



Colección
**Nuevas Tecnologías
y Sociedad**

Acerca de Internet

Hubert L. Dreyfus

Índice general

Agradecimientos	13
Introducción	15
I. El mito de los hipervínculos	23
II. Entre el aprendizaje y la educación a distancia	41
Primera etapa: principiante	47
Segunda etapa: intermedio	48
Tercera etapa: intermedio-avanzado	49
III. La telepresencia y la distancia de lo real	63
IV. Nihilismo, anonimato y compromiso en la época actual	85
Conclusiones	103

Introducción

“Yo te muestro el Superhombre, pues el hombre es algo a superar. ¿Y qué has hecho tú para superarle?”

Friedrich Nietzsche. *Así Habló Zarathustra* (Prólogo).

“¿Por qué pretender ser post-humano? (...) Es cierto que podemos lograr mucho siendo humanos, pero cúspides mucho más altas pueden ser alcanzadas sólo aplicando la inteligencia, determinación y optimismo necesarios para trascender el crisol de lo humano... nuestros cuerpos... restringen nuestras capacidades.”

Cita de Nietzsche y respuesta de Max More, fundador del Instituto Extropiano.¹
<http://www.ct.heise.de/tp/english/inhalt/co/2041/1.html>.

Internet no es sólo una innovación tecnológica; es un nuevo *tipo* de innovación que saca a relucir la verdadera esencia de la tecnología. Hasta ahora, los precursores de este campo han producido dispositivos o aparatos que satisfacen necesidades reconocidas previamente, logrando algunos efectos inesperados. En su momento, Alexander Graham Bell pensó que el teléfono sería útil para la comunicación en el mundo empresarial, pero que no tendría éxito en los hogares, y ni hablar de usarlos mientras se camina por la calle. Igualmente, Henry Ford, que imaginó los automóviles como un medio seguro, económico e individual de transporte, jamás imaginó que llegarían a destruir el centro de las ciudades o participar en la liberación sexual en adolescentes.

1. Aunque los extropianos van demasiado lejos, algunas de sus ideas están bien articuladas en libros de pretensiones serias tales como el de Hans Moravec (1988). *Mind Children*. Cambridge, MA: Harvard University Press.

Internet es diferente. Originalmente fue concebida para la comunicación entre científicos, lo cual hoy por hoy es en sí un efecto secundario. Hoy nos encontramos con que Internet es una masa gigantesca y demasiado dinámica como para imaginar que satisfaga *alguna* necesidad en especial. Cada nuevo uso que se le confiere nos depara sorpresas. Si la esencia de la tecnología es hacer *todo* accesible y mejorable, entonces Internet es el dispositivo tecnológico perfecto: la culminación de aquella tendencia a flexibilizarlo todo hasta el límite, materializada en la digitalización e interconexión de toda realidad posible.² Lo que nos permite la Red es literalmente ilimitado, y eso naturalmente ha llevado a algunos a luchar en favor de cambios antes impensables. Se nos dice, entre otras cosas, que por sus modos de conectar y proveer acceso a información, Internet traerá una nueva era de prosperidad económica, que nos llevará al desarrollo de buscadores inteligentes para dar justo con lo que necesitamos, resolver los problemas de educación de masas, conectarnos con toda la realidad, permitirnos una identidad aún más flexible de la que ya podamos tener y, de este modo, dotar de nuevos sentidos a nuestras vidas.

Parecería que estamos frente a un caso de auge y consiguiente caída de los *punto com*, donde ha habido incontables áreas con proyectos *web* que han prometido nuevas formas de vida, mucho más satisfactorias. Como iniciativas dieron mucho de qué hablar, pero pocos resultados sostenibles en el tiempo.³ De hecho, los investigadores de la Universidad de Carnegie-Mellon quedaron sorprendidos al descubrir el porcentaje de personas que, con acceso a la World Wide Web, se sentían desoladas y deprimidas. En el *Times* de Nueva York se recoge que:

“Los resultados del proyecto del millón y medio de dólares fueron totalmente contrarios a las expectativas de los científicos sociales que lo diseñaron, así como

2. Una tendencia que Martin Heidegger da por definitiva para nuestra comprensión moderna de lo que es ser parte del todo. Ver Martin Heidegger (1994). “La pregunta por la técnica” (Cap. 1). *Conferencias y artículos*. Barcelona: Ediciones del Serbal.

3. No significa esto que no se hayan producido innovaciones reales. Los nuevos modos de vincular información han afectado a las bibliotecas; los cursos de páginas web para estudiantes universitarios les han permitido atender a lecciones y sostener discusiones sin salir de sus habitaciones; la telerrobótica ha hecho posible el control de un vehículo en Marte y algún día millones de televidentes, sin duda, podrán ver este vehículo desplazarse sobre alguna superficie lejana. El correo electrónico ha abierto posibilidades sorprendentes, desde disidentes políticos echando mano sobre reformas, hasta abuelos orgullosos que envían las últimas fotografías digitales de sus nietos a todos sus amigos. El punto aquí es que todos estos desarrollos, de por sí extraordinarios, son pequeños comparados con las predicciones.

a las organizaciones que lo patrocinaron y financiaron... Robert Kraut, profesor de Psicología social del Instituto de Interacción Humana Computarizada de la Universidad de Carnegie-Mellon declaró estar 'estupefacto con los resultados, porque son contra-intuitivos con respecto a lo que sabemos sobre el uso social de Internet... No se trata de casos extremos: hemos estudiado a adultos normales y sus familias y, en promedio, para aquellos que más utilizaban Internet, la vida empeoró.'"⁴

A. Harmon (30 de agosto de 1998). "Researchers Find Sad, Lonely World in Cyberspace". *The New York Times*.

Los investigadores resumen sus conclusiones como sigue:

"Esta investigación examina el impacto social y psicológico de Internet en 169 personas, pertenecientes a 73 hogares, durante sus primeros dos años de contacto con Internet... En esta muestra, la Red era utilizada fundamentalmente para comunicaciones y, sin embargo, su uso prolongado suscita cierta pérdida de comunicación con otros miembros del hogar, reducción del círculo social e incremento en depresión y sensación de soledad."

R. Kraut, M. Patterson, V. Lundmark, S. Kiesler, T. Mukophadhyay y W. Scherlis (1998). "Internet Paradox: A Social Technology that Reduces Social Involvement and Psychological Well-being?". *American Psychologist* (vol. 53, núm. 9, pág. 1017-1031).

La conclusión de los autores es que hay una pérdida de presencia corporal entre las personas:

"Las amistades *on-line* tienden a ser más limitadas que las que cuentan con cierta cercanía física... Puesto que los amigos en Internet no conocen un mismo entorno cotidiano, tendrán menos posibilidad de crear un contexto para las conversaciones, haciendo las discusiones más difíciles y generando un apoyo menos aplicable. Incluso los lazos fuertemente establecidos, cuando pasan a sostenerse en la distancia gracias a la comunicación electrónica, son propensos a cambiar de forma y quizás disminuir en intensidad en comparación con las relaciones que se mantienen con contacto físico. Las aplicaciones de comunicación interpersonal hoy pre-

4. En ese mismo artículo, Harmon escribe que: "Aquellos participantes que se sentían más solos y deprimidos al inicio del estudio de dos años –tal como se determinó a través de un cuestionario estándar administrado a toda la muestra– no demostraron ser más propensos al uso de Internet. Más bien, según los investigadores, el uso de la Red parecía causar una caída del bienestar psicológico."

valecientes en Internet son neutrales con respecto a las relaciones sólidas, o incluso tienden más a debilitarlas que a promoverlas.”⁵

R. Kraut, M. Patterson, V. Lundmark, S. Kiesler, T. Mukophadhyay y W. Scherlis (1998). “Internet Paradox: A Social Technology that Reduces Social Involvement and Psychological Well-being?”. *American Psychologist* (vol. 53, núm. 9, pág. 1017-1031).

Estos hallazgos tan sorprendentes demuestran que la *incorporeidad* que experimenta el usuario de Internet tiene efectos tan profundos como inesperados. Podemos presumir que afecta a las personas de modo distinto con respecto a la mayoría de instrumentos comunicacionales porque puede llegar a ser el mecanismo principal de relación entre esos mismos usuarios y el resto del mundo. En vista de estas sorpresas y decepciones, queremos saber cuáles son los beneficios y la amenazas de vivir *on-line*. Sólo así podremos echar luz sobre nuestras preocupaciones acerca de qué implica asumir el reto de vivir en Internet.

Los más entusiastas del mundo virtual aducen que la promesa de largo alcance de Internet está en que cada uno de nosotros será capaz de trascender los límites de su cuerpo. En palabras de John Perry Barlow, uno de los más prominentes exponentes de estas posturas, la frontera electrónica es “un mundo a la vez omnipresente y ausente, en el que no viven los cuerpos”.⁶ Cuando se refieren al cuerpo, estos visionarios no sólo circunscriben una dimensión física con cabeza, tronco, brazos, piernas y la habilidad para moverse en el mundo, sino también los estados de ánimo que permiten nuestro interés por las cosas, nuestra ubicación en un contexto particular donde debemos lidiar con objetos y personas, así como las distintas maneras a través de las cuales nos exponemos a decepciones y fracasos, a los accidentes o a la muerte. Durante el resto del libro interpretaremos esta corporeidad como un término amplio que incluye todos los aspectos de nuestra finitud y vulnerabilidad.

5. Pareciera que la ausencia física puede llevar también a una suerte de soledad moral. Cuando Larry Froistad confesó ante su grupo de apoyo por correo electrónico que había asesinado a su hija, los miembros del grupo le ofrecieron sostén emocional; sólo uno de ellos pensó que era necesario denunciarlo a la policía. Ver Amy Harmon (30 de abril de 1998). “On-line Thoughts and Off-line Killing”. *The New York Times*: “Para la Sra. De Carlo, la naturaleza de la comunicación *on-line* –que produce una distancia tanto física como psicológica entre los participantes– hace a sus compañeros olvidar responsabilidades *off-line* tales como llevar un asesinato confeso ante la justicia.”

6. J.P. Barlow (8 de febrero de 1996). “A Declaration of the Independence of the Cyberspace”. Davos, Suiza (http://www.eff.org/publications/John_Perry_Barlow/barlow_026.declaration).

Yeats lamentaba que su alma estuviese “atada a un animal en proceso de desaparición” (“Sailing to Byzantium”, en *The Toser*, 1928). Y es fácil ver con cierta fascinación la idea de una evolución humana que deje atrás al cuerpo animal que mantiene atrapada la identidad lingüística y cultural. ¿Quién no querría convertirse en una masa incorpórea capaz de estar en cualquier parte del universo y hacer copias de seguridad de sí mismo para evitar el sufrimiento o evadir la muerte? Los visionarios de la virtualidad y muchos otros estarían felices de liberarse de las deformidades, la depresión, las enfermedades, la edad o la muerte, una promesa que no han escatimado los futuristas inspirados en los ordenadores como Hans Moravec⁷ y Ray Kurzweil,⁸ y que se encuentra tipificada en la Red (¿dónde más?) por grupos internacionales como los extropianos, cuyo líder, Max More, ha sido citado en el epígrafe al inicio de este texto. Incluso otros sabios más cercanos a la vida cotidiana suscriben la idea de que estamos entrando en un nuevo nivel de civilización. De acuerdo con el consultor empresarial Esther Dyson “el Ciberespacio es la tierra del conocimiento. Su explotación constituye el más puro y eminente motivo de la humanidad”.⁹

Dejar atrás la corporeidad habría satisfecho también a Platón, quien veía en el cuerpo la tumba del alma¹⁰ y, siguiendo a Sócrates, aceptaba que la tarea más alta del ser humano era “morir en cuerpo” y ser pura mente. En palabras del mismo Sócrates: “al despreciar el cuerpo y evitarlo, buscando la independencia, el filósofo supera a todos los demás”¹¹.

Resulta curioso que los extropianos digan seguir a Nietzsche y no a Platón cuando afirman que debemos trascender nuestra humanidad. De hecho, la postura antiplatónica de Nietzsche con respecto al cuerpo aparece en el mismo libro que los extropianos tanto se complacen en citar. En la sección llamada “De los que desprecian el cuerpo”, Nietzsche pone en boca de Zarathustra –y casi en respuesta a los extropianos– un mensaje claro: “No sigo vuestros pasos, despreciadores del cuerpo, pues considero

7. Moravec (ob. cit.)

8. R. Kurzweil (2000). *The Age of Spiritual Machines*. Nueva York: Penguin.

9. E. Dyson y otros (1994). “Cyberspace and the American Dream: A Magna Carta for the Knowledge Age. Release 1.2”. Washington DC: The Peace and Progress Foundation.

10. Platón. *Górgias* (492e7-493a5). Dice Sócrates: “Escuché una vez a uno de nuestros sabios decir que estamos muertos, y que nuestro cuerpo (soma) es una tumba (sema).”

11. Plato (1954). *Phaedo. The Last Days of Socrates* (pág. 84). Baltimore, MD: Penguin. En castellano, Platón. *Fedón. Apología de Sócrates*.

que no sois puentes tendidos hacia el Superhombre”¹², después de haber dicho:

“¡Con qué orgullo dices: ‘YO’! Pues, aunque no lo quieras creer, tu cuerpo y su gran razón...no dicen ‘yo’ sino que constituyen el ‘YO’... Detrás de tus pensamientos y de tus sentimientos, hermano, hay un amo poderoso, un sabio desconocido que se llama sí mismo. Habita en tu cuerpo. Es tu cuerpo.”

Friedrich Nietzsche (2000). *Así Habló Zaratustra* (pág. 34). Madrid: Edimat libros (trad. Francisco J. Carretero).

Nietzsche pensaba que lo más importante para los seres humanos no eran sus capacidades intelectuales, sino las posibilidades emocionales e intuitivas del cuerpo. En su implacable cuestionamiento del platonismo y el cristianismo, e incluso ante las formas más abstrusas de la ciencia y la tecnología, este autor apostaba por la trascendencia de nuestras limitaciones humanas hacia el superhombre. Con ello quería fomentar la idea de que los seres humanos, en vez de continuar negando su muerte y finitud, concibieran una fuerza afirmativa de sus cuerpos y su mortalidad.

En este sentido, las cuestiones que aquí nos interesan son las siguientes: ¿Podemos seguir adelante sin nuestros cuerpos? ¿Es el cuerpo un simple residuo en nuestra evolución como animales –una limitación de libertades que hoy por hoy podemos superar, como sostienen los extropianos– o, por el contrario, juega un rol fundamental en nuestra vida espiritual e intelectual, como apunta Nietzsche? De llevar razón este último, las supuestas virtudes de Internet como fuente de liberación ante las limitaciones corporales constituirían, contrariamente, su talón de Aquiles.

Como filósofos no nos hemos propuesto condenar algunos usos de Internet y celebrar otros. La pregunta aquí es aún más especulativa: ¿Qué sucedería si Internet pasara a ocupar un lugar central en nuestras vidas: si se convirtiera en lo que Joseph Nye, decano de la Escuela de Gobierno Kennedy de la Universidad de Harvard, llama “una irresistible cultura alternativa”? Y también, en la medida que vamos integrando gran parte de nuestras vidas al ciberespacio, ¿nos haremos suprahumanos o infrahumanos?

12. Friedrich Nietzsche (2000). *Así Habló Zaratustra* (pág. 55). Madrid: Edimat libros (trad. Francisco J. Carretero).

Al elaborar sobre estas cuestiones también debemos sentirnos invitados a tener la apertura necesaria para aceptar la posibilidad de ganar libertades nunca antes conseguidas como seres humanos. Entrar en el ciberespacio definitivamente puede implicar dejar atrás nuestra corporalidad con sus formas animales, emociones, intuiciones, situaciones y vulnerabilidades, pero ello no excluye la posibilidad de perder, necesariamente, algunas de nuestras capacidades fundamentales: nuestra habilidad para dotar de sentido a las cosas, distinguir lo relevante de lo irrelevante, dar importancia al éxito o el fracaso como condiciones necesarias para el aprendizaje, y nuestra necesidad de máximo aferramiento al mundo con el que diseñamos nuestro sentido de realidad de las cosas. De hecho, esperamos analizar en qué medida suprimir el cuerpo conlleva suprimir la capacidad para conceder relevancia, desarrollar habilidades, percibir la realidad y dotarla de sentido. Si ese es el canje, el futuro de vivir en la Red, después de todo, puede no resultar tan atractivo.

Sumario

Capítulo 1. Las limitaciones de los hipervínculos. La aspiración a lograr una gestión inteligente de la información y el fracaso de la inteligencia artificial (IA). Se intenta revelar cómo la forma y el movimiento del cuerpo juegan un rol crucial en la construcción del sentido de mundo, de manera que perder la corporeidad conllevaría la *pérdida de la habilidad para adjudicar relevancia*.

Capítulo 2. El sueño de la educación a distancia. La importancia de lo relevante en el proceso de enseñanza/aprendizaje. Aprendizaje y necesidad de imitación. Sin implicación corporal y presencial *no podemos adquirir destrezas*.

Capítulo 3. La ausencia de la telepresencia. El cuerpo como fuente de sentido de contacto con la realidad. Se revela que la ausencia de parámetros para establecer contextos coordinados, propia de la telepresencia, produce la *pérdida de sentido de la realidad de las personas y las cosas*.

Capítulo 4. Anonimato vs. nihilismo. El sentido de la vida implica un compromiso genuino que, a su vez, requiere asumir riesgos reales. El anonimato y la seguridad que conceden los compromisos *on-line* pueden llevar a una *vida sin sentido*.

Capítulo I

El mito

de los hipervínculos

“El problema de la inteligencia artificial, como se la denomina –la producción de máquinas que se comporten de modo suficientemente similar a los humanos cuando actúan inteligentemente–, no ha sido aún resuelto. Más aún, no hay nada en el horizonte que nos haga pensar que hay algo de luz al final del camino. Los mundos como el de la ‘Inteligencia Artificial’, de los ‘agentes inteligentes’, los ‘sirvientes’ –exagerados en la prensa– son replanteamientos del caos y del problema en el que estamos envueltos.

Nos encantaría tener una máquina que pueda navegar por la Red para encontrar algo específico, o construir nuestros almacenes personales, con información sobre nuestras preferencias y de lo que queremos decir cuando buscamos algo. Pero, sencillamente, no hay nada a ese nivel.”

Michael Dertouzos (laboratorio de Ciencias Informáticas, MIT) (7 de julio del 2000). “The Future of Computing”. *Talk to the Nation, Science Friday*. National Public Radio.

Internet crece, extensa y exuberantemente. En un sondeo reciente se detectaron más de un billón de páginas, y ha seguido expandiéndose a razón de, al menos, un millón de páginas por día.¹ Es incluso característico de la Red que, para el momento en que se lean estas líneas, tales cifras estén ya desactualizadas. Existen cantidades tan astronómicas de información relevante en el espacio virtual, que se hace cada vez más difícil

1. S. Lawrence y C.L Giles, Instituto de Investigaciones NEC (3 de abril de 1998). “Searching the World Wide Web”. *Science* (núm. 280, pág. 98). Más aún, el tamaño no se constituye por la cantidad de páginas o sitios web, pues el número de vínculos en estos espacios es aún mayor.

encontrar un contenido en particular; un problema que surge de la manera en que toda esa información es organizada (más bien, desorganizada) en la Red. Esta maraña de información se “maneja” utilizando hipervínculos que pueden enlazar cualquiera de sus infinitos contenidos entre sí indiscriminadamente o, lo que es lo mismo, por la razón que al diseñador se le ocurra en el momento que ha de establecer la relación. No existe autoridad ni sistemas de catálogos concertados que orienten o circunscriban las asociaciones entre los enlaces.²

Los hipervínculos no han sido introducidos porque sean más útiles para buscar información que los sistemas antiguos de ordenamiento jerárquico. Más bien han sido un dispositivo ordinario de aprovechamiento de la velocidad de los ordenadores para relacionar grandes cantidades de información sin requerir comprenderla o someterla a algún tipo de estructura, ya sea impuesta o generalmente aceptada. No obstante, cuando todo puede vincularse indiscriminadamente y sin obedecer a algún propósito o significado concreto, el tamaño de la Red y la arbitrariedad entre sus enlaces hacen extremadamente difícil para un usuario encontrar exactamente el tipo de información que busca.

La cuestión de la búsqueda de información relevante dentro de un cuerpo de hipervínculos interconectados forma parte de las innovaciones de Internet. La manera tradicional de organizar información dependía de alguien –un zoólogo, un bibliotecario o un filósofo– que produjera un esquema de clasificación de acuerdo a los significados de los términos envueltos en la organización, así como a los intereses de los usuarios.³ Las personas podían, entonces, buscar dentro de este esquema en correspondencia con lo que entendían por las diferentes categorías y la información presente en el sistema. En este sentido, para utilizar la información había que depender de lo descrito y empleado con anterioridad, organizaciones

2. Ha habido litigios interesantes para regular la libertad de vincular cualquier información a voluntad, donde la parte demandante ha reclamado al demandado por el establecimiento de vínculos con su página web. Obviamente, esto es apenas una fracción minúscula y es improbable que tenga efectos significativos sobre las formas de conducción de Internet, pero sin duda refleja de modo simple lo que queda de la antigua y decadente práctica de pretender establecer al menos *algunos* límites a la eventual vinculación de cualquier información con cualquier otra.

3. El sistema decimal Dewey se organizó de esta manera; ni siquiera permitía que un mismo ítem pudiese ser catalogado bajo dos categorías. Si bien hoy en día los bibliotecólogos tienen mayor libertad para archivar la misma información bajo diferentes rubros, como en el caso, por ejemplo, de la Filosofía de la Religión que podría encontrarse presumiblemente tanto en Filosofía como en Religión, se puede ver una taxonomía jerárquica relativamente acordada.

de bases comunes de significados para facilitar a los usuarios una manera sistemática de dar con lo que considerasen relevante para sus propósitos.

David Blair, profesor de Sistemas de computación e información de la Universidad de Michigan,⁴ sostiene que los esquemas de clasificación más “tradicionales” estaban explícita o implícitamente relacionados con cierto tipo de prácticas. Las ciencias serían el ejemplo obvio en este sentido, pero otros tipos de prácticas menos formales subyacen a este tipo de ordenamientos, tales como las prácticas agrícolas, donde el campesino debe ser capaz de identificar diferentes tipos de plantas, animales, pestes, enfermedades, condiciones climatológicas, estaciones, etc. Mientras algunos vínculos en Internet pueden asociar sitios ajustados a determinadas prácticas comunes, la mayoría no reúnen ninguna práctica en particular. Y sin la presencia de alguna práctica que oriente qué ha de vincularse con qué, la proliferación de enlaces puede crecer infinitamente.⁵

Desde los tiempos de Aristóteles, nos hemos acostumbrado a organizar la información jerárquicamente en categorías cada vez más amplias, donde las superiores contienen a las inferiores. En este sentido, por ejemplo, pasamos de cosas, a seres vivientes, a animales, a mamíferos, a perros, a *collies*, y a Lassie. Cuando la información se organiza en una base de datos jerárquica, el usuario puede seguir los enlaces relevantes, precisamente porque se adscribe a cierto tema o cierta clase de información, previamente organizada, dentro del cual intenta alcanzar un nivel de especificidad determinado. El usuario tiene que tener, por ejemplo, cierto interés en los animales antes de saber lo que quiere averiguar sobre las tortugas, y una vez que se ha adentrado en la categoría de animales, no puede examinar los datos sobre problemas de la física cuántica sin abandonar o retroceder en su interés principal.

La información organizada en hipervínculos en Internet, por el contrario, no cuenta con criterios de organización de categorías, clases o miem-

4. El libro de David Blair titulado *Language and Representation in Information Retrieval* (Nueva York, Elsevier Science, 1990) fue escogido como Mejor Libro de 1991 por la Sociedad Americana de Ciencias de la Información. Y su autor también fue reconocido como “Investigador más destacado del año” en 1999, por la misma institución.

5. Blair sugiere que “al creer que organizamos cierta información para un uso o práctica específica, olvidamos que tal uso o práctica también ayuda a organizar la información. Estos casos, que hemos denominado ‘arreglos naturales’ de la información, implican que la selección y organización de la información a disposición de una práctica es un proceso interactivo. Y lo que la Red evita o desalienta es este tipo de relación natural entre información y práctica”.

bros. El criterio parece ser la libre interconexión entre todos los elementos, sin jerarquías y en un mismo nivel. Los hipervínculos nos permiten enlazar directamente una entrada de datos con otra siempre que estén relacionadas de algún modo, así sea remoto. Así, tenemos la impresión de estar apenas a unos pocos pasos de cualquier página dentro de la inmensidad de Internet. En las bases de datos de hipervínculos, los usuarios han de navegar por una red muy vasta de información en la que todo es igualmente accesible y nada es privilegiado: entre los sitios que se nos ofrecen en el navegador con información sobre tortugas, podríamos acceder a una página llamada “Tortugas comparadas con liebres”, y caer instantáneamente en la paradoja de Zenón, el filósofo de la Grecia antigua.

Se puede ilustrar la oposición entre los anteriores y los nuevos métodos de organización y búsqueda de información, las ventajas de cada uno, si comparamos la antigua cultura de las librerías con lo que hoy posibilitan los hipervínculos. Las oposiciones demuestran la transformación de un modo semántico de estructuración de la información orientado por el significado, en otro basado en la sintaxis formal donde el significado pierde toda relevancia. La tabla 1.1 muestra la sistematización de algunas de las oposiciones.

Tabla 1.1

Comparación entre viejos y nuevos sistemas de búsqueda de información	
Cultura anterior de librerías	Cultura del hipervínculo
Clasificación a. Estable b. Organizado jerárquicamente c. Definido por intereses concretos	Diversificación a. Flexible b. De un solo nivel c. Permite todas las asociaciones posibles
Selección meticulosa a. Calidad de las ediciones b. Autenticidad de los textos c. Eliminación de materiales antiguos	Acceso a todo a. Inclusión entre ediciones b. Disponibilidad de textos c. Todo se guarda
Colecciones permanentes a. Preservación de textos determinados b. Navegación según intereses concretos	Colecciones dinámicas a. Evolución intertextual b. Navegación lúdica

Ciertamente, el usuario de una librería virtual no será más un sujeto moderno con identidad fija y deseos de encontrar un modelo de mundo

más completo y confiable⁶ sino, más bien, un ser postmoderno y multi-forme abierto a explorar nuevos horizontes. En cualquier caso, este no está interesado en congregar datos sobre aquello que le resulte *significativo*, sino *en conectarse con tanta información en red como le sea posible*.

Los nuevos navegantes de Internet celebran esta proliferación de información como una contribución a la nueva forma de vida que coloca la sorpresa o la fascinación por encima del significado y la utilidad. Especialmente aquellos que se jactan de rechazar cualquier forma de jerarquía o autoridad sin preocuparse por resolver el problema práctico de encontrar información relevante, encuentran cierto atractivo en esta idea: los teóricos y artistas posmodernos ven en los hipervínculos un modo de liberarse de los especialistas que, desde el anonimato, organizan las bases de datos y deciden por los demás lo prescindible de lo imprescindible. Aparte, el hecho que la *cantidad* de conexiones sea valorada por encima de cualquier juicio sobre su calidad, apoya ciertos referentes de la idea estadounidense de democracia, como bien ha dicho Fareed Zacaria, editor del Foreign Affairs, al afirmar que “Internet es profundamente irrespetuosa con las tradiciones, el orden establecido y la jerarquía, lo cual es tremendamente norteamericano.”⁷

Quienes requieren utilizar los datos disponibles, no obstante, tratan de encontrar información significativa y relevante de acuerdo a sus necesidades, y el hecho que los hipervínculos conformen bases de datos asociando cualquier información sin criterios claros o estables dificulta sustancialmente esa tarea. Dado que los hipervínculos se crean por incontables razones o propósitos, y considerando que se trata de una sola clase de enlace, el usuario no puede introducir los significados relevantes para arribar a la información buscada. En otras palabras, el problema está en que, desde el punto de vista del significado, todos tienen el mismo valor. En palabras de un investigador, buscar información en Internet ni siquiera se puede comparar con buscar una aguja en un pajar, sino con buscar una aguja determinada

6. El concepto actual de “sujeto moderno” viene de principios del siglo XVII –gracias a Luther, la imprenta, y las nuevas ciencias– cuando las personas comenzaron a considerarse individuos autosuficientes. Descartes introdujo la idea de que el sujeto es lo que subyace bajo los estados mentales en permanente transformación, y para Kant, el sujeto en tanto secularizador de todo, tenía que ser libre y autónomo. Como veremos en el capítulo 4, Soren Kierkegaard concluyó que cada persona es un sujeto llamado a tomar parte dentro de una determinada identidad que definirá quién es y qué es significativo. (ver capítulo 4).

7. *The New York Times* (9 de enero del 2000).

en un gran saco de agujas. En vista de la ausencia de cualquier contenido semántico en la determinación de los enlaces, cualquier pretensión de buscar en Internet termina convirtiéndose en una técnica sintáctica y formal de manipulación de símbolos sin sentido, que trata de localizar contenidos semánticos relevantes y significativos.

Nótese que las dificultades del uso de operaciones mecánicas carentes de significado en la búsqueda de información relevante no son exclusivas ni llegaron con Internet, sino que se encuentran cada vez que una persona intenta buscar información de acuerdo a cierto interés en una base de datos que no está organizada para satisfacer ese interés. En un caso típico, un grupo de investigadores puede buscar los trabajos publicados sobre un tópico concreto y encontrar que las palabras en los títulos de los artículos no permiten a la máquina de búsqueda localizar la información precisa que requieren.

Para comprender esta dificultad, es importante distinguir entre búsqueda de datos y de información. David Blair ha explicado la diferencia como sigue:

“Los Sistemas de Gestión de Bases de Datos han revolucionado el tratamiento y la búsqueda de información –podemos llamar a un directorio de asistencia y averiguar prácticamente cualquier número de teléfono de cualquier lugar en Estados Unidos o Canadá, acercarnos a un cajero automático en una ciudad lejana a nuestro hogar y extraer efectivo de nuestra cuenta, ir a una oficina de venta de boletos en Michigan y adquirir un asiento numerado para una obra de teatro, la siguiente semana, en San Francisco, etc.–, gracias, en parte, a estos sistemas desarrollados en los últimos 35 años de manera extensiva y confiable.

La búsqueda de datos funciona sobre entidades como “nombres”, “direcciones”, “números de teléfono”, “balances de cuentas bancarias”, “números de seguridad social”, ítems que típicamente representan referencias claras e inequívocas. Pero si bien muchos documentos se pueden representar con referencias o significados concretos –como por ejemplo, autor y/o título–, muchas de las búsquedas de información se guían más por el *contenido intelectual* de los documentos que se espera encontrar (por ejemplo, cuando se buscan los estudios económicos que analizan las perspectivas de inversión en Europa Central para industrias de servicio). Las descripciones de *contenido intelectual* casi nunca están determinadas, y en un campo tan vasto como el World Wide Web (WWW) se hacen irremediablemente imprecisas y vagas para cualquier cosa que no sea una búsqueda demasiado general.

Buscar un URL conocido en Internet puede ser simple y muy fácil porque tiene la precisión y dirección que requiere la búsqueda de datos. Pero buscar una página web con contenidos intelectuales determinados a través de operadores de búsquedas puede ser muy difícil y, en ocasiones, imposible”

Prueba del próximo libro de David Blair, *Wittgenstein, Language and Information*.

La diferencia entre búsqueda de datos y de documentos puede resumirse en la siguiente tabla:

Tabla 1.2

Diferencias entre la búsqueda de datos y de documentos.	
Búsqueda de datos	Búsqueda de documentos
1. Directa (“quiero saber X”)	1. Indirecta (“quiero saber acerca de X”)
2. Necesaria relación entre lo que se busca y lo que se encuentra efectivamente.	2. Relación probabilística entre lo buscado y los documentos efectivamente encontrados.
3. Criterio de éxito: adecuación	3. Criterio de éxito: utilidad
4. Escalar categorías ascendenteamente no representa mayor complicación.	4. Escalar categorías ascendenteamente representa una grave complicación.

Antes de la llegada de Internet y sus máquinas de búsqueda, la solución típica al problema consistía en emplear seres humanos –es decir, especialistas que entendían la documentación– que colaboraran en la descripción de los contenidos de manera que, eventualmente, los usuarios pudiesen localizar la información requerida. El problema hoy es que resulta materialmente imposible tener suficientes catalogadores para generar índices de Internet –tanto por su tamaño como por su velocidad de crecimiento.

Para comprender la magnitud de este problema de acceso, podemos considerar una analogía planteada por Blair:

“Supongamos que queremos encontrar un libro entre los muchos que hay a disposición. Sería como dar con un individuo entre muchos en un espacio reducido, lo cual no implica un problema demasiado complicado, incluso si asumíramos que nuestras descripciones del libro o la persona son relativamente generales. Supongamos, ahora, que tenemos que encontrar un libro en una biblioteca de 50.000 títulos. Si bien cualquiera ha estado en una biblioteca de estas dimensiones, es difícil imaginar cuán complicada puede ser esta tarea. Para ello, considérese la situación

de muchos campos profesionales de béisbol en Estados Unidos, con capacidad para unas 50.000 personas, y se comprenderá cómo encontrar a alguien en un juego cuyas entradas han sido totalmente vendidas, en un lugar como, por ejemplo, Fenway Park, puede ser verdaderamente complicado. Ahora bien, nuestra tarea es aún mucho más formidable: incluso sabiendo que nuestras guías de búsqueda de un individuo en particular sean generales como por ejemplo edad media, de cabello oscuro, ojos oscuros, de cerca de 1,75 metros de altura y delgado, debemos imaginar que esto sería equivalente a buscar un libro en una biblioteca moderadamente grande, digamos, de unos 100.000 ejemplares. Algo así como buscar a un individuo en un concierto de los Rolling Stones en el Central Park de Nueva York. E incluso tras este último ejemplo, queda mucho para comprender las proporciones de la magnitud de una búsqueda en Internet: buscar a través de millones de fuentes de contenido actualmente a disposición en el universo virtual, empleando sólo las herramientas de búsqueda disponibles, equivale a buscar a alguien particular en la ciudad de Nueva York, con los datos generales que tenemos: cabello y ojos oscuros, edad media y delgado.”

D. Blair (1994). “Hill it scale up? Thoughts and Intellectual Access in the Electronic Networks”. En: Okerson (ed.). *Gateways, Gatekeepers, and Roles in the Information Omniverse*. Washington, DC: Association of Research Libraries: Office of Scientific and Academic Publishing.

Frente a tan enorme problema, los investigadores que trabajan en sistemas de búsqueda de información han optado por la inteligencia artificial (IA) y, desde los sesenta, han estado trabajando para lograr que los ordenadores –que son máquinas sintácticas sensibles únicamente a modos de ingreso de información– se comporten como seres humanos –que son sensibles a la semántica y a los significados. Así visto, es natural que se haya recurrido a la inteligencia artificial como la base en la generación de programas que encuentren información tras reconocer criterios de relevancia del autor humano de la búsqueda. Al principio, los programadores de la inteligencia artificial cayeron en el optimismo de pensar que, con unos pocos millones de datos, lograrían representar todos los hechos acerca del mundo que daba significado a las personas. Pero al final de los setenta y principio de los ochenta se toparon con que, para producir inteligencia artificial, tendrían que sistematizar explícitamente el sentido común compartido por las personas. Una tarea hercúlea ingente, cuando no materialmente imposible.⁸

8. Ver H. Dreyfus (1992). *What Computers (Still) Can't Do*. Cambridge, MA: MIT Press, 3^a ed.

El más destacado promotor de esta iniciativa es Douglas Lenat,⁹ quien comienza por asumir que el conocimiento que compone el sentido común no se corresponde con el tipo de saber presente en las encyclopedias, sino, muy por el contrario con el que dan por supuesto quienes escriben las encyclopedias. Es un *trasfondo* tan obvio que apenas se nota. Según Lenat, para comprender un artículo sobre George Washington, por ejemplo, puede ser necesario saber hechos tales como que pasó por el Capitolio de los Estados Unidos en determinada ocasión, o las circunstancias en las que murió. Así, en 1985, este autor se lanzó a la tarea de congregar, durante los 10 años siguientes, este sentido común en “un único agente inteligente cuya base de conocimiento contenga (...) millones de entradas”.¹⁰

Hoy han pasado quince años desde entonces y, al menos, unos 15 millones de dólares se han invertido en el desarrollo de CYC, una base de datos de conocimientos orientada a que los ordenadores interpreten aspectos del sentido común como factores en la información, lo cual supondría un primer paso en la resolución del problema de las búsquedas. Para demostrar el uso de la CYC, Lenat ha desarrollado un sistema fotográfico de búsquedas que exemplifica el rol esencial que juega el conocimiento de sentido común. Tal sistema ha de encontrar imágenes *on-line* de acuerdo con ciertos títulos de búsqueda, no con los millones de imágenes que circulan en Internet, sino con la modesta cantidad de 20 fotografías. Sobre su experimentación ante tal demostración, un profesor de Stanford ha hecho los siguientes comentarios:

“La demostración del CYC ha sido llevada a cabo con 20 imágenes. La búsqueda solicitada, ‘alguien relajándose’ produjo una imagen de tres hombres en la playa con sus tablas de surf. CYC localizó esta imagen asociando el verbo ‘relajándose’ con los atributos previamente introducidos de la imagen. Tan sólo con 20 imágenes, el sistema no ha funcionado muy bien.”

V. Pratt (6 de abril de 1994). *CYC Report*. Stanford University.

Incluso al asumir que este sistema funciona, es necesario resaltar que lo hace única y exclusivamente porque los programadores del CYC han

9. Ver D. Lenat; R.V. Guha (1990). *Building Large Knowledge-Based Systems*. Nueva York: Addison Wesley.

10. *Ibid.*, pág. 4.

explicitado como *conocimiento* algunos de los significados generales de nuestra *comprensión* de la relajación, el ejercicio, el esfuerzo, etc., sobre la inevitable condición del cuerpo. Como casi todo nuestro entendimiento sobre lo que implica estar en un cuerpo es intuitivo, orientado por la inmediatez y las reacciones cotidianas, tenemos todas las razones para dudar que sea posible organizarlo correctamente en una base de datos computarizada.

Obviamente, en nuestra vida cotidiana, nuestra corporalidad no constituye mayor problema: podemos encontrar respuesta a inquietudes sobre nuestro cuerpo simplemente utilizándolo o imaginando lo que ocurriría si hicieramos algo en concreto. Así, por ejemplo, entendemos que los ejercicios abdominales no son relajantes simplemente al imaginarnos haciendo unas cuantas series. Luego, una fotografía de una persona haciendo ejercicios abdominales tendría que ser categorizada por los programadores de CYC dentro de “acciones de esfuerzo físico llevadas a cabo por alguien”. Sólo así CYC podría “deducir” que esta persona no está relajándose. Pero en general, podríamos generar un número infinito de preguntas necesarias o hechos relacionados a nuestra corporalidad, al cuerpo como tal. Tantas que resultaría imposible almacenarlas en un compendio explícito de conocimientos.

CYC no tiene cuerpo, y por tanto, como hemos visto, tendría que contemplar todos los hechos acerca de la corporalidad para poder encontrar la información en su base de datos. Aún siendo esto todavía posible, el programa no podría *comprender* cómo utilizar todos los hechos que ha registrado para responder a nuevas preguntas sobre el cuerpo. Si, por ejemplo, un usuario preguntara al CYC si es posible masticar chicle y silbar al mismo tiempo, éste no podría responder adecuadamente a pesar de tener toda la información sobre la masticación de chicle y el silbido por separado (al menos hasta que un programador humano y con cuerpo propio haga el trabajo de imaginar la respuesta y la añada a la base de respuestas del CYC). En cualquier caso, el número de hechos alrededor del cuerpo que tendríamos que hacer explícitos para integrar a la base de datos es, sencillamente, infinito, y afortunadamente quienes tenemos cuerpo no tenemos la necesidad de almacenarlo enteramente.

Falta por considerar que, incluso aunque fuese materialmente posible hacer explícito todo aquello que implica tener un cuerpo y organizarlo en la base de datos del CYC, nos quedaría por resolver el problema más

general de cómo mantener actualizados los datos de acuerdo al ritmo de cambios del mundo. Cualquier sitio web de título sencillo y directo que incluya las palabras “Bill Clinton” puede formar parte de una lista de temas de distinto interés, a su vez categorizables dentro de subgrupos, para un número indeterminado de usuarios, cuyas preocupaciones, en todo caso, cambiarían en la medida que las noticias van cambiando. Un día la política exterior puede ser el tema de interés principal, y al otro, un reportaje a alguna estrella. Visto que lo relevante de Clinton cambia con los días, sería menester diseñar un procedimiento que haga seguimiento y organice los cambios que ocurren a diario en el mundo, de modo que el ordenador pueda actualizarse según la adaptación permanente de contenidos en la organización global de los sitios web referidos a Bill Clinton.

En este sentido, está claro que es imposible determinar el número de cambios que ocurren en el mundo todo el tiempo: las fechas cambian, así como la formación de las nubes, como el peso de Clinton, su edad, lugar, puntos de vista, etc. Y, sin embargo, sólo unos pocos de estos cambios son suficientes para que a alguien, en algún momento, se le ocurra encontrar un sitio web sobre Clinton. Cualquier procedimiento para actualizar la información presentada en Internet tendría, por todo esto, que ignorar casi todos los cambios que acontecen en el mundo actual, y en la vida de Clinton, para tomar en cuenta sólo los más relevantes.

Si miles de millones de seres humanos responden sólo a los cambios que son relevantes en virtud de sus cuerpos e intereses, no es de asombrar que no se haya podido programar un ordenador para que responda certamente según criterios de relevancia. De hecho, el problema para reconocer sólo aquellos cambios que son relevantes, en un contexto determinado, ha sido uno de los mayores desafíos para la inteligencia artificial desde sus comienzos en los sesenta. Se denomina el “problema de marco” y permanece irresuelto hasta nuestros días.

Lenat ha comprendido que el problema de la relevancia amenaza todo su proyecto y se ha embarcado, como es lógico, en la empresa de reemplazar el sentido de relevancia basado en significados por formalismos lógicos, comenzando por proponer dos tipos de axiomas de relevancia: el *específico* y el *general*. Detrás de la idea del axioma específico se encuentra la premisa de que las diferentes secciones de la base de conocimientos “pueden ser organizadas por su relevancia para la resolución de problemas

explícitos”.¹¹ Así, por ejemplo, si a CYC se le ordena encontrar datos relevantes en el área del diseño de chips, el programa será guiado en su búsqueda por un axioma que indique que la sección de informática es más relevante que la de botánica (aunque esta última no necesariamente sea totalmente descartada, por cuanto podría servir como fuente para una o más analogías útiles).¹²

Y no sería sólo la analogía lo que concedería relevancia a ciertos hechos tentativamente alejados de la búsqueda. Consideremos el caso de una carrera de caballos en la que un apostador sabe que su jinete favorito sufre de alergia, se da cuenta de que le han asignado un carril cubierto de heno y se lanza a cambiar de apuesta. Proponernos que un programa interprete un cambio así de insignificante como relevante para la situación-problema del apostador nos coloca nuevamente frente al “problema de marco”. En efecto, cada conocimiento que tenemos podría asociarse a cualquier cosa de infinitas maneras posibles. En esos casos, sólo la comprensión de los significados presentes nos permitiría seleccionar lo que es relevante frente a una situación particular, de lo cual se deduce que estos tipos de axiomas *específicos*, en vez de ayudar a resolver, incrementan de modo dramático el problema de la relevancia general.

Frente a esta dificultad están los axiomas de relevancia *general*, que son formalizaciones de afirmaciones tales como “es necesario considerar solamente aquellos eventos cercanos en el tiempo a los eventos o proposiciones a tratar”.¹³ Esto no sólo nos devuelve al problema anterior del apostador, la alergia del jinete y la carrera, sino que además amenaza con introducir un número incontable de hechos en torno a las pistas de carrera que no harían sino complicar la pretensión de clasificar la relevancia. Guha y Levy, tratando de defender este axioma, sostienen que “es raro que un evento ocurra y que... [sea después] de un tiempo considerable... que repentinamente se manifiesten sus resultados”.¹⁴ Pero las promesas o casi todos los problemas de salud, por sólo nombrar dos ejemplos, cuentan justamente con esa característica: sus efectos podrían encontrarse, precisamente, en un futuro distante. Igualmente, toda clase de hechos históricos o

11. R. V. Guha; W. Pratt (1992). “Microtheories: An Ontological Engineer’s Guide”. *Reporte Técnico de la MCC* (número CYC-050-92, pág. 15).

12. Ibíd.

13. R. V. Guha; A.Y. Levy (1990). “A Relevance Based Meta Level”. *Reporte Técnico de la MCC* (número CYC-040-90, pág. 7).

14. Ibíd.

psicológicos relevantes para el presente inmediato de una persona están relacionados con su pasado más o menos lejano.¹⁵

Hace quince años, cuando Lenat se embarcó en este proyecto, pensó que diez años serían suficientes para que CYC fuese capaz de leer artículos en los periódicos y catalogar las novedades en la base de datos sin ayuda humana. Así son los sueños de quienes esperan que dispositivos de inteligencia artificial encuentren y provean a cada persona de la información que está interesada en encontrar. Pero en palabras de Michael Dertouzos, citadas como epígrafe al principio de este capítulo, estos cambios no han ocurrido. El mensaje es, como lo ha expresado Don Swanson, que “las máquinas no pueden reconocer el significado de las cosas, y por ello no pueden reproducir aquello que el juicio humano, en principio, incorpora dentro del proceso de clasificación e indexación de documentos”.¹⁶

El fracaso de los proyectos de inteligencia artificial, como en el caso de Lenat, debe orientar nuestra atención hacia la importancia de nuestros cuerpos a la hora de dar sentido y significado al mundo que nos rodea. Efectivamente, y como señalan los ejemplos anteriores, nuestra forma de vivir se organiza por y para la corporalidad que nos caracteriza: somos criaturas envueltas en cuerpos, con manos y pies, con interioridad y exterioridad, dotados para balancearnos en nuestro campo gravitacional, con mejor capacidad de movimiento hacia delante que hacia atrás; seres que se cansan y que alcanzan objetos con atravesar e intervenir el espacio que nos rodea, superando o evadiendo los obstáculos en el camino, etc. Nuestros intereses corporales se extienden en el mundo que nos rodea con tal amplitud que no notamos cómo construimos los significados del cuerpo mismo.¹⁷ Apenas podríamos caer en cuenta de ello en una situación de total desorientación, tal como, por ejemplo, aparecer en un mundo extraño, rodeados de criaturas de cuerpos radicalmente diferentes a los

15. En una conversación, Guha comentó que Lenat había descubierto que eran necesarios cientos de axiomas de relevancia y que, luego de llegar a miles, abandonaron la tarea porque parecía carecer de todo sentido. Y es que la necesidad de tantos axiomas conlleva la necesidad de criterios de mayor orden de relevancia que permitan discriminar, en un caso dado, axiomas más o menos relevantes, haciendo el trabajo infinito.

16. D. Swanson (antiguamente decano de la Escuela de Bibliotecología de la Universidad de Chicago) (1998). “Historical Note: Information Retrieval and the Future of an Illusion”. *Journal of the American Society for Information Science* (vol. 32. núm. 2, pág. 92-98).

17. El rol que juega el cuerpo en nuestra capacidad de experimentar el espacio, el tiempo y los objetos ha sido desarrollado en detalle en S. Todes (2001). *Body and World*. Cambridge, MA: MIT Press.

nuestros –seres esféricos o gaseosos, a modo ilustrativo– o, quizás, al observar la presencia de estas mismas criaturas en nuestro planeta.

Obviamente, sería de gran ayuda sistematizar nuestro sentido “encarnado” de lo relevante. Es indudable que ello sería útil en tanto base común para que otras personas con cuerpos e intereses similares a los nuestros encuentren información en las bases de datos o sitios web disponibles. No obstante, como demuestra la gestión de Lenat en su intención de explicitar y sistematizar el sentido común, no hay razón para esperar que sea posible formalizar la comprensión humana, ello en virtud de su fundamento corporal. Las esperanzas colocadas en la inteligencia artificial para resolver el problema de la relevancia han sido abandonadas de plano. El hecho ahora es que existe una vasta y creciente cantidad de información en el mundo, y a todas luces, nuestra única manera de acceder a ella es a través de ordenadores sin cuerpo, que no comparten nuestro mundo y que, por ende, no “entienden” el significado de nuestros documentos y sitios web.

Disociarnos de nuestro sentido común acerca del mundo, tal como requiere el uso de los ordenadores, implicaría hacer las cosas de la manera adecuada para la informática, esto es, intentar organizar lo que resulta relevante tamizando su semántica por los atajos del formalismo lógico. Existe, para ello, toda una industria dedicada a la gestión de búsqueda de la información, a producir programas y máquinas virtuales que aproximen el sentido de relevancia humano a través de la manipulación de símbolos, carentes de significado pero legibles para el ordenador.

En este sentido, los investigadores de la búsqueda de información también distinguen entre la *recuperación* y la *precisión*. En una situación ideal, un usuario encontraría, efectivamente, un 100 por ciento de documentos relevantes, esto es, encontraría todos los documentos que busca, y solamente los documentos que busca. La *recuperación* es el porcentaje de documentos relevantes para una búsqueda, mientras que la *precisión* es el porcentaje de documentos relevantes entre los encontrados. Ambas categorías no son independientes entre sí, de modo que el usuario se encuentra permanentemente en la difícil situación de tener que abandonar un criterio por el otro: al maximizar la *recuperación*, la *precisión* disminuye, y viceversa. En nuestros días, podemos afirmar que una búsqueda de 100 por ciento de *recuperación* y de *precisión* es prácticamente un ideal inalcanzable.

Ambos criterios se hacen más difíciles de maximizar en la medida en que el sistema se expande. Y dado el tamaño descomunal de Internet, se estima que las máquinas de búsqueda llegan a recuperar con precisión, en el mejor de los casos, el 2 por ciento de sitios relevantes. Blair explica por qué este dato tan interesante no ha sido observado hasta ahora:

“A pesar del tamaño y la dificultad de dar con contenidos precisos en las búsquedas por Internet, casi toda la publicidad que se ha hecho al WWW ha sido positiva. Un pionero del área, Don Swanson, notó este fenómeno hace décadas y lo llamo ‘la falacia de la abundancia’: un error que los usuarios cometen cuando emplean un sistema de búsqueda en superficies amplias y consiguen algunos documentos útiles. Swanson comentó que en un sistema suficientemente extenso... prácticamente cualquier búsqueda puede dar con documentos interesantes. El error está en pensar que, por dar con algunos resultados relevantes, el sistema sea lo suficientemente funcional, pues nunca se sabe cuántos documentos más apropiados quedaron fuera”.

D. Blair (Comunicación personal)

De hecho, la creencia actual en el desarrollo creciente de los sistemas de búsqueda hasta llegar al punto de que sean capaces de encontrar únicamente aquellos documentos requeridos por el usuario, sólo sería posible con una nueva taxonomía como la de Aristóteles o el sistema decimal de Dewey, que capture de algún modo cómo el mundo se divide en partes. Sin embargo, no hay solución metafísica que salve al mundo de los hipervínculos.

Las primeras máquinas de búsqueda simplemente creaban índices de palabras asociadas a una lista de documentos que las contenían, basados en puntajes sobre repeticiones y discriminando cuándo las palabras aparecían en el título, en el texto, la sinopsis, etc. Con todo, en opinión de los investigadores, estas técnicas tan sólo tienen un 10 por ciento de probabilidad de encontrar un documento de utilidad para una búsqueda.

Las llamadas máquinas de popularidad (*popularity engines*), analizan las actividades de millones de internautas, encontrando los sitios que más visitas reciben y el tiempo que emplean los usuarios en ellos, lo que permite conseguir los mejores resultados de búsqueda. Si bien esto ha incrementado la efectividad hasta un 20 por ciento, la lógica de asociar entradas y tiempo empleado en los sitios web como criterios para que los internautas encuentren lo que necesitan, presume que cualquier persona quedaría

satisfecha con los resultados de una búsqueda hecha por muchos otros usuarios. La satisfacción vendría dada, en este sentido, por la cantidad de entradas en un documento y el tiempo empleado en este.

Así, el problema vendría a ser de “similitud”. Todo es similar de infinitas maneras posibles –por ejemplo, este libro es similar a usted en cercanía al suelo, en estar ambos hechos de materia, reflejar la luz, acumular polvo, etc. Sucede que solamente notamos las similitudes que nos son relevantes en virtud de nuestra corporalidad e intereses, y por ello es imposible que los ordenadores puedan generar juicios útiles de similitud. No nos debe sorprender que el método de conteo de entradas funcione sólo para casos en que las búsquedas sean *idénticas* y no con peticiones que apunten a lo mismo expresadas de modo diferente. En este sentido, como comenta Gordon Rios, “los análisis sobre redes de datos de gran escala (de más de 100 millones de registros) demuestran que cerca de la mitad de las búsquedas son únicas y, por tanto, no cuentan con registros previos de entradas. Ello requeriría que se agreguen búsquedas ‘similares’ a cada gestión, lo cual nos devolvería directamente al origen del problema de la similitud.”

La última técnica desarrollada para buscar sitios relevantes sustituye el conteo de entradas por un análisis de los documentos encontrados. Empleando las anotaciones que los autores producen para vincular sus páginas a otras, las máquinas de búsqueda más avanzadas han mejorado la precisión en algunos aspectos. Sin embargo, el problema de la similitud vuelve a aparecer porque el espacio de posibilidades de búsqueda es mucho mayor que los textos concretos empleados para caracterizar los enlaces. Aparte, estas anotaciones abren el camino a quienes se dedican a enviar mensajes “basura” (*spammers*) y que, ya desde las máquinas de popularidad, venían aprovechándose de estos medios para mejorar el rango de acción de sus “actividades”.

El problema aquí está en que la medida de popularidad y los descriptores de vínculos tienden a ocultar páginas que, aunque sean menos visitadas, pueden perfectamente ser relevantes para el contexto de una búsqueda particular. Si, por ejemplo, buscamos documentos escritos por el famoso investigador Michael Jordan, algunas máquinas organizadas según número de entradas ignorarían por completo cualquier información que no tenga que ver con el jugador de baloncesto de igual nombre. Otras, diseñadas para usuarios más técnicos, seguramente nos resultarán

más funcionales, pero no serán de utilidad para quienes quieren saber del famoso jugador. De este modo, podríamos pensar que las especificaciones técnicas y los documentos insertos e instalados en las bases de datos de las máquinas son determinantes para el tipo de resultados accesibles al usuario.

El científico Gordon Rios, de Inktomi, el mayor proveedor de servicios de búsqueda por Internet, resume la situación actual como sigue:

“Hemos conducido estudios internos de gran escala que indican que, contemplando los principales criterios de búsqueda: textos, comportamiento de los internautas, estructura de enlaces y anotaciones, alcanzaríamos un 20-30 por ciento de precisión para búsquedas razonables. Y hemos intentado ir más allá generando modelos estadísticos complejos que operen simultáneamente con todas estas fuentes de información. Pero la mayoría en esta industria sabe que, al final de cada esfuerzo, terminamos dando cabezazos contra la pared... Es muy difícil producir lo que estos sistemas quieren lograr.”¹⁸

G. Rios (Comunicación personal)

Ese 30 por ciento es lo que podemos esperar de estos sistemas, y no debería sorprendernos. Hemos visto que el criterio de relevancia no puede operar sin contar con sentido común y que este último es imposible sin un fundamento “corporal” de relación con el mundo en el que estamos inmersos. Las técnicas de búsqueda de información en la web tienen una desventaja de partida: intentan aproximarse a la construcción del significado o el sentido de relevancia humana sin contar con un cuerpo y, por tanto, sin posibilidad de sentido común. No cabe esperar, por ende, que estos programas nos conduzcan por un camino al final del cual encontraremos la panacea computarizada del universo de información relevante, accesible y satisfactoria para los intereses de cualquier usuario. Aún menos cuando ni siquiera se ha logrado lo suficiente con los criterios de relevancia actuales.

18. Rios comenta, además, que “algunas máquinas potencian las búsquedas empleando criterios de proximidad. Así, no sólo se evalúa con cierta puntuación los documentos de mayores resultados de palabras encontradas, sino que se toma en cuenta la cercanía entre las palabras como factor relevante. Una vez más podríamos preguntar cuál es la manera correcta de evaluar la proximidad, si ésta varía según el contexto o el tema, claramente, de manera tal vez transparente para el usuario, pero opaca para el ordenador”. Este autor también destaca que muchas búsquedas (cerca de la mitad) sólo se introducen una vez y, de estos casos, posiblemente sea la mitad la que encuentre un resultado satisfactorio.

Don Swanson resume este punto en las siguientes palabras:

“La consistente indexación y búsqueda automatizada en forma totalmente efectiva no es posible. Nuestros juicios de relevancia... implican un conocimiento sobre quiénes somos, el tipo de mundo en que habitamos y el porqué de las búsquedas que hacemos. Es difícil imaginar un mecanismo... capaz de adquirir esta clase de autoconocimiento, o que funcione sin éste.”

D. Swanson (1998). “Historical Note: Information Retrieval and the Future of an Illusion”. *Journal of the American Society for Information Science* (vol. 32. núm. 2, págs. 92-98).

Considerando todo lo anterior, por un lado, el ciberespacio no cuenta con nuestras capacidades corporales para aprehender significados y, por otro, el sentido de relevancia se mantiene más allá de nuestro alcance virtual. Cabe preguntarse cómo es que los internautas encuentran información relevante para sus intereses en alguna medida. En este capítulo hemos sugerido que, desde la perspectiva humana, el mundo no es una masa volátil compuesta de billones de datos sin sentido global, sino más bien un campo de significados organizados por personas con cuerpo, deseos, intereses y propósitos. Con esto no pretendemos adentrarnos en el misterio de cómo nuestro cerebro se las arregla para sintonizar con elementos de nuestro interés en un momento dado. Quisiéramos al menos poner en evidencia que, en vista de que el mundo está organizado por agentes con cuerpo y no por ordenadores, es importante replantear nuestra manera de producir sentido a la hora de encontrar la información que queremos. Una cosa es segura, los internautas que cambian el mundo del cuerpo por el lenguaje sintáctico de la web y sus máquinas de búsqueda, tendrán que resignarse a escarbar entre montañas de documentos con la esperanza de encontrar, a veces, la información que desean.