
CONTESTACIÓN AL DISCURSO DE INGRESO DE D. JOSÉ JUAN PAZOS
ARIAS EN LA REAL ACADEMIA GALLEGA DE CIENCIAS

Fernando Pérez González

Sr. Presidente de la Real Academia Gallega de Ciencias,
Sr. Rector de la Universidad de Santiago,
Sr Rector de la Universidad de Vigo,
Autoridades, compañeros de la Academia, señoras e señores.

En su ilustrado y ameno discurso, el Profesor Pazos nos ha guiado por la selva de la información en que se ha convertido nuestro mundo. El abaratamiento del almacenamiento de los datos ha hecho que estos sigan creciendo incontroladamente hasta el punto de que una parte de la otrora llamada ‘aldea global’ es apenas visible por culpa de la maleza. Y este crecimiento desmesurado no ha hecho más que empezar, pues, como hemos escuchado, cada uno de nuestros futuros objetos no dejará de expectorar datos indolentemente. El saber sí ocupa lugar, pero el lugar no cuesta nada.

Adicción de datos, infoxicación, glotonería de datos, infosaturación, devoramiento de información, infobesidad, asfixia por datos, hambruna de tiempo, sobrecarga cognitiva... Son muchas las expresiones que se han ido acuñando para describir la ansiedad y los síndromes producidos por el exceso de información. A ellos, el profesor Pazos ha añadido certeramente el *síndrome de Diógenes* por la acumulación de información inútil. Trastornos estos que recuerdan al que sufrió el hidalgo Alonso Quijano, enajenado por la lectura de montañas de libros de caballerías.

Un reciente documental mostraba hasta dónde puede llegar la infoadicción en un ciudadano norteamericano, que ha decidido monitorizar cada segundo de su vida y emplea más de 700 sistemas de recogida de datos, 11 dispositivos vestibles y 71 redes WiFi, para buscar patrones en sus actividades. Almacena cada paso que da, cada bocado de comida que ingiere, todas las señales biológicas que puede captar, toma fotos automática y periódicamente de aquello que

tiene ante sí... Y este pobre gargantúa de los datos afirma desconsolado: “el problema de tener más datos es que nos hace ansiar más datos”.

De hecho, a otra escala, la angustia producida por la imposibilidad de aprehender todo el conocimiento existente no es nueva. El eclesiastés bíblico ya lamentaba que “no hay fin de hacer muchos libros” y fue Séneca padre quien escribió que “la abundancia de libros es distracción”. Leibniz, por su parte, se dolía de que “la horrible masa de libros siga creciendo”.

Los 44 zettabytes que contendrá Internet en 2020 son sólo una meta volante en el camino a la *Biblioteca de Babel* de Borges. Así como la inmensa mayoría de los libros de la Biblioteca no alcanza a tener sentido y sólo llena estantes interminables, también el diluvio de datos inservibles anega la red. Theodore Sturgeon, autor de ciencia ficción estadounidense, formuló su propio Principio de Pareto de la utilidad: “el 90% de todo es basura”. Sturgeon usó en realidad la palabra inglesa ‘crap’, que Uds. pueden traducir en formas más escatológicas.

Hay una página web llamada Petittube que muestra vídeos de Youtube que nadie ha visto nunca, en una sucesión aleatoria e inacabable de escombros visuales. Accediendo a un sitio denominado Forgetify, uno puede disfrutar de alguna de las más de 4 millones de canciones de Spotify que nadie ha escuchado jamás. El fenómeno no es exclusivo de Internet: un reciente estudio indicaba que, por término medio, cada estadounidense recibe en su hogar 189 canales de televisión por cable, pero sólo ve 17. Lo versificaba Bruce Springsteen en los años 80: “cincuenta y siete canales, y no ponen nada”.

Muy pertinentemente, el profesor Pazos ha disertado sobre la revolución que se avecina para extraer valor de este maremagnum. La envidiable simplicidad onomástica anglosajona ha comprimido en el término ‘big data’ (datos masivos) toda una colección de técnicas para el análisis y manipulación de conjuntos ciclópeos de datos heterogéneos. Muchas de las técnicas empleadas convivían ya bajo el paraguas de una disciplina conocida con el sugestivo nombre de “minería de datos”, porque de desenterrar gemas de información se trata. La diferencia es que ahora se procura extraer conocimiento de depósitos considerablemente mayores de datos: hay más gemas, pero también mucha más tierra que remover. Borges afirmaba que en algún tomo de su biblioteca estarían “las autobiografías de los arcángeles y la relación verídica de [... la propia] muerte”, pero, para encontrarlas, harían falta técnicas eficientes, capaces de

gestionar y buscar en la colosal biblioteca sin tener que hojear todos y cada uno de los volúmenes.

Son numerosísimos los ejemplos recientes de éxitos de los 'datos masivos' en áreas como la medicina, la banca, los seguros, las telecomunicaciones, el comercio minorista, la seguridad, la industria o la logística. Mi favorito es un sistema denominado Watson y desarrollado por IBM para participar en *Jeopardy*, un veterano y popular concurso de la televisión americana. Jeopardy invierte el tradicional concurso de preguntas y respuestas: en lugar de formular cuestiones a los concursantes, va dando pistas sobre la respuesta y son los contendientes quienes deben adivinar cuál es la pregunta. Para retorcerlo más, los nombres de las categorías y las propias pistas están llenos de dobles sentidos, juegos de palabras e ironías. A pesar de todo ello, Watson fue capaz de vencer en 2011 a dos de los mejores concursantes de todos los tiempos. Para lograr semejante hazaña, Watson combinó tecnologías como el procesado del lenguaje natural, el aprendizaje máquina o la representación del conocimiento, con hardware de procesamiento masivo en paralelo. Mientras aprendía, Watson engulló un total de 15 terabytes de conocimiento humano. Con todo, sigue deslumbrando su desempeño en la gestión de la información contextual: la categoría, el tipo de contestación pedido, y el tiempo, lugar y género implícitos en cada pista. Se preguntarán Uds. si este poderoso ingenio no ha sido empleado para propósitos más provechosos: efectivamente, el Memorial Sloan Kettering Cancer Center lo utiliza como apoyo para la toma de decisiones en el tratamiento del cáncer de pulmón.

En la segunda parte de su discurso, el Prof. Pazos ha expuesto admirablemente las bases tecnológicas de los sistemas recomendadores, materia en la que con su grupo de investigación del Departamento de Ingeniería Telemática de la Universidad de Vigo ha adquirido una merecida reputación internacional. Podríamos pensar que los recomendadores son una invención moderna, pero antes hemos visto cómo egregios pensadores lamentaban el crecimiento desmesurado en la literatura disponible, por lo que como mínimo el problema no es nuevo. Una afortunada casualidad hace que el pazo en el que estamos se encuentre relacionado con los recomendadores de antaño. En efecto, en este edificio tiene su sede el Instituto de Estudios Gallegos Padre Sarmiento, dedicado al benedictino Martín Sarmiento, insigne figura de la Ilustración española.

El padre Sarmiento era consciente, a mediados del siglo XVIII, de un hecho clave para los bibliófilos: tras la invención de la imprenta, los libros se habían, como diríamos ahora, “comoditizado”, esto es, no sólo había infinidad de volúmenes publicados, sino que muchos se habían convertido en objetos comunes. Así pues, a petición del marqués de Aranda, Sarmiento preparó su “*Catálogo de algunos libros curiosos y selectos para una librería de un particular que desee componer de tres a quatro mil thomos*”. Decía el fraile: “desde la invención de la imprenta hasta el tiempo presente, se ha escrito infinito, y hay mucho óptimo, mucho mediano, y mucho más pésimo y despreciable. Por esto, es más difícil la elección entre todo”. En definitiva, el *Catálogo de libros curiosos y selectos* no es más que un recomendador de libros para lectores interesados en salirse del *mainstream*. No muy diferente en espíritu de los sistemas recomendadores de música alternativa, de películas *indie* o de restaurantes étnicos de hoy en día.

La diferencia esencial entre el recomendador de Sarmiento y los que nos ha descrito el profesor Pazos es la personalización automática. Si el fraile diera con otro marqués demandante de un catálogo, habría de alternar la soledad de su celda y las visitas a su vasta biblioteca, hasta compilar una lista del agrado de su patrocinador, proceso éste lento y fatigoso. Los recomendadores modernos comparten la necesidad de modelar y capturar los intereses del cliente, pero, además, pueden aprender de las interacciones pasadas del usuario con otros sistemas y también a medida que reciben realimentación, sea implícita o explícita. Como hemos escuchado, la realimentación explícita es un poco más molesta para el usuario, pero más atinada. Mi experiencia personal refrenda este juicio: comencé a utilizar Netflix hace unos años cuando vivía en EE.UU. Netflix es un servicio de vídeo bajo demanda de enorme popularidad en Norteamérica y recientemente desembarcado en nuestro país. Pues bien, desde que decidí puntuar cada película que veía, las recomendaciones mejoraron palpablemente. Aprendida la lección como usuario, ahora soy explícito en mis valoraciones en otros de los recomendadores que uso, como el de Amazon para libros o Tripadvisor para restaurantes.

Otra característica de los recomendadores automáticos es la adaptación al contexto, porque también somos nuestras circunstancias. Por último, el elemento que probablemente más ha tenido que ver en la popularidad de los recomendadores modernos, es la existencia de algoritmos de filtrado, esto es, de adecuación de contenidos a los perfiles de los usuarios. De entre los que el Prof. Pazos ha citado, en mi opinión es el llamado “filtrado colaborativo” no sólo

el más exitoso, sino también el que mejor aprovecha la dimensión social de Internet. El recomendador de Netflix identifica de entre sus millones de usuarios aquellos con intereses similares a los nuestros y nos propone películas que ellos han valorado favorablemente. Así pues, el tino del recomendador de Netflix es también una cura para el ego: por muy singulares que nos sintamos, nuestros gustos concuerdan con los de decenas de miles de habitantes del planeta.

En la última parte de su discurso, D. José Juan Pazos nos ha guiado por algunos de los sistemas recomendadores que él y su equipo han desarrollado, desde el más conocido de la televisión, pasando por un tutor inteligente para plataformas de teleenseñanza, por la publicidad personalizada y los contenidos y destinos turísticos, hasta llegar a un novedoso recomendador para tutores de educación especial que genera automáticamente planes terapéuticos y que demuestra que nuestro nuevo miembro de la Academia no sólo es un excelente investigador, sino también una excelente persona.

Los miembros que somos de la Real Academia Gallega de Ciencias también hacemos de recomendadores cuando llega el momento de proponer nuevas incorporaciones. Imaginarán Uds.-- y acertarán--que, por fortuna, y como diría Sarmiento, "hay mucho óptimo" y, por desgracia, poco espacio en nuestras estanterías. D. José Juan Pazos accede hoy a nuestra institución por su brillante trayectoria como investigador, inicialmente en el campo de las metodologías formales de diseño y desarrollo de software y, posteriormente, en la personalización, recomendación y adaptación de contenidos mediante tecnologías semánticas, actividades financiadas por un buen número de proyectos de diferentes convocatorias autonómicas, nacionales y europeas, y contratos con instituciones y empresas. Como resultado de estos trabajos ha publicado más de 200 artículos en revistas y actas de congresos, que han sido copiosamente citados. Además de Subdirector de la Escuela de Ingenieros de Telecomunicación en el periodo 1994-1998, ha sido Director del Departamento de Ingeniería Telemática de 2001 hasta 2004. Y es que D. José Juan ha sido siempre capaz de encontrar tiempo para investigar a pesar de las muchas veces ingratas tareas de gestión; las últimas, como gestor del programa de Fomento de la Investigación e Innovación Empresarial de la Dirección Xeral de I+D+i y desde 2015 como colaborador de la Agencia Nacional de Evaluación y Prospectiva. Por si no fuera bastante, ha sido autor de dos libros de texto que forman parte de la bibliografía recomendada en numerosas universidades españolas e iberoamericanas. No

es casualidad que el Prof. Pazos recibiese el premio al mejor profesor de segundo ciclo de la titulación de Ingeniería de Telecomunicación.

Pero de que es un magnífico docente puedo dar fe porque un servidor formaba parte del puñado de supervivientes de la primera promoción de alumnos de Ingeniería de Telecomunicación de Vigo que tuvo en cuarto curso la fortuna de recibir las primeras clases impartidas por el Prof. Pazos, recién aceptada la oferta de dejar la empresa de Madrid en que trabajaba. Fueron aquellos tiempos de grata solidaridad entre estudiantes y profesores de casi su misma edad. Clases de teletráfico de D. José Juan que yo, inclinado hacia las matemáticas aplicadas, disfrutaba especialmente, maravillado de que la teoría de colas ofreciese soluciones cerradas a problemas de dimensionamiento de recursos limitados, ya fuese el número de surtidores en una gasolinera, ya el número de routers en una red informática.

Nacido en Baiona, padre de dos hijos, es aficionado a la vela y el senderismo. Le gusta viajar poniendo por delante cultura y gastronomía a sol y playa. Lector en su tiempo libre, prefiere evadirse de su profesión con ensayos sobre economía, historia y filosofía, por lo que es posible que, sin saberlo, nos hayamos recomendado libros mutuamente a través de Amazon. Lo que sí nos hemos recomendado, y en grandes dosis, es resignación frente a la displicencia de aquellos que deberían apostar por la Ciencia, pero se contentan con los fuegos de artificio. Decía precisamente Sarmiento que “si la joya que antiguamente se daba en Barcelona al que mejor discurriese sobre propuestos asuntos del arte de trovar, se hubiera determinado que algunas veces se diese al que mejor escribiese sobre puntos de Matemáticas, Física experimental, Náutica, Agricultura o Maquinaria no hubiera reinado tanto tiempo la barbarie”.

D. José Juan: reciba mi afectuosa enhorabuena por este ingreso. La Academia se honra de contar con Ud. entre sus miembros, sabedora de que pondrá todo su empeño y resolución en la consecución de su fin último, que no es otro que el desarrollo de Galicia mediante el conocimiento.

Bienvenido a esta Academia.