

El Proyecto Gutenberg (1971-2009)

EL PROYECTO GUTENBERG (1971-2009)

MARIE LEBERT

Traducido por ANNA ÁLVAREZ y revisado por Marie Lebert

NEF, Universidad de Toronto, 2009

Copyright © 2009 Marie Lebert

RESUMEN

Fundado por Michael Hart en julio de 1971 cuando estudiaba en la Universidad de Illinois (Estados Unidos), el Proyecto Gutenberg se propone difundir gratuitamente por vía electrónica el mayor número posible de obras literarias. Michael escribió: "Consideramos el texto electrónico como un nuevo medio de comunicación, sin verdadera relación con el papel. La única semejanza es que distribuimos las mismas obras, pero en cuanto la gente se haya acostumbrado, no veo cómo el papel aún podría competir con el texto electrónico, sobre todo en las escuelas." Al generalizarse el uso de la web en los años 90, el Proyecto Gutenberg cobra nueva vida y alcanza una difusión internacional. A lo largo de los años, miles de voluntarios fueron digitalizando con paciencia cientos de obras en modo texto. Primero las colecciones eran esencialmente anglófonas, pero con el tiempo se fueron haciendo multilingües, con sitios web ubicados en varios países (Estados Unidos, Canadá, Europa y Australia) sin contar los que quedan por crear.

UNA APUESTA LANZADA HACE 38 AÑOS

Gestación

¿Cuáles son las raíces del proyecto? Cuando Michael Hart estudiaba en la Universidad de Illinois, en los Estados Unidos, el laboratorio informático (Materials Research Lab) de su universidad le asignó 100 millones de dólares de tiempo de ordenador. El 4 de julio de 1971, día de la fiesta nacional, Michael digitó en el teclado de su ordenador "The United States Declaration of Independence" (Declaración de independencia de los Estados Unidos, firmada el 4 de julio de 1776), en mayúsculas, pues las letras minúsculas aún no existían. El texto electrónico representaba 5 K (kilobytes).

Pero el envío de un archivo de 5 K a las cien personas que formaban la red de la época hubiera acabado con el sistema, provocando la implosión de la red. Por aquel tiempo el ancho de banda aún era ínfimo. Michael difundió un mensaje que explicaba dónde se almacenaba el texto (que no contenía aún ningún hipervínculo, pues habría que esperar unos 20 años más para la web). Poco después, seis personas descargaron este archivo.

Sobre la marcha, Michael decidió dedicar ese crédito de tiempo de unos millones de dólares a la búsqueda de obras del dominio público que estuvieran disponibles en bibliotecas, y a la digitalización de éstas. Optó por almacenar los textos electrónicos bajo la forma más sencilla posible, en formato ASCII, para que esos textos pudieran ser leídos sin problema con cualquier ordenador, plataforma y programa. El conjunto de páginas encuadradas, forma tradicional del libro, se convierte en un texto electrónico que se puede desplegar de par en par, con letras capitales para los términos escritos en cursiva o en negrita y los términos subrayados en la versión impresa.

Poco después, Michael definió así la misión del Proyecto Gutenberg: poner a disposición de todos, por vía electrónica, el mayor número posible de obras del dominio público: "Nosotros consideramos el texto electrónico como un nuevo medio de comunicación, sin verdadera relación con el papel", explicó más

tarde, en agosto de 1998. "La única semejanza es que ambos difundimos las mismas obras, pero en cuanto lagente se haya acostumbrado, no veo cómo el papel aún podría competir con el texto electrónico, sobre todo en las escuelas."

Tras haber digitado "The United States Declaration of Independence" en 1971, Michael prosiguió con sus esfuerzos en 1972 digitando "The United States Bill of Rights" (Declaración de derechos de los Estados Unidos). Esa declaración incluye las diez primeras enmiendas añadidas en 1789 a la Constitución de los Estados Unidos (ratificada en 1787), y define los derechos individuales de los ciudadanos y los poderes respectivos del gobierno federal y de los Estados. En 1973, un voluntario digitó "The United States Constitution" (Constitución de los Estados Unidos) en su totalidad.

Perseverencia

Año tras año, la capacidad del disquete aumenta con regularidad – el disco duro no existe aún – de modo que los archivos pueden ocupar cada vez más espacio. Algunos voluntarios se animan a digitalizar la Biblia, compuesta de varios libros, cada cual puede tratarse por separado y ocupar un archivo diferente. Michael también empieza a digitar la obra completa de Shakespeare, con la ayuda de voluntarios, una obra de teatro tras otra, cada una en un archivo. De hecho aquella edición nunca se ha podido poner en línea, debido a que entre tanto, había entrado en vigor una ley de copyright más rigurosa, ya no destinada a proteger el texto de Shakespeare, pasado desde hace tiempo al dominio público, sino los comentarios y notas de aquella edición. Otras ediciones anotadas que sí habían pasado al dominio público se pusieron en línea algunos años más tarde.

Por las mismas fechas, el internet, aún embrionario en 1971, despegó verdaderamente en 1974, tras la creación del protocolo TCP/IP (transmission control protocol/internet protocol). En 1983, le red está en plena expansión.

De 10 a 1.000 ebooks

En agosto de 1989, el Proyecto Gutenberg pone en línea su décimo texto,

"The King James Bible", publicada por primera vez en 1611 y cuya

versión más conocida es la de 1769. El conjunto de los archivos del

Antiguo Testamento y del Nuevo Testamento representa 5 M (megabytes).

En 1990, ya hay 250.000 internautas, y el modelo estándar vigente es el disquete de 360 Ko. En enero de 1991, Michael digita "Alice's Adventures in Wonderland" (Alicia en el país de las maravillas) de Lewis Carroll (publicado en 1865). En julio del mismo año, digita "Peter Pan" de James M. Barrie (publicado en 1904). Cada uno de estos dos clásicos de la literatura infantil cabe en un disquete estándar.

Luego llega la web, operativa a partir de 1991. El primer navegador, Mosaic, aparece en noviembre de 1993. Al generalizarse el acceso a la red, resulta más fácil hacer circular los textos electrónicos y reclutar voluntarios. El Proyecto Gutenberg perfecciona su método de trabajo, y logra digitalizar un texto al mes en 1991, dos textos al mes en 1992, cuatro textos al mes en 1993 y ocho textos al mes en 1994. En enero de 1994, el Proyecto Gutenberg celebra su centésimo libro con la puesta en línea de "The Complete Works of William Shakespeare" (La obra completa de William Shakespeare). Shakespeare escribió la mayor parte de su obra entre 1590 y 1613. A continuación la producción sigue aumentando, con una media de 8 textos al mes en 1994, 16 textos al mes en 1995 y 32 textos al mes en 1996.

Como lo podemos observar, entre 1991 y 1996, las producciones se han duplicado cada año. Mientras sigue con la digitalización de los libros, Michael coordina también a partir de entonces el trabajo de decenas de voluntarios. A partir de finales de 1993, el Proyecto Gutenberg se articula en torno a tres grandes sectores: (a) "Light Literature" (literatura de entretenimiento), que incluye por ejemplo "Alice's Adventures in Wonderland" (Las aventuras de Alicia en el país de las maravillas), "Peter Pan" o "Aesop's Fables" (Las Fábulas de Esopo); (b) "Heavy Literature" (literatura "seria"), que incluye por ejemplo la Biblia, las obras de Shakespeare o "Moby Dick"; (c) "Reference Literature" (literatura de referencia), que reúne enciclopedias y diccionarios, por ejemplo el "Roget's Thesaurus". Más adelante, esta presentación en tres sectores es sustituida por una clasificación detallada con más secciones.

El Proyecto Gutenberg pretende ser universal, tanto a nivel de las obras que selecciona como del público al que se dirige, ya que la meta es poner la literatura a disposición de todos, sin limitarse al público habitual de estudiantes y docentes.

El sector dedicado a la literatura de entretenimiento se propone atraer frente a la pantalla a un público muy diverso, por ejemplo a niños y a sus abuelos en busca del texto electrónico de "Peter Pan" después de haber visto la película Hook, de la versión electrónica de "Alicia en el país de las maravillas" tras haber visto una adaptación televisiva, o del origen de una cita literaria tras visionar un episodio de Star Trek. En casi todos los episodios de Star Trek se citan libros disponibles en versión digital en el Proyecto Gutenberg.

El objetivo, entonces, es que el público, esté o no esté acostumbrado al contacto con el libro impreso, pueda encontrar con facilidad los textos escuchados en conversaciones, películas, músicas, o leídos en otros libros, periódicos y revistas. Gracias al uso del formato ASCII, los archivos electrónicos no ocupan mucho espacio. Se pueden descargar fácilmente a través de la línea telefónica. La búsqueda textual es igual de sencilla. Basta con utilizar la función "buscar" presente en cualquier programa.

En 1997, la producción sigue siendo de 32 títulos al mes. En junio de 1997, el Proyecto Gutenberg pone en línea "The Merry Adventures of Robin Hood" (Las alegres aventuras de Robin Hood) de Howard Pyle (publicado en 1883). En agosto de 1997, pone en línea su milésimo texto electrónico, "La Divina Commedia" (La Divina Comedia) de Dante Alighieri (publicada en 1321), en su idioma original, el italiano. Michael Hart escribió en agosto de 1998: "Mi proyecto es cargar 10.000 textos electrónicos en el internet. Si pudiera conseguir subvenciones más importantes, me gustaría llegar hasta un millón y ampliar también el número de nuestros usuarios potenciales pasando de un 1, x% a un 10% de la población mundial, lo que representaría la distribución de 1.000 veces un billón de textos electrónicos en lugar de un solo billón."

De 1.000 a 10.000 ebooks

Entre 1998 y 2000, la media sigue constante, con 36 textos al mes. En mayo de 1999, las colecciones cuentan con 2.000 libros. El 2000° texto es "Don Quijote" (publicado en 1605), en castellano.

Disponible en diciembre del año 2000, el 3.000° título es el tercer volumen de "A l'ombre des jeunes filles en fleur (A la sombra de las muchachas en flor) de Marcel Proust (publicado en 1919), en su idioma original, en francés. La media pasa a 104 libros al mes en 2001.

Puesto en línea en octubre de 2001, el 4.000° texto es "The French Immortals Series" (Obras de los Inmortales franceses), en versión inglesa. Publicado en 1905 por la Maison Mazarin (París), este libro reúne varias ficciones de escritores laureados por la Academia de la lengua francesa (Académie française), como Emile Souvestre, Pierre Loti, Hector Malot, Charles de Bernard, Alphonse Daudet y otros.

Disponible en abril de 2002, el 5.000° texto es "The Notebooks of Leonardo da Vinci" (Los cuadernos de Leonardo da Vinci), que datan de inicios de siglo XVI. En 2008, este texto aún se encuentra en el Top 100 de los libros descargados con más frecuencia.

En 1988, Michael Hart decidió digitalizar "Alice's Adventures in Wonderland" (Las aventuras de Alicia en el país de las maravillas) o "Peter Pan" porque su versión digital cupo en un disquete de 360 Ko, o sea el estándar de la época. Quince años más tarde, en 2002, ya disponemos de disquetes de 1,44 M (megabytes) y resulta muy sencillo comprimir archivos en archivos ZIP. Un archivo estándar puede contener tres millones de caracteres, o sea más de los necesarios para un libro de tamaño mediano, ya que una novela de 300 páginas digitalizada en formato ASCII representa 1 M (megabyte). Un libro voluminoso cabe en dos archivos ASCII, descargables en versión normal o en versión comprimida. Se necesitan unas cincuenta horas para seleccionar un libro de tamaño mediano, comprobar que ha pasado al dominio público, escanearlo, corregirlo, formatearlo y compaginarlo.

Algunos números de libros se reservan para más tarde, por ejemplo el número 1984 (eBook #1984) para la novela epónima de George Orwell, publicada en 1949, que aún está lejos de pasar al dominio público.

En 2002, las colecciones se amplían de 203 títulos al mes. Al llegar la primavera de 2002, ya cubren un cuarto de las obras del dominio público de libre acceso en la web, inventariadas de forma casi exhaustiva por la Internet Public Library (IPL). Un buen resultado, fruto del trabajo paciente de miles de voluntarios activos en numerosos países.

1.000 libros en agosto de 1997, 2.000 libros en mayo de 1999, 3.000 libros en diciembre de 2000, 4.000 libros en octubre de 2001, 5.000 libros en abril de 2002, 10.000 libros en octubre de 2003. El 10.000° libro es "The Magna Carta", que fue el primer texto constitucional inglés, firmado en 1215.

Entre abril de 2002 y octubre de 2003, las colecciones se duplican, pasando de 5.000 a 10.000 libros en dieciocho meses. En 2003, se digitalizan de promedio unos 348 libros al mes.

Diez mil libros: una cifra impresionante si uno piensa en lo que aquello representa en términos de páginas escaneadas, releídas y corregidas. Este crecimiento rápido se debe a la actividad de Distributed Proofreaders

(DP), un sitio web concebido en 2000 por Charles Franks para repartir la corrección de las obras entre varias personas. Los voluntarios eligen uno de los libros cuya digitalización ya ha empezado para releer y corregir determinada página. Cada cual trabaja a su ritmo. A título indicativo, en el sitio web se aconseja repasar una página al día. Representa poco tiempo en un día, pero mucho para el proyecto. En agosto de 2003, se pone a disposición un CD "Best of Gutenberg" con una selección de 600 libros. En diciembre de 2003, fecha en la que el Proyecto Gutenberg alcanza la cifra de 10.000 libros, se graba en un DVD la casi totalidad de los libros (9.400 libros). Se envían gratuitamente el CD y el DVD a cualquier persona que lo pida. Luego cada uno queda libre de realizar todas las copias que pueda y de distribuirlas a su alrededor.

De 10.000 a 20.000 ebooks

En diciembre de 2003, las colecciones rondan los 11.000 libros. Varios formatos están representados (HTML, XML, RTF, etc.) pero el formato principal - y obligatorio - sigue siendo el ASCII. Hay 46.000 archivos en conjunto, lo que equivale a una capacidad total de 110 G (gigabytes). El 13 de febrero de 2004, fecha en la que Michael Hart da una conferencia en la sede de la UNESCO en París, las colecciones abarcan exactamente 11.340 libros en 25 idiomas diferentes. En mayo de 2004, los 12.500 libros disponibles representan 100.000 archivos en veinte formatos diferentes, lo que equivale a una capacidad total de 135 G (gigabytes), y ésta debería duplicarse cada año con la agregación de más de 300 libros al mes (338 libros en 2004).

Al mismo tiempo, el Project Gutenberg Consortia Center (PGCC), lanzado en 1997 para reunir colecciones de libros digitales en un punto de acceso único, es afiliado oficialmente al Proyecto Gutenberg en 2003. En otro ámbito, se inicia un proyecto europeo a iniciativa del Proyecto Rastko, basado en Belgrado (Serbia). Distributed Proofreaders Europe nace en diciembre de 2003, y el Proyecto Gutenberg Europa en enero de 2004, con cien libros disponibles para abril de 2005. Aparecen libros en varios idiomas con el fin de reflejar la diversidad lingüística que prevalece en Europa, y a largo plazo la intención es proponer documentos en cien idiomas.

En enero de 2005, el Proyecto Gutenberg celebra sus 15.000 libros, al poner en línea "The Life of Reason" de George Santayana (publicado en 1906).

En junio de 2005, se alcanzan los 16.000 libros. Si bien en febrero de 2004 sólo había textos en 25 idiomas, en julio de 2005 están representados 42 idiomas, entre los cuales el sánscrito y las lenguas maya. En diciembre de 2006 figuran 50 idiomas. El 16 de diciembre de 2006, varios idiomas cuentan con colecciones de más de 50 títulos: el inglés (17.377 libros), el francés (966 títulos), el alemán (412 títulos), el finlandés (344 títulos), el holandés (244 títulos), el español (140 títulos), el italiano (102 títulos), el chino (69 títulos), el portugués (68 títulos) y el tagalog (51 títulos).

Lanzado en agosto de 2001, el Proyecto Gutenberg of Australia celebra sus 500 libros en julio de 2005. En la misma fecha, el Proyecto Gutenberg Canadá aún está en gestación, así como un Proyecto Gutenberg en Portugal y uno en las Filipinas.

En diciembre de 2006, el Proyecto Gutenberg alcanza los 20.000 libros. El 20.000º libro es un libro audio, "Twenty Thousand Leagues Under the Sea" (Veinte mil leguas de viaje submarino), la versión inglesa de la novela de Julio Verne (publicada en 1869). En 2006, se digitalizan de promedio unos 345 libros nuevos al mes. Aunque se hayan necesitado 32 años para digitalizar los 10.000 primeros libros – de julio de 1971 a octubre de 2003 – bastará con sólo tres años y dos meses, de octubre de 2003 a diciembre de 2006, para digitalizar los 10.000 libros siguientes.

Por la misma fecha, el Proyecto Gutenberg of Australia ronda los 1.500 libros (se alcanzará esta cifra en abril de 2007) y el Proyecto Gutenberg Europa cuenta con 400 libros.

En enero de 2006 se crea la sección Project Gutenberg PrePrints para recoger nuevos documentos que por su interés merecerían estar en línea, pero que no pueden ser incorporados a las colecciones existentes sin procesamiento previo por voluntarios: colecciones incompletas, calidad insuficiente, necesidad de conversión a otro formato, etc. Esta sección comprende 379 títulos en diciembre de 2006, y más de 2.000 títulos dos años más tarde.

Decenas de miles de ebooks

Se estrena el sitio web Project Gutenberg News a principios de noviembre de 2006 a instigación de Mike Cook. Este sitio es complementario al de la carta de información semanal y mensual que ya llevaba muchos años de existencia. El sitio presenta por ejemplo las estadísticas de producción semanales, mensuales y anuales desde 2001. La producción semanal se eleva a 24 libros en 2001, 47 libros en 2002, 79 libros en 2003, 78 libros en 2004, 58 libros en 2005, 80 libros en 2006 y 78 libros en 2007. La producción mensual es de 104 libros en 2001, 203 libros en 2002, 348 libros en 2003, 338 libros en 2004, 252 libros en 2005, 345

Thank You for previewing this eBook

You can read the full version of this eBook in different formats:

- HTML (Free /Available to everyone)
- PDF / TXT (Available to V.I.P. members. Free Standard members can access up to 5 PDF/TXT eBooks per month each month)
- Epub & Mobipocket (Exclusive to V.I.P. members)

To download this full book, simply select the format you desire below

