al. (1999), "Biological simulations in drug discovery", *Drug Discovery Today*. **4**: 10-16.
- Noble, D. & Colatsky, T. J. (2000), "A return to rational drug discovery: computer-based models of cells, organs and systems in drug target identification", *Emerging Therapeutic Targets*. **4**: 39-49.
- Nölting, B. et al. (2004), "Evolutionary Computer Programming of Protein Folding and Structure Predictions", *Journal of Theoretical Biology*. **229**: 13-18.
- Norman, D. A. (1985), *El aprendizaje y la memoria*. Madrid: Alianza.
- \_\_\_\_\_ (1990), *La psicología de los objetos cotidianos*. Madrid: Editorial Nerea.
- \_\_\_\_\_ (1992), *Turn signals are the facial expressions of automobiles*. Reading, MA: Addison-Wesley.
- \_\_\_\_\_ (1993), *Things that make us smart. Deffending human attributes in the age of the machine*. Reading, MA: Addison-Wesley.
- \_\_\_\_\_ (1997), *The Invisible Computer*. Cambridge, MA: .MIT Press.
- \_\_\_\_\_ (2004), *Emotional design. Why we love (or hate) everyday things*. New York: Basic Books.
- Ouzounis, C. A. (2002), "Bioinformatics and the Theoretical Foundations of Molecular Biology", *Bioinformatics*, vol. 18, no. 3, pp. 377-378.
- Ouzonis, C. A. & Valencia, A. (2003), "Early bioinformatics: the birth of a discipline- a personal view", *Bioinformatics*, vol. 19, no. 17, pp. 2176-2190.
- Owens, S. R. (2003), "Revolution or evolution?", *EMBO Reports*. **4**: 742-743.
- Perkins, D. N. (1986), *Knowledge as design*, Hillsdale, N. J: Lawrence Erlbaum Associates.
- \_\_\_\_\_ (1992), *Smart Schools: Better learning and thinking for every child*. New York: Free Press.

- Picard, R. W. (1997), *Affective Computing*. Cambridge, MA: MIT Press.
- Rapp, D. N. et al. (2003), "The impact of digital libraries on cognitive processes: psychological issues of hypermedia", *Computers in human behavior*. **19**: 609-628.
- Richardson, R. C. 1997, "Natural and Artificial Complexity", *Philosophy of Science*. **64**: S255-S267.
- Simon, M. (2005), *Emergent Computation. Emphasizing Bioinformatics*. New York: Springer-Verlag.
- Sloman, A. (1978), *The Computer Revolution in Philosophy*. Cambridge, MA: Harvard University Press.
- Smith, J. M. (2000), "The Concept of Information in Biology", *Philosophy of Science*, vol. 67, no. 2, pp. 177-194.
- Studt, T. (2003), "Bioinformatics-transforming data into models", *R&D Magazine*, vol. 45, no. 8, pp. 20-24.
- Sunshine, C. (2003), *República.com*. Barcelona: Paidós.
- Terry, R. (2005), "Funding the Way to Open Access", *PLOS Biology*, vol. 3, no. 3, pp. 1-6.
- Thagard, P. (1988), *Computational philosophy of science*. Cambridge, MA: MIT Press.
- \_\_\_\_\_ (1992), *Conceptual revolutions*. Princeton, NJ: Princeton University Press.
- \_\_\_\_\_ (2006), *Hot Thought: Mechanisms and Applications of Emotional Cognition*. Cambridge, MA: MIT Press.
- Tichy, W.F. (1998), "Should Computer Scientists Experiment More?", *Computer*, May, pp. 32-40.
- Valdés-Pérez, R. E. (1999), "Principles of human-computer collaboration for knowledge discovery in science", *Artificial Intelligence*. **107**: 335-246.
- Vallverdú, J. (2004), "Frequentistes versus bayesians: uns apunts des del sentit comú", *Comprendre*, vol. VI, no. 2004/1, pp. 79-94.
- Van Dijk, J. (1998), *Imagination: Popular Images of Genetics*. London: Macmillan.
- Varenne, F. (2001), "What does a computer simulation prove ?", *Simulation in Industry*, Proc. of The 13th European Simulation Symposium, Marseille, France, October 18-20th, 2001, ed. by N. Giambiasi and C. Frydamn, SCS Europe Bvba, Ghent, pp. 549-554.
- Vaughan, D. (ed.), (1996), *To Know Ourselves*. Washington: U.S. Department of Energy & HGP.
- Wickware, P. (2000), "Next-generation biologists must straddle computation and biology", *Nature*. **404**: 683-684.
- Winsberg, Eric (2003), "Simulated Experiments: Methodology for a Virtual World", *Philosophy of Science*. **70**: 105-125.
- Wolf, R.S. y Yaeger, L. (1993), *Visualization of Natural Phaenomena*. New York: Springer-Verlag.